Carnegie Mellon University

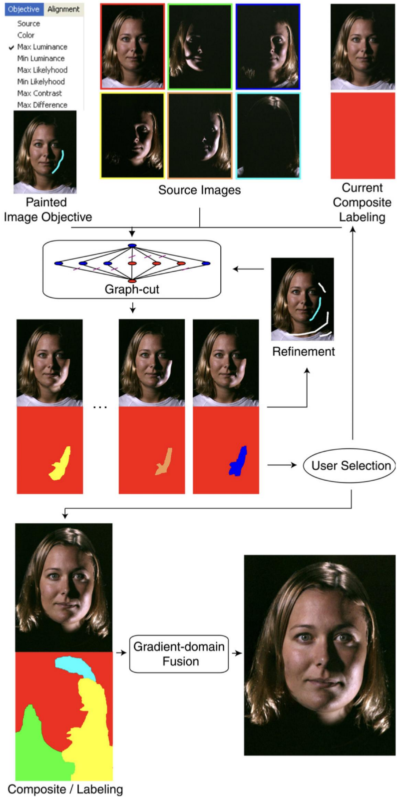
Interactive Digital Photomontage - Agarwala et. al, SIGGRAPH 2004

Presented by:
Rohan Rao (rgrao)
Harsh Sharma (hsharma2)

High-level overview of photomontage framework

- Input: set of source images, **an image stack**. Usually formed with fixed camera pose, but can align images using Lucas-Kanade alignment or other methods
- Select an image objective, apply it globally or locally: luminance, contrast, brightness, etc
- Select a secondary seam objective: colors vs colors+edges vs colors+gradients
- Select a relative weighting for these objectives - can create multiple small regions with more seams (low seam objective) or larger regions with less seams (high seam objective)
- Graph-cut optimization by balancing the two types of objectives with relative weighting to identify the best seams and cuts across the image stack
- Gradient domain fusion of multiple composites for seamless blending (Optional)

Framework



https://youtu.be/Rp7uDRdQRFc

Image Objectives

1. **Designated color:** a specific desired color to either match or avoid
2. **Minimum or maximum luminance:** the darkest or lightest pixel in the span.
3. **Minimum or maximum contrast:** the pixel from the span with the lowest or highest local contrast in the span
4. **Minimum or maximum likelihood:** the least or most common pixel value in the span (subject to a particular histogram quantization function).
5. **Eraser:** the color most different from that of the current composite.
6. **Minimum or maximum difference:** the color least or most similar to the color at position p of a specific source image in the stack

Graph Cut Optimization

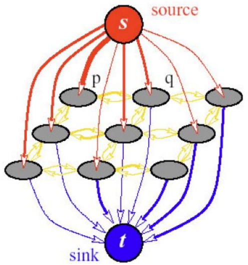
- Suppose we have n source images . To form a composite, we must choose a source image S_i for each pixel p.
- Label is the mapping between pixel and source image denoted by L(p).
- Seam exists between two neighboring pixels p,q in the composite if L(p) = L(q).
- Provide hard constraints with user sketches, soft constraints with seams
- Optimization -
  - Outer loop - across all labels
  - Inner loop - For a label α, optimize the following cost function

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

Data penalty - Distance to the image objective

Interaction penalty - Distance to the seam objective

  - After t_{th} iteration,$L_{t+1}(p) = L_t(p)$ or $L_{t+1}(p$



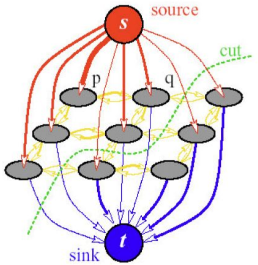


Image Objective penalty (C_d) definitions

1. **Designated color**(most or least similar): the Euclidean distance in RGB space of the source image pixel $S_{L(p)}(p)$ from a user-specified target color. We supply a user interface for the selection of a pixel in the span that is used as the color target.
2. **Minimum or maximum luminance:** the distance in luminance from the minimum (maximum) luminance pixel in a pixels span.
3. **Designated image:** 0 if L(p) = u, where S_u is a user-specified source image, and a large penalty otherwise.
4. **Minimum or maximum likelihood:** the least or most common pixel value in the span (subject to a particular histogram quantization function).
5. **Eraser**: the Euclidean distance in RGB space of the source image pixel S L(p) (p) from the current composite color.
6. **Minimum or maximum difference**: the Euclidean distance in RGB space of the source image pixel $S_{L(p)}(p)$ from $S_u(p)$, where S u is a user-specified source image.

Carnegie Mellon University

Seam Objective penalty (C_i)

Seam objective is 0 if L(p) = L(q).

Otherwise:

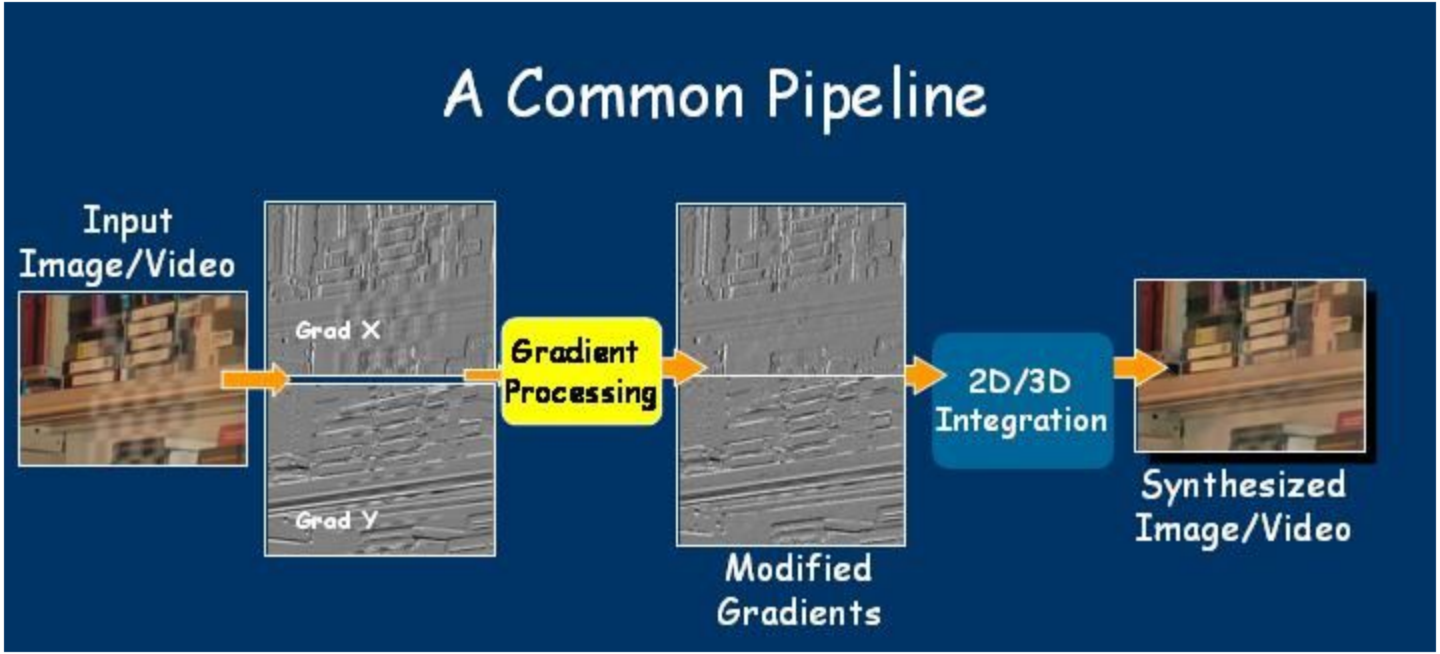
$$C_i(p,q,L(p),L(q)) = \begin{cases} X & \text{if matching "colors"} \\ Y & \text{if matching "gradients"} \\ X+Y & \text{if matching "colors \& gradients"} \\ X/Z & \text{if matching "colors \& edges"} \end{cases}$$

where

$$\begin{aligned} X &= \|S_{L(p)}(p) - S_{L(q)}(p)\| + \|S_{L(p)}(q) - S_{L(q)}(q)\| \\ Y &= \|\nabla S_{L(p)}(p) - \nabla S_{L(q)}(p)\| + \|\nabla S_{L(p)}(q) - \nabla S_{L(q)}(q)\| \\ Z &= E_{L(p)}(p,q) + E_{L(q)}(p,q)) \end{aligned}$$

Carnegie Mellon University

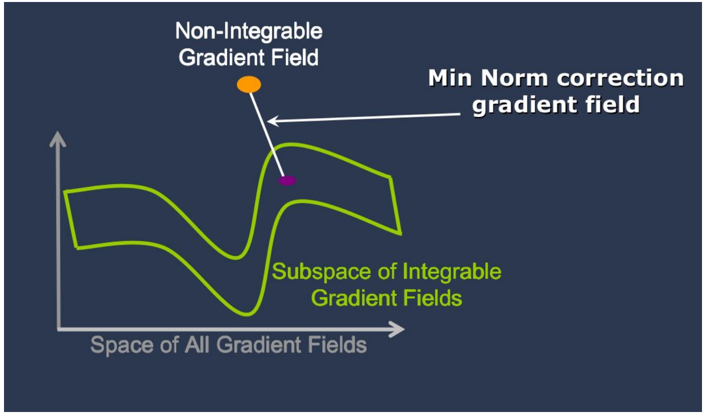
Intensity Gradient Manipulation

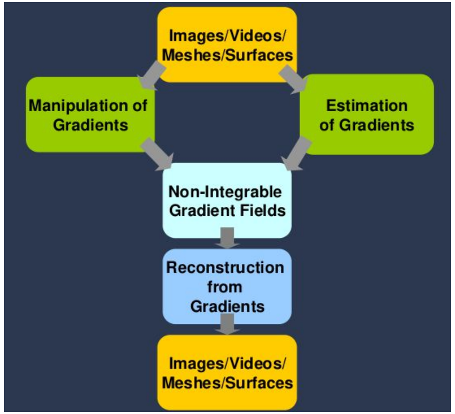


Gradient Domain Fusion

Even after seam objective penalty, the source images may be too dissimilar for a graph-cut alone to result in visually seamless composites. If the graph-cut optimization cannot find ideal seams, artifacts may still exist. Then GDP step becomes useful.

- Try to match the gradients across two images instead of the raw pixel luminance values
- Reintegrate the gradient field using methods like Poisson or Frankot-Chellappa integration
- In case of synthesized gradient fields, they usually need to be projected into the subspace of integrable gradient fields





Ref: ICCV'07 course on *Gradient Domain Manipulation Techniques* - MERL, Agrawal and Raskar

Results on different objectives

Designated source image objective

Figure 1 From a set of five source images (of which four are shown on the left), we quickly create a composite family portrait in which everyone is smiling and looking at the camera (right). We simply flip through the stack and coarsely draw strokes using the *designated source* image objective over the people we wish to add to the composite. The user-applied strokes and computed regions are color-coded by the borders of the source images on the left (middle).

Carnegie Mellon University

Results on different objectives

Creating a Time series -

Maximum likelihood objective for background, followed by Maximum difference objective image

Figure 5 To capture the progression of time in a single image we generate this stroboscopic image from a video sequence. Several video frames are shown in the first column. We first create a background image using the *maximum likelihood* objective (second column, top) and then add it to the stack. Then, we use the *maximum difference* objective to compute a composite that is maximally different from the background (second column, bottom). A lower weight for the image objective results in fewer visible seams but also fewer instances of the girl (third column, top). Beginning with the first result, the user removes the other girls by brushing in parts of the background and one of the sources using the *designated source* objective (third column, bottom) to create a final result (right).

References

1. Aseem Agarwala, Mira Dontcheva, Maneesh Agrawala, Steven Drucker, Alex Colburn, Brian Curless, David Salesin, and Michael Cohen. 2004. **Interactive digital photomontage**. *ACM Trans. Graph.* 23, 3 (August 2004), 294–302. DOI:https://doi.org/10.1145/1015706.1015718

2. Y. Y. Boykov and M. -. Jolly, **"Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images,"** *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, Vancouver, BC, Canada, 2001, pp. 105-112 vol.1, doi: 10.1109/ICCV.2001.937505.

3. Amit Agrawal and Ramesh Raskar, **Gradient Domain Manipulation Techniques in Vision And Graphics**, http://www.amitkagrawal.com/ICCV2007Course, ICCV 2007 Course, Mitsubishi Electric Research Labs (MERL).