Style Transfer for Headshot Portraits

Overview

Introduction

Quite often, it is not feasible to capture studio-like quality headshot portraits, since we might not have the required lighting conditions, studio setup, etc. By transferring local statistics, we can transfer these feature-specific lighting styles and retouching effects onto the input image from a selected example image.

In this manner, we would be able to achieve studio-quality portraits from photos captured by a casual photographer.

Problem Statement

Given a subject portrait (input image) and a professionally taken headshot portrait (example image), match the appearance/style of the subject to that of the example.

Proposed Approach

- Morphological matching
- Dense SIFT flow matching
- Pyramid based energy transfer
- Finishing touches

As outlined in the 2014 paper - https://people.csail.mit.edu/yichangshih/portrait_web/2014_portrait.pdf

Workflow specifics

Morphological matching

- Detect facial landmarks using facial templates
- Align example image to the input image based on affine transforms
- Morph example image to match the input image considering segments of the facial templates

Related Resources

1. Face Alignment through Subspace Constrained Mean-Shifts
2. Personal Photo Enhancement Using Example Images
3. Feature-Based Image Metamorphosis

Dense SIFT flow matching

- Fine tune the morphological transform of the example image by considering dense SIFT flows among the two images.
- Using normal optical flows won’t work since there might be changes which occur across many pixels.

Related Resources

1. SIFT Flow: Dense Correspondence across Scenes and its Applications

Pyramid based image transfer - 1

- We first construct a laplacian pyramid of the image, except that we don't downsample the image as we proceed through the levels.
- Energy of a given image is defined as the square of the image convolved with a gaussian filter => S[l] = conv(L^2[l], G(l+1)) where l is the current level and G(l+1) is the gaussian at the next level.

Pyramid based image transfer - 2

- At each level we perform the following steps.
  - Calculate the energies of the image (before the warping steps we did earlier) and then apply the same warping transforms to transform the energies. This is required since applying energy calculation directly on the warped image might give false peaks etc at times.
  - We then calculate the gain between the energy maps of both the morphed example image and the input image.
  - This gain is nothing but ⇒ sqrt(energy(example) / (energy(input) + ε))
  - We set the output image at this level to be Input image at current level multiplied with the gain ⇒ L[out] = L[in] * gain

Pyramid based image transfer - 3

- At times there are issues with the gain maps, so we clamp it's limits between 0.9 and 2.8 and then apply a gaussian averaging to smoothen the map.
- Once we have the output images at each of the layers we will merge them to form our final image. For the last residual layer we consider the output of the example image directly. And reconstruct the image using this output gaussian pyramid.
- We use the CIE-Lab color space to work with each channel independently in all the steps above. For each of the images we assume them to be a double representation (0.0 to 1.0 range for each pixel).

Related Resources

1. Preattentive texture discrimination with early vision mechanisms

Finishing touches - 1

- Define a face mask which we use to remove the background. This is done using a grab cut along with the matting laplacian.
  - conv(img, gaussian) = conv(img*mask, gaussian) / conv(mask, gaussian)
- Background transfer
  - Directly replace the background in the input image with the background in the example image.
  - Any missing data is extrapolated using inpainting.

Finishing touches - 2

- Eye highlight transfer
  - Consider specular reflections in example image eyes and transfer them to the input image.
  - This is done by locating the iris first using circular arc detection around eye position given by the face template.
  - We then move over the highlight to the input image averaging existing highlights if any on the input image. This is done using k-means to detect pixels corresponding to the highlight.

Related Resources

1. “GrabCut” — Interactive Foreground Extraction using Iterated Graph Cuts
2. A Closed Form Solution to Natural Image Matting
3. High confidence visual recognition of persons by a test of statistical independence
4. PatchMatch: A Randomized Correspondence Algorithm for Structural Image Editing

Members

Vishal Batchu

Sri Aurobindo Munagala