Intrinsic Image Decomposition

Submitting: Dor Litvak
Advisor: Dr.Oren Freifeld

1 Background

Intrinsic image analysis is the problem of decomposing an image into various scene characteristics and is a useful midlevel description of the scene.
We will discuss an image decomposition into two specific scene characteristics: Shading and Reflectance. Assuming a Lambertian surface model, where the perceived illumination is constant from all angles of incidence, the observed image decomposes into the product of the intrinsic shading and reflectance images. The reflectance image contains the albedo of the object surface, whereas the shading image captures the amount of reflected light from the surface.

Intrinsic image analysis is also important for other fields of computer vision. The shading image can be exploited in shape-from-shading algorithms to reveal the underlying 3D structure of an object or to infer elements of the scene illumination, such as the number, location, and color of the light sources. The reflectance image improves many segmentation algorithms, where shading effects often introduce artifacts.

2 The method

We present a two-steps Bayesian Non-parametric approach that contains an alternation between the reflectance and the shading estimation.

We will work in the log domain where the log of the observed image, x, is assumed to be generated from the sum of the log shading and the log reflectance image.[1]

$$log(X) \sim log(S) + log(R) \tag{1}$$

X is the observed image, S is the shading image, and R is the reflectance image.

Denote $X_{i=1}^{K}$ images of the same instance taken under different lighting conditions [3]. We wish to evaluate for each image the shading and the reflectance, denote as $S_{i=1}^{K}$, $R_{i=1}^{K}$. Our method is an iterative method alternating between the reflectance and the shading estimation. We first evaluate the image reflectance using Versatile Hierarchical Dirichlet Processes, estimation is done over all the images together. Then, for each image separately, we will calculate the shading by solving a Least-squares problem.

2.1 Reflectance Estimation

The log reflectance image, denoted $R_{i=1}^{K}$, is estimated by using Versatile Hierarchical Dirichlet Processes (vHDP) [2].
Consider the following setting, $X_{i=1}^{K}$ are K images of the segments under different lighting conditions. If we

can find the "real" colors of the features, we can restore the shading. Given K images, we wish to partition each image into segments consistently across all frames. If a segment appears in more than one image, each pixel that belongs to this segment should get the same global cluster across all of the images. The locations, as well as the number of the segments, are unknown. Each image is a data group where each pixel is associated with a 5D data point = (a 3D RGB value, a 2D location). The color is the global part as the colors of an object are often similar across different images. The location is the local part as an object's location usually varies across images.

Clustering by color and location of the images using Versatile Hierarchical Dirichlet Processes (vHDP) [2] would give estimated clusters for each segment in the image. The reflectance image $R_{i=1}^{K}$ are images where each pixel is the mean color of his global cluster.

2.2 Shading Estimation

The log shading image, denoted g, is the result of a Least Squares problem as described next.

$$argmin_g \ \lambda||Ag||_2^2 + ||Id \cdot g - y||_2^2 \tag{2}$$

Where A is the neighborhood matrix, y - is the noisy shading observation, $y = log(X) - log(S)$, and g is the shading we want to estimate. Id is the identity matrix and λ is a parameter controlling the influence of the neighborhood of a specific pixel in the shading image.

This Least Squares problem has a closed form solution:

$$(Id + \lambda A^T A)g = y \tag{3}$$

g is the solution of a sparse linear system ($Id + \lambda A^T A$ is a sparse matrix).

2.2.1 The neighborhood matrix:

A is the neighbors matrix, a very sparse matrix where each row has only two cells differ from 0. A can be build using the following algorithm:

Algorithm 1 Building A

Result: A neighborhood matrix
A = zeros()
for *each two pixels in the image* X_i, X_j **do**
  if X_i, X_j *are neighbors* **then**
    Add a row to the A matrix with N cells, in the i-th index place 1 and in the j-th index place -1.
    (X_i is a neighbor of X_j)
    Add another row to the A matrix with N cells, in the j-th index place 1 and in the i-th index place -1.
    (X_j is a neighbor of X_i)
  end
end

Where N is the number of pixels in the image X. The A matrix describe a constraint over the pixels, if pixels are called "neighbors" they need to be similar in color.

3 Conclusions

Using the state-of-the-art method for clustering grouped data [2] and alternation process between shading and reflectance estimation, we are getting a low computation process that can handle a wider set of problems. We are hoping to obtain state-of-the-art results of image decomposition to shading and reflectance. The code is written in Julia language.

References

[1] Jason Chang, Randi Cabezas, and John W Fisher. Bayesian nonparametric intrinsic image decomposition. In *European conference on computer vision*, pages 704–719. Springer, 2014.

[2] Or Dinari and Oren Freifeld. Scalable and flexible clustering of grouped data via parallel and distributed sampling in versatile hierarchical dirichlet processes. 2020.

[3] Roger Grosse, Micah K. Johnson, Edward H. Adelson, and William T. Freeman. Ground-truth dataset and baseline evaluations for intrinsic image algorithms. pages 2335–2342. International Conference on Computer Vision, 2009.