2021/4/25 – 2021/4/30

1. 参加新员工上岗培训
2. SfT探索

从12年到17年，有一系列sft的工作，建立在不同的形变假设上，例如：isometric, conformal, etc.

形变模型也有几种不同的选择，例如：thin-plate-spline, free-form-deformation, mesh vertex location deformation.

Thin-plate-spline模型在该作者的文章中经常用到，所以我把相关知识了解后，做了后续记录(Appendix)。

考虑到关键点在材质较少的图片上不易取得，我用图片的颜色差异作为监督信息学习形变模型的参数，发现效果也不好（在材质较少图片中）。因此，可能需要尝试下另外的更加鲁棒的监督信息或者其他的形变模型。

我在实现thin-plate-spline的时候是用了pytorch中的grid\_sample来确保可微操作，之前作者的文章中大多使用Gauss-Newton迭代求解。两者实现有所差异。

另外还有一种基于mesh vertex location deformation模型的方法：Direct, Dense, and Deformable: Template-Based Non-Rigid 3D Reconstruction from RGB Video. 后续可以尝试实现下。

一些相关的文章思想和摘要在后面的Appendix.

# Appendix

# Thin Plate spline

手表上有字

中度可信度描述已自动生成

<https://zhuanlan.zhihu.com/p/227857813>

<https://lccurious.github.io/2019/01/29/Thin-Plate-Spline/>

<https://khanhha.github.io/posts/Thin-Plate-Splines-Warping/>

<https://www.cse.wustl.edu/~taoju/cse554/lectures/lect07_Deformation2.pdf>

is Radial Basis Function (depends only on the distance between the input and some fixed point), w is the weight for different radial basis. is the control points, there are N control points.

Suppose each point has a v value (hight), which means . The above equation means that we use a surface to fit all the points.

means that we use a flat surface to fit the point, the last term allows us to add some curvature to this flat surface.

电脑屏幕的照片上有字

中度可信度描述已自动生成 手机屏幕的截图

描述已自动生成 文本, 信件

描述已自动生成

Let

示意图

低可信度描述已自动生成

Then

徽标

中度可信度描述已自动生成

图片包含 文本

描述已自动生成

Review Thin Plate Spline Warp: <http://vision.ucsd.edu/kriegman-grp/papers/cvpr05b.pdf>

黑色的钟表

中度可信度描述已自动生成

When noise exists in the target value V, we need to introduce the regularization parameter to control the amount of smoothing in TPS. That is the submatrix is replaced by .

In this paper, we are interested in warping 2D points using TPS defined by pairs of control points (with hat sign). Therefore, we need **two TPS functions** for x and y coordinate separately, which is:

黑色的钟表

中度可信度描述已自动生成

Use above equation to obtain the TPS model, the w and a parameteres

钟表的特写

中度可信度描述已自动生成

, the j-th row of Q is

Use above equation to compute any transformed points. We denote [B, Q] as M.

# Efficient TPS Warp Estimation

Most of prior works follow the above procedure. Find the correspondences, obtain the TPS coefficients, then the points of interest are transformed to new locations using the interpolated spline.

We consider the problem of modeling non-rigid deformations in images based on object appearance.

Suppose we are given a template of an object with control points, and a target image of the same object undergoing non-rigid motion.

In the first frame, we need to **initialize the target region** by some object detector. The control points are only used to determine the amount of non-rigid deformation. We can position control points on a regular grid as below:

社交网络的鱼

中度可信度描述已自动生成

Back-warped image: warp the target images to template image, update TPS by minimizing the template image and back-warped image.

黑色的钟表

中度可信度描述已自动生成

We pre-compute during the initialization, since they are constant, the control points are pre-defined (constant and the fixed template)

文本

描述已自动生成

# Feature-Based Deformable Surface Detection with Self-Occlusion Reasoning

Actually, it combines feature-based and pixel-based methods when the surface self-occludes.

Pixel-based method: the intensity discrepancy between images is used as a cue to compute the deformation.

Feature-based method: detect features first, and then matched, last compute the deformation.

However, pixel-based methods have limitations: being iterative and ‘local’ they generally **heavily rely on the initialization**.

Feature-based methods do not need an initialization, but have two main challenges: 1) erroneous matches; 2) surface self-occlusions, very few features are detected.

Our core idea to cope with self-occlusions is to model the motion field as being piecewise smooth instead of globally smooth. Use local smoothness to filter out erroneous matches.

* Our method can detect and discard outliers based on smoothness on the local scale.
* Our method can find which parts of the template are self-occluded in the input image.
* A feature-based warp estimation method that prevents the warp to fold
* A pixel-based warp estimation method fine-tunes the warp parameters.

With a set of inlier matched points, we search for the global warp function , parameterized by the matrix that minimizes a distance criterion between every point in and . We assume that given the correct parameter vector, maps any point in template image to its match in input image.

文本

低可信度描述已自动生成

# Image Deformation Model

TPS (with radial basis functions) and FFD (with cubic B-spline)

Let be a point coordinate vector in the template image. The warp maps 2d points from the template to the input image. It depends on a set of 2D control points stacked in the parameter matrix . The general parameter warp is defined as:

is some nonlinear lifting function.

E.g., In the FFD case, we consider a regular grid of control points, **covering a specific area of the template image**, and a modified grid of control points stacked inside . The warp function is:

where are cubic B-spline interpolation coefficients evaluated at the normalized coordinates of point (with 16 closest control points). Thus the lifting function is thus composed of B-Spline coefficients for each point p.

# Linear Least Squares Warp Estimation

The loss function is composed of a data term based on the average distance between warped point and the match points, and a smoothing term that controls the smoothness of the motion field.

Data term: 钟表的特写

描述已自动生成, which can be written as:

文本, 信件

描述已自动生成

Smoothing term: 手机屏幕截图

低可信度描述已自动生成

For TPS, smoothing term has a closed-form solution. For the other kinds of warps, it can be approximated by a Riemann sum,

钟表的特写

描述已自动生成

In all cases, can be written as a quadratic function of using a constant matrix . Which is . Therefore,

文本, 信件

描述已自动生成

The influence matrix transform the input image point coordinates in in the warp parameter matrix .

# Outlier Rejection

Using neighboring points (defined by Delaunay triangulation) to estimate the warp function, and measure the distance between and warped point from . We consider that this distance should be smaller than a chosen threshold, and is a probable inlier.

Choosing the template image coordinates to evaluate the distance is reasonable since we know every detail about it.

文本

描述已自动生成 图表

描述已自动生成

# A Global Self-Occlusion Resistant Warp Estimator (not needed in our case)

# Pixel-Based Registration Refinement

The registration accuracy obviously depends on the keypoint detection and matching. However, most of the methods often lose many points near self-occlusion boundaries or high deformed areas.

We propose to use a pixel-based registration approach to refine the surface parameters. The data term of pixel-based methods is:

钟表的特写

中度可信度描述已自动生成

(Warp input image and compare the template with the warped image)

is the template region of interest (usually the whole template image). The Euclidean norm used to compare pixel intensities or colors. Maybe we need some more robust loss. The region represents all pixel positions in the template region of interest that are not self-occluded .

# Minimizing the cost function

A straightforward way to obtain the minimum of the cost is to use the Gauss-Newton iterative optimization method. The warp parameters are additively updated at each iteration () for some increment .

文本, 信件

描述已自动生成文本, 信件

描述已自动生成

文本

描述已自动生成

一些文字和图片的手机截图

描述已自动生成

# A Pixel-Based Approach to Template-Based Monocular 3D Reconstruction of Deformable Surfaces

# Can We Jointly Register and Reconstruct Creased Surfaces by Shape-from-Template Accurately?

Most existing SfT methods require well-textured surfaces that deform smoothly, which is a significant limitation. Due to the sparsity of correspondence constraint and strong regularizations, they usually fail to reconstruct strong changes of surface curvature such as **surface creases**.

*We tend to implicitly model the creases with a dense mesh-based surface with a robust bending energy term. This robust term will deactivate curvature smoothing automatically where need (the crease locations*)

SFT: the template is composed of a texture map and a model of the object’s 3D shape.

Most Sft methods break down when surface creases, since 1) feature correspondences are not usually dense enough to tell us where creases occur; 2) use smoothed parameterizations which prefer smooth rather than creased solutions.

图片包含 形状

描述已自动生成

Since we do not know the creased locations, we can not modify the spline to permit high changes in curvature. **Instead, our solution is to implicitly model creases through an adaptive bending energy prior acting on a high-resolution non-parametric surface mesh**.

# deformation models and Priors in SfT

Two main ways to model deformations: 1) thin-shell models (only for surface); 2) volumetric models (object surface and interior volume are modeled).

Most methods use some forms of dimensionality reduction to reduce the problem’s search space, such as smoothness.

Spline models such as b-splines and thin-plate-splines reduce dimension by definition (since they model the deformation by a set of control points).

Thin-plate splines enforce global smooth deformations, and are *not suitable* to model surface creases. It is possible to model high-frequency and/or discontinuous deformations with **b-splines** by changing the spline’s order and introducing repeated control points [include papers B-spline surface]

# Monocular Template-based Reconstruction of Inextensible Surfaces

Since the camera is projective, the sightlines intersect at the camera center and are not parallel to each other. The consequence is that the distance between two points increases with their depths.

However, the distance is maximum in the template image, thus, the depth can be constraints.

The main idea is the Inextensible surfaces, the distance between two points becomes smaller after deformation, which is shown as follows:

Depth is maximized under inextensible constraints.



