Digital Image Processing

Optimizing Color Consistency in Photo Collections

Team name: thankyoudoctor

Shubh Maheshwari - 20161170

Trunapushpa - 20161173

27th November, 2018

Github Link

Abstract

With dozens or even hundreds of photos in today's digital photo albums, editing an entire album can be a difficult task. Existing automatic tools operate on individual photos without ensuring consistency of appearance between photographs that share content.

We used a method that operates by efficiently constructing a graph with edges linking photo pairs that share content and consistent appearance is achieved by globally optimizing a quadratic cost function over the entire graph.

Our method automatically enforces consistent appearance of images that share content without any user input. When the user does make changes to selected images, these changes automatically propagate to other images in the collection, while still maintaining as much consistency as possible.

Introduction

Problem Statement

The goal of the project is to understand and implement algorithms for consistent editing over a collection of photos. When the user does make changes to selected images, these changes automatically propagate to other images in the collection, while still maintaining as much consistency as possible.

Motivation

With dozens or even hundreds of photos in today's digital photo albums, editing an entire album can be a daunting task. Existing automatic tools operate on individual photos without ensuring consistency of appearance between photographs that share content.

Constraints

In the process of imposing color consistency, we attempt to balance between achieving color consistency and preserving the dynamic range and natural appearance of individual photos.

Overview

The basic steps followed are :

Matching Photos in a Collection : A lightweight link prediction mechanism (Non-Rigid Dense Correspondence (NRDC)) is used to suggest candidate pairs that are likely to yield meaningful matches.

Appearance Consistency Optimization: Ensuring that pixels depicting the same content have the same color across different images. Avoiding unsightly visual artifacts, such as gradient reversals or severe loss of contrast. attempting to preserve as much as possible the original dynamic range of each photo.

Transformation: Transform all the images to obtain the color consistency among all images.

Approaches or Work Performed

Image Correspondence

Histogram matching

We tried histogram matching of the source image with the reference image in multiple colour spaces.

Source

Histogram Equalisation in RGB space

Reference

Histogram Equalisation in HSI space

In the above case the method seems to work.

Now consider the following test case:

Source

Histogram matching in RGB space

Reference

Histogram matching in HSI space

We find that the above method of equalizing the histogram doesn't work well when the images are less similar in the sense of colour and exposure setting, like in the above example the backgrounds are different in the source and the reference image.

Thus, we need to do first need to find better ways to find the correspondences between images and find a transformation model that can effectively convert the source image.

SIFT(Scale-invariant feature transform):

Sift is widely used in computer vision for image matching. Using Sift Keypoints we can find important key points which show correspondence between 2 images. These key points can be used for aggregating consistent regions.

Source

Reference

SIFT

NRDC

NRDC (Non-Rigid Dense Correspondence)

NRDC is an efficient method for recovering both a reliable local sets of dense correspondences between similar regions, and a global non-linear parametric color transformation model. This method is designed for pairs of images depicting similar regions acquired by different cameras and lenses, under non-rigid transformations, under different lighting, and over different background.

Calculating NRDC for image:

1. Nearest-neighbor search: Given a source image S and a reference image R, we compute the Nearest Neighbor Field from S to R, i.e., for each patch u ∈ S we seek a transformation T^u such that $T_u = \text{argmin} \| S_u - R_{T(u)} \|_2$. T^u consists of rotation, scaling, translation, bias and gain.

   We use small overlapping patches of eight by eight pixels, and a four dimensional feature vector per pixel, which includes the three channels of Lab color space and the magnitude of the luminance gradient at each pixel.

2. Aggregating consistent regions: Consistent regions of can be added to give dense correspondence regions in both source and reference image. For u, v ∈ S with matched transformations T^u,T^v , we define consistency error as

$$C(u,v) = \frac{\|T^v(v_c) - T^u(v_c)\|_2}{\|T^u(u_c) - T^u(v_c)\|_2}$$

   If C(u, v) < τ_{local} we add the patch to the consistent regions.

3. Global color mapping: Global color transformation model has two purposes
    a. Improving correspondence algorithm for reduces latent space transformation parameters
    b. Create the final image using the parametric model
4. Search constraints: We reduce the parametric space of the transformation. If reliable region is less than, 1% we state no correspondence.

We continue the above process until we detect no change in the reliable regions and hence we get our local dense correspondence field and color transformation model

Reference (Left) Source (Middle) NRDC Transformed (Right)

Why not SIFT:

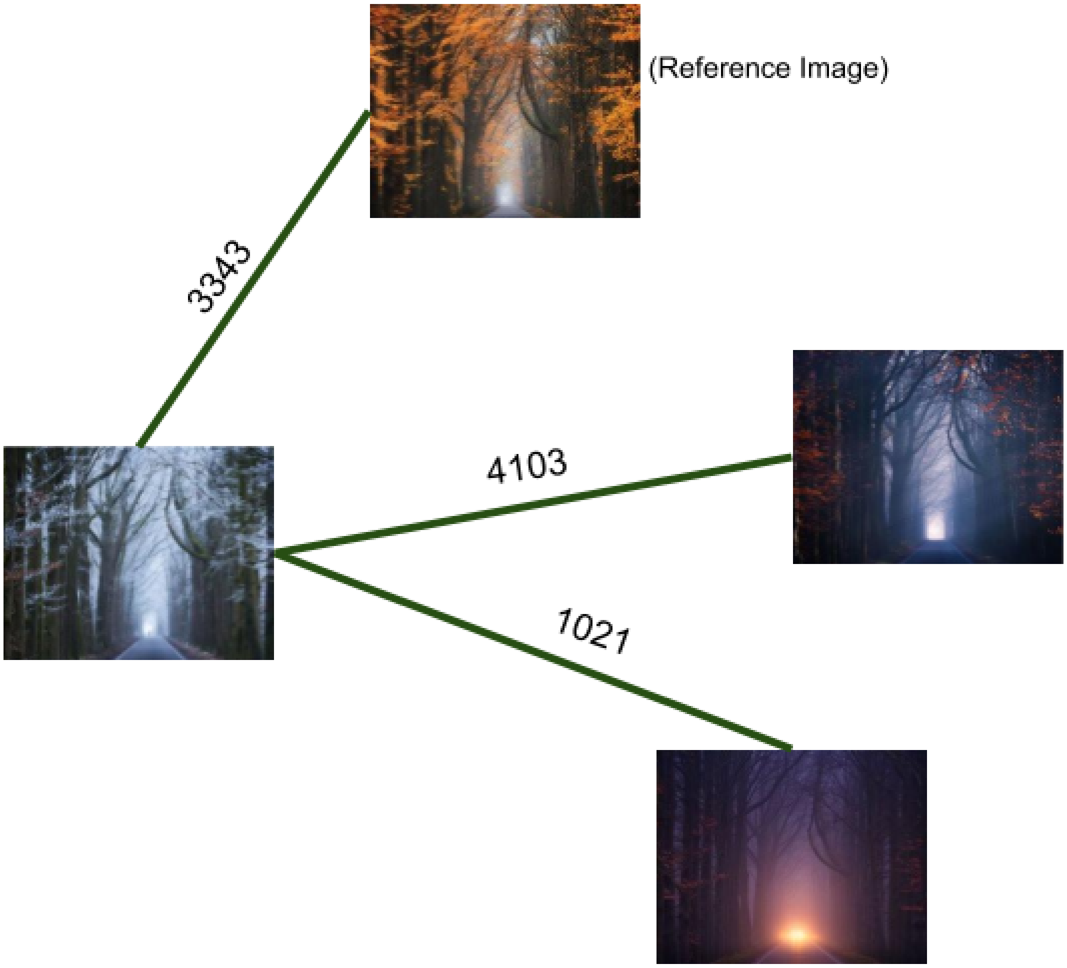
SIFT feature are very sparse and contain error while creating the consistent regions that cannot be filtered easily in presence of non-rigidly moving people. Hence sift was unable to

Source and Reference (First Column), Consistent Regions in Source and Reference and using SIFT (Second Column) and Consistent Regions in Source and Reference and using SIFT (Third Column)

Approaches or Work Performed

Match Graph

For a given set of photo images, our goal is to ensure consistent appearance (similar color and exposure) of shared content across multiple photos of the collection, and to maintain this consistency as the user performs interactive adjustments on selected images. To achieve this goal, as a pre-processing step, we construct a match graph G = {V,E} whose vertices V = {Ii} n i=1 represent individual photos in the collection and whose edges E contain information regarding the correspondences between photo pairs. E is calculated by finding the consistent regions of the correspondence. Using this match graph we minimize a quadratic cost function that penalizes color differences between matching areas.



Approaches or Work Performed

Appearance Consistency Optimization

In order to achieve consistent appearance between multiple photos in a collection we attempt to strike a balance between three potentially contradictory goals:

1. Ensuring that pixels depicting the same content have the same color across different images
2. Avoiding unsightly visual artifacts, such as gradient reversals or severe loss of contrast.
3. Attempting to preserve as much as possible the original dynamic range of each photo.

Given an images $v_i, v_j \in G(V)$ we would like to find a transformations f_i, f_j such that

$$\{\hat{f}_i\}_{i=1}^{n} = \underset{\{f_i\}_{i=1}^{n}}{\arg\min} \sum_{i \neq j} A(f_i, f_j) + \sum_{i=1}^{n} C_{soft}(f_i)$$
$$\text{subject to: } C_{hard}(f_i), \forall i \in \{1, \ldots, n\}$$

Here $A(f_i, f_j)$ is a pairwise affinity term that penalizes color differences between shared content, while $Csoft(f_i)$ is a unary term that softly enforces certain constraints on the color transformations, in addition to hard constraints enforced on them by $Chard(f_i)$.

Affinity Term $A(f_i, f_j)$:

The pairwise affinity term $A(f_i, f_j)$ is defined using the weighted SSD (sum of squared differences) between color-mapped pairs of matching pixels.

$$A(f_i, f_j) = \sum_{\mathbf{p}} w_{i,j}(\mathbf{p}) \left| f_i(I_i(\mathbf{p})) - f_j(I_j(M_{i,j}(\mathbf{p}))) \right|^2$$

Here w(i,j) is the confidence corresponding the v_i, v_j images and M(i,j) is partial pixel-wise mapping that maps pixels in v_i to v_j

Approaches or Work Performed

Color transformation model

We define the color transformation model f_i, an expressive global parametric model very similar to the global transformation defined in NRDC. It is global color transformation because it maps colors which were not mapped by the dense consistent regions.

The model fits three monotonic curves, one per channel of the RGB color space. Also its stores a linear transform to accommodate saturation changes.

Each curve is a smooth piecewise-quadratic spline with 7 knots at (0, 0.2, 0.4, 0.6, 0.8, 1), which translates to 8 degrees of freedom per curve. This model is flexible enough to compensate for a variety of common appearance differences, such as gamma curves, S-curves, color temperature changes, and other common global operators. We use quadratic programming to solve for the curves' degrees of freedom for the Appearance Consistency Optimization.

Affinity Term:

Pairwise affinity term, penalizing color diff. between shared content. We have already defined above in Appearance Consistency Optimization.

C Soft (f_i):

. The shape of each curve is regularized via unary soft constraint term:

$$\begin{aligned} C_{soft}(f_i) &= \lambda_1 \textstyle\sum_{x\in\{0,1\}} |f_i(x)-x|^2 \\ &+ \lambda_2 \textstyle\sum_{x\in\{0.2j-0.1\}_{j=1}^5} |f_i(x)-x|^2 \\ &+ \lambda_3 \textstyle\sum_{x\in\{0.2j-0.1\}_{j=1}^5} |f_i''(x)|^2. \end{aligned}$$

Hyperparameters:

1. λ1 and λ2 together try to pull the curve towards identity mapping
2. λ3 controls the smoothness of the curve by penalizing large deviations in the second derivative
3. λ1 preserves the dynamic range and λ2 to control preservation of original appearance

C Hard (f_i):

. The curve is forced to be strictly monotonic and to cross the x axis right of the orig

$$C_{hard}(f_i): \begin{array}{ll} \text{i.} & 0.2 \leq f_i'(x) \leq 5, \quad \forall x \in \{0.2j - 0.1\}_{j=1}^{5} \\ \text{ii.} & f_i(0) \leq 0 \end{array}$$

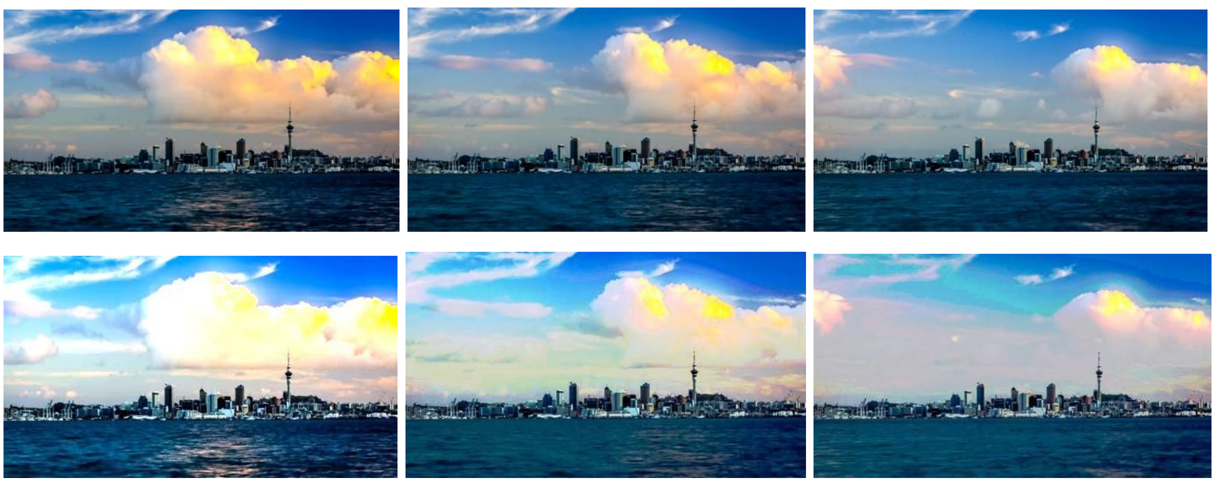
Why not use the dense correspondence given by the above algorithms?

We chose to use the global parametric color model described above over a local one for several reasons

1. A global model is more robust to errors in the correspondence.
2. A global model is easier to regularize.
3. A global parametric model is more efficient to optimize due to the small number of unknowns it involved.

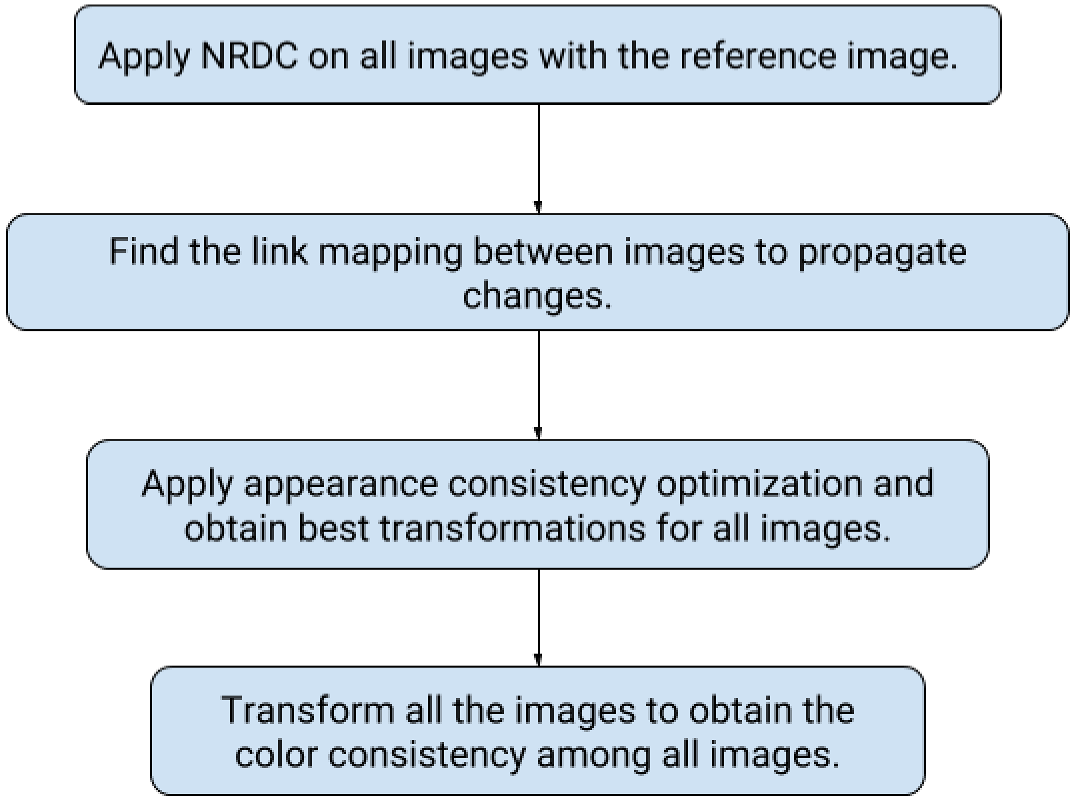
Propagating changes in a photo album

This method automatically enforces consistent appearance of images that share content without any user input. When the user does make changes to selected photos in an album, these changes automatically propagate to other images, while still maintaining as much consistency as possible. This makes it possible to interactively adjust an entire photo album in a consistent manner by manipulating only a few images.



In the above photo album only the contrast of the first image was increased. The rest of the changes were automatically propagated.

Workflow


Apply NRDC on all images with the reference image.
Find the link mapping between images to propagate changes.
Apply appearance consistency optimization and obtain best transformations for all images.
Transform all the images to obtain the color consistency among all images.

Application in photo albums

Input Images (1st image is source)

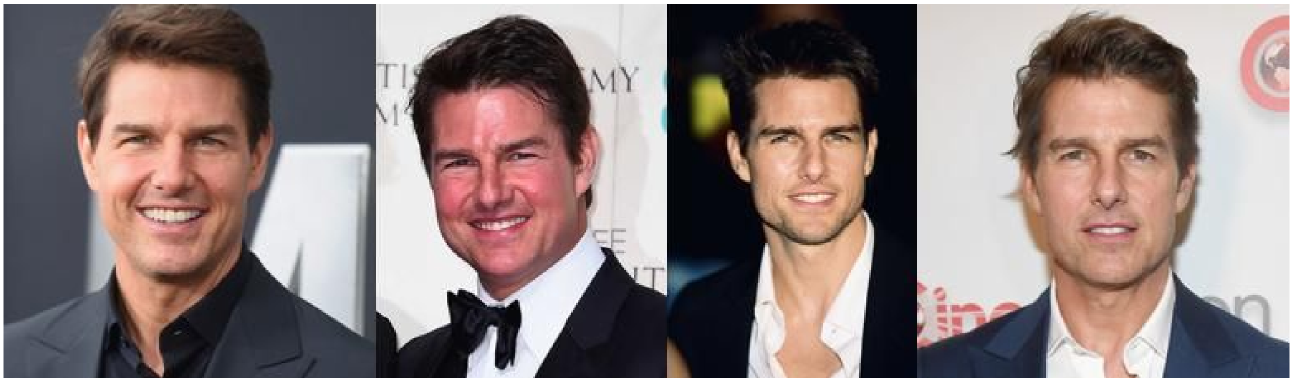
Output Images

Input Images (2nd image is source)

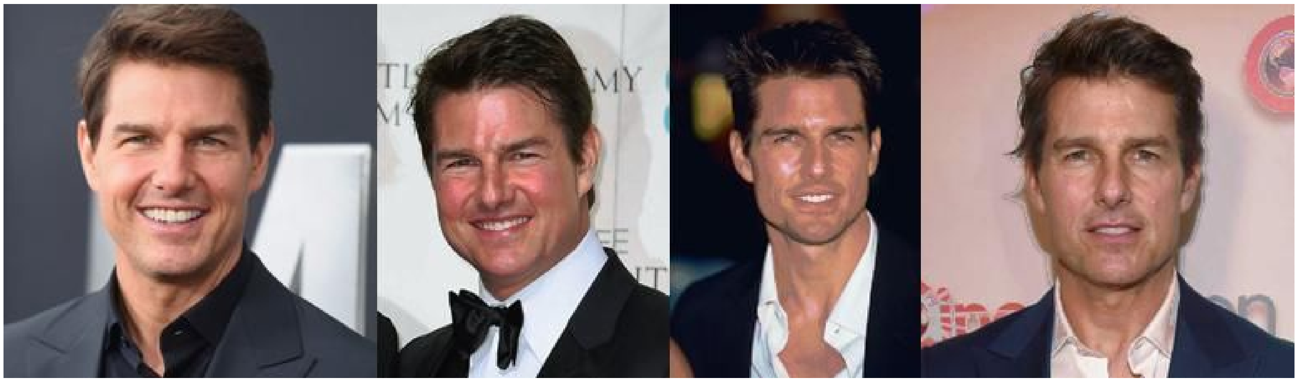
Output Images

Input Images (Top left image is source)

Output Images



Input Images (1st image is source)



Output Images

Input Images (1st image is source)

Output Images

Application in Panorama Stitching

Test 1

Panoramic Stitching without colour correction

Panoramic Stitching after colour correction

Test 2

Panoramic Stitching without colour correction

Panoramic Stitching after colour correction

Propagating user edits in a video

This method automatically enforces consistent appearance of images that share content without any user input. When the user does make changes to selected frames in a video, these changes automatically propagate to other images in the video, while still maintaining as much consistency as possible. This makes it possible to interactively adjust an entire video in a consistent manner by manipulating only a few frames.

Reference Video: https://www.youtube.com/watch?v=rUWxSEwctFU

In the above video on changing contrast of just the first frame, the following video is obtained by applying this method:
https://drive.google.com/file/d/1-PqmA90_nomWKzyg7KzLrAYKTipPQgCr/view?usp=sharing.

We observe that the change is propagated to all the frames with similarity.

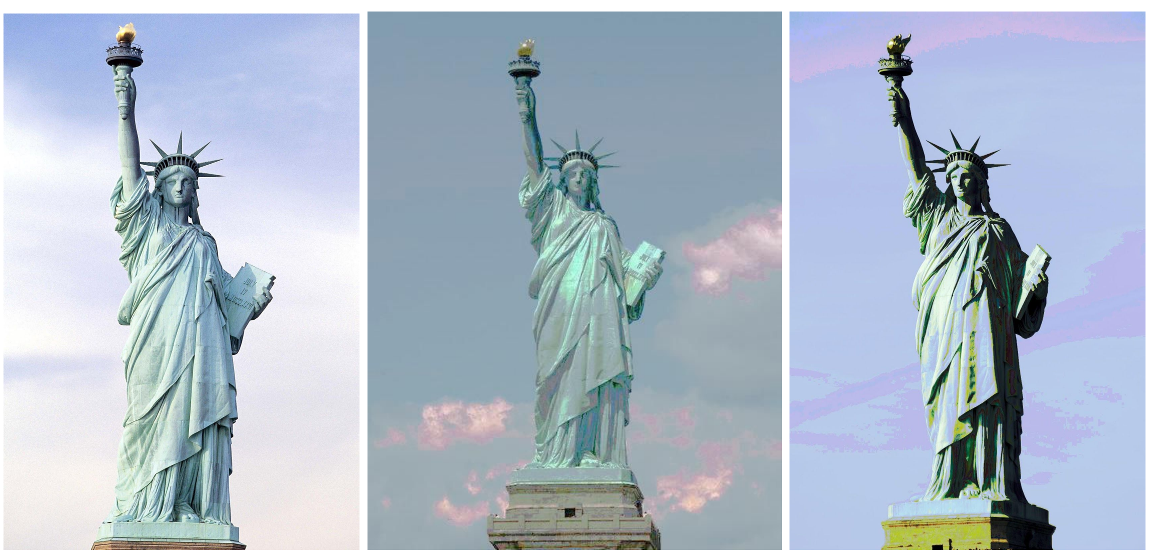
Video showing this can be found here:

https://drive.google.com/open?id=1vFWglOoh7_XSBmmFYEgJm76QSEw_Au5A

Failure cases

The following case gives undesired changes in the output.

Input Images (1st image is source)



Output Images

NRDC has difficulty in finding reliable correspondences in very large smooth regions, such as clear sky in the above case.

Analysis and Discussion

- Histogram Matching does not good result with non-significant similarities between the images. Thus colours which which appear in only any one of the images will fail to produce meaningful mapping for colours.
- SIFT feature mapping gives sparse correspondence regions which is error prone hence we use NRDC which has a very limitations.
- NRDC is a state of the art method the find correspondences for non rigidly objects.
- Hence we are using NRDC.
- To find the colour transformation model we need to solve a quadratic cost function over the entire graph.
- We observe that images with high intensity intensity coherency had a dense correspondence which are transformed more than the ones having sparse correspondence.
- The time for making the dense graphs and its further propagation is time consuming, hence, we need an even better mechanism to predict the links and propagate.

Limitations

- This method currently accounts for global color appearance variations only
- The color transfer model is based on RGB curves and does not model saturation changes well.
- This method works best when the collection contains a substantial amount of shared content, as typically occurs in personal and professional collections.

Division of Work

- Shubh
    - Non-Rigid Dense Correspondence (NRDC)
    - Match graph
    - Propagating User Edits
- Trunapushpa
    - Appearance Consistency Optimization
    - Color transformation model
    - Apply method on panorama and videos

Future Work

1. Accelerating Match Graph Construction

    NRDC is very computationally expensive, instead we could only use NRDC when their is a higher likelihood of correspondence edge being present between 2 images. To calculate the likelihood of edge being present we can train an SVM classifier where the kernel matrix would be defined by using the sift features.

2. Link Prediction Strategy

    As most of the images would have only correspondence with some images we will obtain a sparse weight matrix.

Acknowledgement

1. We would like to thank our mentor, TA Ashwin Pathak for his guidance in completing the project.
2. We used the beta implementation of NRDC provided by Yusheng Yves Xie on github. Link
3. We would like to thank HaCohen et al. for writing such an incredible paper.

References

1. HACOHEN, Y., SHECHTMAN, E., GOLDMAN, D. B., AND LISCHINSKI, D. 2011. Non-rigid dense correspondence with applications for image enhancement. ACM Trans. Graph. 30, 4, 70:1–70:9