Group Work - Computer Vision

Uncalibrated Photometric Stereo Constrained by Intrinsic Reflectance Image and Shape From Silhoutte; A Reimplementation

Arthur Clarysse
arthur.clarysse@vub.be
Vrije Universiteit Brussel
Brussels, Belgium

Jens Dumortier
jens.dumortier@vub.be
Vrije Universiteit Brussel
Brussels, Belgium

Guillaume Tibergyn
guillaume.tibergyn@vub.be
Vrije Universiteit Brussel
Brussels, Belgium

1 Summary

The reimplementation in this group project was based on the paper "*Uncalibrated photometric stereo constrained by intrinsic reflectance image and shape from silhoutte*" by Hashimoto et al.. In this paper, they try to estimate the surface normal, which in itself is not a very challenging problem, However, in our case, the direction of the light sources is unknown, making it much harder to solve. Mainly because of its ambiguous nature. Fortunately, it is possible to solve this problem by adding two constraints: the intrinsic reflection (albedo) and an approximate normal. This results in a reasonable estimate of the surface normal; the estimation and constraints are shown in Figure 1.

The main goal of the original paper is to implement an uncalibrated photometric stereo pipeline that estimates surface normals from multiple images without prior knowledge of light directions. This pipeline first estimates the intrinsic reflectance (or albedo) and the estimated normal, which can be used as constraints. Finally, using the same matrices, we can calculate the guide normal and use it to make a 3d representation of the image.

As a dataset, we use the photometric stereo dataset made by Harvard [Xiong et al. [n. d.]]. The dataset contains multiple shaded images of different animals, including cats, frogs, turtles, and more. To evaluate our results, we will calculate the average error of the estimated depth map with the actual results provided in the dataset.
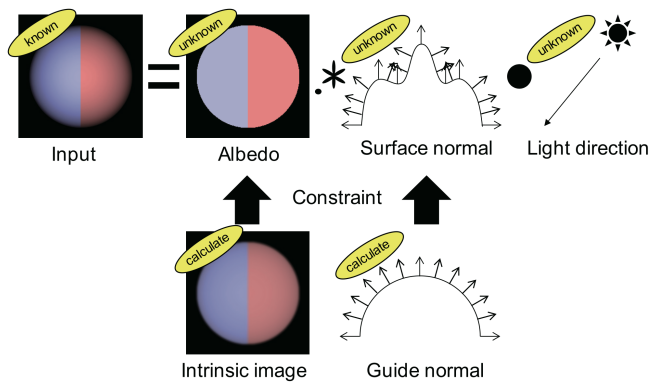

Figure 1: Overview of the known and estimated parts of the paper

2 Implementation

2.1 Dataset & Preprocessing

We first started by collecting the data from dataset made by Xiong et al., then, we selected the animal of which we want to make a 3D representation. In our case, we chose the cat image, one of the twenty input images can be observed in Figure 2a.

2.2 Albedo Estimation

After the preprocessing, step, we can estimate the albedo. To do this, we can start by calculating the average image, this results in an image where shades an shadows are reduced. Unfortunately, some shading effects remain, we can solve this by applying a bilateral filter.

The result from this step is an estimated albedo, and is shown in Figure 2b. This result is still unusable in normal photometric stereo techniques, but fortunately, we can use the framework of uncalibrated photometric stereo to solve this.

2.3 Singular Value Decomposition

In this step, we want to calculate the true surface matrix and the true light matrix, S and L respectively.

1. We start by calculating the shadow matrix $\hat{I}$ (intrinsic illumination matrix) by dividing an input image by the albedo.

$$\hat{i}_{pf} = \frac{i_{pf}}{\hat{a}_p}$$

With
- i_{pf}: The pixel p of the image f
- $\hat{a}_p$: The albedo of image f

This results in the shading image $\hat{I}$ with on its i-axis the images and on its j-axis the pixels.

2. Then we can perform singular value decomposition froem which we can calculate the pseudo surface matrix S' and pseudo light matrix L'.

$$I = U'W'V'^T$$

With
- $S' = U'$
- $L' = W'V'^T$

3. Finally, we can calculate the true surface matrix S and true light matrix L by disambiguting the matries using the ambiguity matrix A.

$$S = S'A$$

$$L = A^{-1}L'$$

2.4 Constant Albedo Constraint

In the shading images, the albedo has been divided out, so for pixels where our albedo estimate is reliable, the albedo should be constant (we normalise it to 1). Following the original paper, we use this idea to further reduce the ambiguity between the pseudo surface matrix S' and the pseudo light matrix L'.

First, we only consider pixels whose estimated albedo $\hat{a}_p$ is larger than a threshold T_a, so that the division in the shading image is stable. For these pixels we know that the norm of the true surface vectors should be constant, i.e. the squared albedo is equal to one. We express the remaining ambiguity between S' and L' by a 3×3 matrix A and write the true surface vectors as

$$s_p = A^\top s'_p.$$

The squared norm of s_p can be written as

$$\|s_p\|^2 = s_p^\top s_p = s_p'^\top B s'_p,$$

where $B = AA^\top$ is a symmetric matrix with six unknown entries. Enforcing a constant albedo now becomes the constraint

$$s_p'^\top B s'_p = 1$$

for all selected pixels. By expanding this equation for all pixels we obtain a linear system in the six unknown entries of B, which we solve in a least-squares sense.

Once B is known, we compute its eigenvalue (or singular value) decomposition and recover the ambiguity matrix A by taking the square root of its eigenvalues. We then update the pseudo surface and light matrices as

$$S'' = S'A, \qquad L'' = A^{-1}L'.$$

After this step, only a remaining orthogonal ambiguity is left between S'' and L''.

2.5 Guide Normal Constraint

Even after applying the constant albedo constraint, the solution is still ambiguous up to a global rotation. In other words, there exists an orthogonal matrix R such that

$$S = S''R, \qquad L = R^\top L'',$$

and we still need to determine R. To resolve this last ambiguity, we use a coarse *guide normal* that approximates the overall shape of the object.

First, we detect the object region and its silhouette from the input images. From this silhouette we compute an approximate shape using simple image-based geometry, and from this shape we obtain a guide normal matrix $\tilde{S}$, which contains approximate surface normals for the same pixels as S''. These normals are not very accurate in the details, but they capture the global orientation and rough shape of the object.

We then find an orthogonal matrix R that best aligns the updated pseudo surface matrix S'' with the guide normals $\tilde{S}$. In practice, we solve

$$\tilde{S} \approx S''R$$

in a least-squares sense, and afterwards project R onto the set of orthogonal matrices (for example via an SVD-based orthogonalisation). Finally, we obtain the true surface and light matrices as

$$S = S''R, \qquad L = R^\top L''.$$

In this way, the guide normal is only used to fix the remaining three degrees of freedom (a global rotation), so the global pose and coarse shape follow the guide normal, while the finer details of the surface are preserved from the uncalibrated photometric stereo solution.
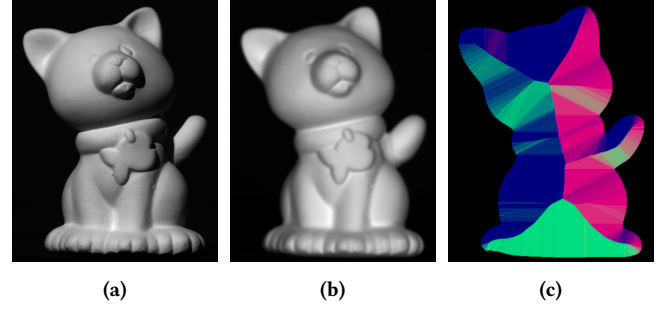

| (a) | (b) | (c) |

Figure 2: Intermediate results of our uncalibrated photometric stereo pipeline: (a) input image, (b) estimated albedo, and (c) guide normal map

3 Results

The output of the pipeline is the guide normal, which can easily be visualized, the results can be seen in Figure 3a. From this result, a height map can be made using the Frankot-Chellappa algorithm, this result is visible in Figure 3b.
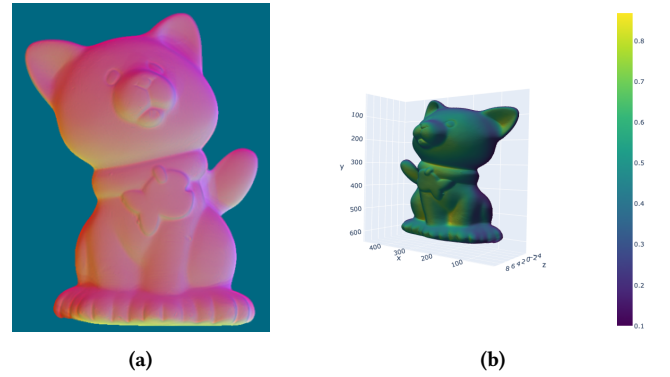

| (a) | (b) |

Figure 3: The 3D results visualized

4 Discussion

...

5 Reproducibility Checklist

- ☐ Code provided & documented
- ☒ Data included
- ☒ Environment included
- ☒ Run instruction included
- ☒ Random seeds
- ☐ Expected outputs
- ☐ Algorithm pseudo code included

References

Shuhei Hashimoto, Daisuke Miyazaki, and Shinsaku Hiura. 2019. Uncalibrated photometric stereo constrained by intrinsic reflectance image and shape from silhoutte. In *2019 16th International Conference on Machine Vision Applications (MVA)*. IEEE, Tokyo, Japan, 1–6. doi:10.23919/MVA.2019.8758025

Ying Xiong, Ayan Chakrabarti, Ronen Basri, Steven J. Gortler, David Jacobs, and Todd Zicker. [n. d.]. From Shading to Local Shape. https://vision.seas.harvard.edu/qsfs/index.html