本周我们回到了 face2face 的论文，借鉴他们做表情移植使用的方法，看看我们是否能够借鉴他们的方法以用于自己的工程

Face2face 在进行表情移植时，参考了一篇经典的 paper，即 deformation transfer for triangle meshes。这篇文章主要内容是针对 mesh 变形的研究，以图中的马和骆驼为例，这种方法的输入是**马和骆驼**的图像，以及**马做出一系列动作**的图像，输出是**骆驼做出一系列动作**的图像

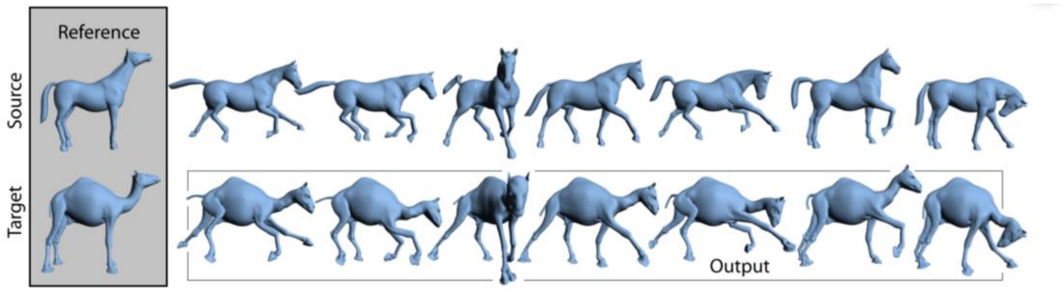

Figure 1: Deformation transfer copies the deformations exhibited by a source mesh onto a different target mesh. In this example, deformations of the reference horse mesh are transfered to the reference camel, generating seven new camel poses. Both gross skeletal changes as well as more subtle skin deformations are successfully reproduced.

我们不妨称马为 source，骆驼为 target，换到我们需要的研究中，马的动作即为摄像机前人做出的表情，target 即为目标视频中人做出的表情，我们输入人做出的表情，经过此方法后可以获得目标视频中人做出相同的表情。

这篇被参考的 paper 使用网格的变形来达到动作变化的效果。我们可以将它的实现过程分为以下 5 步：

1.建立 source 与 target 之间的网格对应，确定哪些网格需要发生变形，以达到同样动作效果

In order to relate the source deformation to the target mesh with the set of triangle indices T, the user supplies a mapping M between the set indices for the source and target triangles:

三角形对应.

一一对应

$$M = \{(s_1,t_1),(s_2,t_2),\ldots,(s_{|M|},t_{|M|})\}. \qquad (5)$$

S:

2.需要保证被不同三角形共享的顶点在不同的三角形中需要被变形到同一位置

ure 3). For the set of target affine transformations $\mathbf{T}_1+\mathbf{d}_1 \ldots \mathbf{T}_{|T|}+\mathbf{d}_{|T|}$ this requirement is

eq6:确保一致性的约束条件

$$\mathbf{T}_j\mathbf{v}_i+\mathbf{d}_j = \mathbf{T}_k\mathbf{v}_i+\mathbf{d}_k, \quad \forall i,\ \forall j,k \in p(v_i), \qquad (6)$$

where $p(\mathbf{v}_i)$ is the set of all triangles that share vertex $\mathbf{v}_i$.

式子中的 T 和 d 指的是针对 mesh 上顶点的变形，其具体定义如图

length of the triangle edges.

An affine transformation defined by the 3×3 matrix $\mathbf{Q}$ and displacement vector $\mathbf{d}$, which, for notational convenience, we write as $\mathbf{Q}+\mathbf{d}$, transforms these four vertices as follows:

$$\mathbf{Q}\mathbf{v}_i + \mathbf{d} = \tilde{\mathbf{v}}_i, \quad i \in 1 \ldots 4. \tag{2}$$

3.综合以上两点，在保证共享顶点变形一致性的前提下，我们要使得目标图像变形的矩阵 T 和源图像变形的矩阵 S 尽量相同，即求解下面这个式子

在保持一致性的同时保证 deformation：
即解决这些优化问题
保证源变化 S_{sj} 与目标变化 T_{tj}
之间差距最小

affine transformations:

$$\min_{\mathbf{T}_1+\mathbf{d}_1 \ldots \mathbf{T}_{|T|}+\mathbf{d}_{|T|}} \sum_{j=1}^{|M|} \left\| \mathbf{S}_{s_j} - \mathbf{T}_{t_j} \right\|_F^2 \tag{7}$$
$$\text{subject to} \quad \mathbf{T}_j \mathbf{v}_i + \mathbf{d}_j = \mathbf{T}_k \mathbf{v}_i + \mathbf{d}_k, \quad \forall i, \ \forall j,k \in p(v_i).$$

The matrix norm $||\cdot||_F$ is the Frobenius norm, or the square root of the sum of the squared matrix elements.

A solution of this optimization problem defines a continuous deformation of the target mesh up to a global translation. The global translation can be defined explicitly by setting the displacement $\mathbf{d}_i$ for any target triangle. In addition, other positional constraints such as foot placement can also be added.

4.我们求解上述式子，可以得到一个目标 transformation T，但更直接的，我们希望直接取得目标图像变形后的点，对此，我们可以将 T 使用目标图像形变后的点来表示：**即通过减法，将位移变量 d 消除，只留下变形矩阵 Q，我们可以将 Q 定义为如下图 eq4 的形式。每个三角形除了三个顶点 v1，v2，v3 外，还有一个正交定义的顶点 v4**

$$\mathbf{Q}\mathbf{v}_i + \mathbf{d} = \tilde{\mathbf{v}}_i, \quad i \in 1 \ldots 4. \tag{2}$$

If we subtract the first equation from the others to eliminate $\mathbf{d}$ and rewrite them in matrix form treating the vectors as columns, we get $\mathbf{Q}\mathbf{V} = \tilde{\mathbf{V}}$ where

$$\begin{aligned} \mathbf{V} &= [\mathbf{v}_2 - \mathbf{v}_1 \ \ \mathbf{v}_3 - \mathbf{v}_1 \ \ \mathbf{v}_4 - \mathbf{v}_1] \\ \tilde{\mathbf{V}} &= [\tilde{\mathbf{v}}_2 - \tilde{\mathbf{v}}_1 \ \ \tilde{\mathbf{v}}_3 - \tilde{\mathbf{v}}_1 \ \ \tilde{\mathbf{v}}_4 - \tilde{\mathbf{v}}_1] \end{aligned} \ . \tag{3}$$

A closed form expression for $\mathbf{Q}$ is given by

$$\mathbf{Q} = \tilde{\mathbf{V}}\mathbf{V}^{-1}. \tag{4}$$

这样一来，变形矩阵就由变形前的点和变形后的点一起表示，由于变形前的点是我们已知的目标图像的点，所以这一项可以直接获得我们需要的变形后的点

将这个式子变形一下，c 包含了源变形的信息，A 包含了目标 mesh 没有发生形变时点的位置，最后求得的 x 就包含了我们需要的形变后的点，详细信息可以阅读下图内容

The solution to this optimization problem is the solution to a system of linear equations. Rewriting the problem in matrix form yields

$$\min_{\tilde{\mathbf{v}}_1 \ldots \tilde{\mathbf{v}}_n} \quad \|\mathbf{c} - \mathbf{A}\tilde{\mathbf{x}}\|_2^2 \tag{9}$$

$\tilde{x}$: 变形后的点位置

c: S的条目

where $\tilde{\mathbf{x}}$ is a vector of the unknown deformed vertex locations, $\mathbf{c}$ is a vector containing entries from the source transformations, and $\mathbf{A}$ is a large, sparse matrix that relates $\tilde{\mathbf{x}}$ to $\mathbf{c}$. Setting the gradient of the objective function to zero gives the familiar normal equations:

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}^{\mathrm{T}}\mathbf{c} \tag{10}$$

A依赖于目标中未形变的部分

计算A^TA:

LU factorization

A的列数与变形点数有关 故列数n = 点数 + 三角形个数

The entries in $\mathbf{A}$ depend only on the target mesh's undeformed vertex locations. Furthermore, the system is separable in the spatial dimension of the vertices. Thus, for each source/target pair, we compute and store the LU factorization of $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ only once. Retargeting any source deformation onto the target mesh only requires performing backsubstitution to solve separately for the x, y, and z components of the deformed target vertices. For efficiency, we use a sparse LU solver [Davis 2003]. Since the columns of $\mathbf{A}$ correspond to the unknown deformed target vertices, and since we add an extra vertex for each triangle, the number of columns of $\mathbf{A}$ (and hence the number rows and columns of $\mathbf{A}^{\mathrm{T}}\mathbf{A}$) is equal to the number of vertices plus the number of triangles of the target mesh. Table 1 lists the vertex and triangle counts for the meshes in this paper, and Table 2 lists the factorization and average backsubstitution times.

5 Correspondence

如何标记目标与源之间的对应点

回到 face2face 的部分，我们现在发生形变的部分就是表情部分，表情是个有明确语义

的部分，它的表达方式应该已经由特征点确定了，所以上图的第一步，源和目标哪些网格互相对应应该也不需要我们多费心。

又由于，点的位置是由 pca 建立起来的，即 vi=M(α,β)，后者为表情参数，所以我们又能将上文中的用点表示的公式修改为用表情表示的公式，即从 eq7：

（Ai 表示摄像头前的人发生表情变化而使用的变化矩阵，V 为变化前的默克尔的脸，V`为发生变化后的默克尔的脸，即我们需要的移植表情后的脸）

As proposed by [27], we first compute the source deformation gradients $\mathbf{A}_i \in \mathbb{R}^{3\times3}$ that transform the source triangles from neutral to deformed. The deformed target $\hat{\boldsymbol{v}}_i = \boldsymbol{M}_i(\boldsymbol{\alpha}^T, \boldsymbol{\delta}^T)$ is then found based on the undeformed state $\boldsymbol{v}_i = \boldsymbol{M}_i(\boldsymbol{\alpha}^T, \boldsymbol{\delta}_N^T)$ by solving a linear least-squares problem. Let (i_0, i_1, i_2) be the vertex indices of the i-th triangle, $\mathbf{V} = [\boldsymbol{v}_{i_1} - \boldsymbol{v}_{i_0}, \boldsymbol{v}_{i_2} - \boldsymbol{v}_{i_0}]$ and $\hat{\mathbf{V}} = [\boldsymbol{v}_{i_1} - \hat{\boldsymbol{v}}_{i_0}, \hat{\boldsymbol{v}}_{i_2} - \hat{\boldsymbol{v}}_{i_0}]$, then the optimal unknown target deformation $\boldsymbol{\delta}^T$ is the minimizer of:

α^T δ_N $\rightarrow$ δ^T

output:

δ^T

$$E(\boldsymbol{\delta}^T) = \sum_{i=1}^{|\boldsymbol{F}|} \left|\left|\mathbf{A}_i\mathbf{V} - \hat{\mathbf{V}}\right|\right|_F^2 . \qquad (7)$$

式子两边同时乘一个 V 逆，再把两项换一下位置，用表情参数来表示点，注意此时的 A 与上面式子不一样了，它表示的是 V 逆的部分，即 target 的没做表情时的脸的信息，b 表示是 eq7 中 A 的部分，即摄像机前人脸发生表情形变的矩阵信息

This problem can be rewritten in the canonical least-squares form by substitution:

$\|C - A\tilde{x}\|$

$$E(\boldsymbol{\delta}^T) = \left|\left|\mathbf{A}\boldsymbol{\delta}^T - \boldsymbol{b}\right|\right|_2^2 . \qquad (8)$$

The matrix $\mathbf{A} \in \mathbb{R}^{6|\boldsymbol{F}|\times76}$ is constant and contains the edge information of the template mesh projected to the expression sub-space. Edge information of the target in neutral

A: 取自 target的未形变部分

b: 取自 source deformation.

仅在

exp部分

所以无需做对应part

现在存在的问题是：

源和目标是否需要在初始状态保持一样的状态？即参考文献中，马和骆驼虽然品种不同，但最初都是保持站直的姿势，而我们 f2f 的部分里，两张脸虽然样子不同，是否都需要处在无表情的初始状态才能保证形变的合理性？这部分我还没有查证，如果可行，我觉得这是可

以实现一种表情移植的方式。