1 Facial Mesh Alignment

In order to create blendshapes from 3D facial scans, movement within scans should be limited to facial movement with as little head movement as possible. If head movement remains within the scans, then this will be reflected in the principal axis produced by Principle Components Analysis (PCA) and as head motion is independent from speech, this is undesirable. Initial data capture can aim to minimise subject head movement, however this cannot be completely eliminated. To remove head motion from the dataset, the scans can be brought into alignment relative to landmark positions on the face which do not move during speech. Such landmarks include the nose, corners of the eyes and cheek bones, as these points which will be stationary during speech. By aligning the meshes based on the variation in these points, the head position of the subjects can be constrained while the mouths are not. To achieve this, Procrustes Analysis can be used.

1.1 Procrustes Analysis

Procrustes Analysis is a statistical shape analysis method used to analyse the differences between two objects [1]. Given matrices $\boldsymbol{X} \in \mathbb{R}^{(n \times f)}$ and $\boldsymbol{Y} \in \mathbb{R}^{(n \times f)}$, each representing the coordinates of n points in an f-dimensional feature space, a common statistical difference metric is the sum of square differences 1.

$$D = \sum_{i=1}^{n} \sum_{j=1}^{f} (x_{ij} - y_{ij})^2 \tag{1}$$

However, two objects which are mathematically similar can still have a significant difference when they are scaled differently and are in different positions, orientations in space. In order to compare two objects regardless of these factors, the differences in orientation, scale and spacial position of the objects must be minimised by bringing the objects into optimal alignment.

1.1.1 Translational Alignment

Translational components can be eliminated by having the centroids of the two objects lie at the same position in space. This can be easily achieved by subtracting the mean of each objects points from itself such that the centroid now lies at the origin.

Let $\bar{x}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij}$ and $\bar{y}_j = \frac{1}{n} \sum_{i=1}^{n} y_{ij}$ where $(j = 1, \ldots, f)$ represent the mean of each feature, such that the centroids of the two objects are given by $C_X = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_f)$ and $C_Y = (\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_f)$ respectively. By subtracting the mean of $\boldsymbol{X}$ and $\boldsymbol{Y}$ from themselves, the centroids C_X and C_Y will be in alignment at the origin, minimising the sum of square differences due to translational components.

1.1.2 Scaling Alignment

Similarly, scaling components can be removed by rescaling the objects by the root mean square distance (2), such that the root mean square distance will be of unit distance for both.

$$D = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{f} x_{ij}^2} \tag{2}$$

1.1.3 Rotational Alignment

In order to align the objects in the same orientation, a rotational matrix which is able to map $\boldsymbol{X}$ to $\boldsymbol{Y}$ as closely as possible must be found. This is described as the Orthogonal Procrustes problem and can be solved with Singular Value Decomposition (SVD). Given matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, the orthogonal matrix $\boldsymbol{R}$ is the matrix which matches $\boldsymbol{A}$ to $\boldsymbol{B}$ as closely as possible, as described by equation (3), where $\|\cdot\|_F$ is the Fobius norm.

$$\boldsymbol{R} = \arg\min_{\boldsymbol{\Omega}} \|\boldsymbol{\Omega A} - \boldsymbol{B}\|_F \quad \text{subject to} \quad \boldsymbol{\Omega}^\top \boldsymbol{\Omega} = \boldsymbol{I} \tag{3}$$

The product of matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ can be decomposed (5), and then the rotational matrix $\boldsymbol{R}$ can be expressed (6).

$$\boldsymbol{M} = \boldsymbol{B}\boldsymbol{A}^\top \tag{4}$$

$$\boldsymbol{M} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top \tag{5}$$

$$\boldsymbol{R} = \boldsymbol{U}\boldsymbol{V}^\top \tag{6}$$

1.1.4 A Simple Application of Procrustes Analysis

This section follows the steps of object alignment using Procrustes Analysis as described above with a simple example. A pair of configurations are given by the matrices below, visualised in figure 1a.

$$\boldsymbol{X} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 3 & 2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{Y} \begin{pmatrix} -3 & -2 \\ -3 & -4 \\ -7 & -4 \end{pmatrix}$$

Before applying alignment, the sum of square differences between $\boldsymbol{X}$ and $\boldsymbol{Y}$ is

$$D = (1+3)^2 + (1+2)^2 + (1+3)^2 + (2+4)^2 + (3+7)^2 + (2+4)^2$$
$$= 213$$

As described in section (1.1.1), the initial objects are translated so that their centroids lie at the origin (Figure 1b). This is then followed by scaling alignment and finally rotational alignment (Figures 1c, 1d). After alignment, the difference is zero as these objects are both similar triangles.

## 1.2 Procrustes Analysis for Facial Mesh Alignment

The process of Procrustes Analysis described above deals with objects in which all coordinate points are used to find the optimum alignment, however this is inappropriate for facial mesh alignment as not only would the position of the head be aligned to, but as would the motion from speech, of which the aim is to maintain variation. However, a subset of points of a pair of objects can be used to align the entire object. By selecting
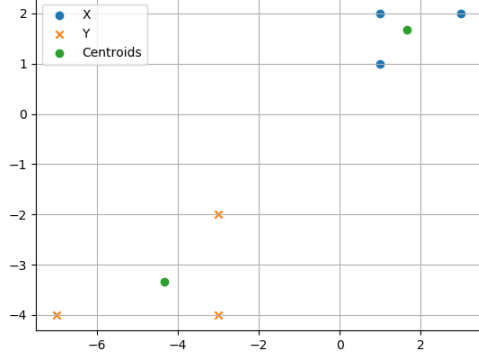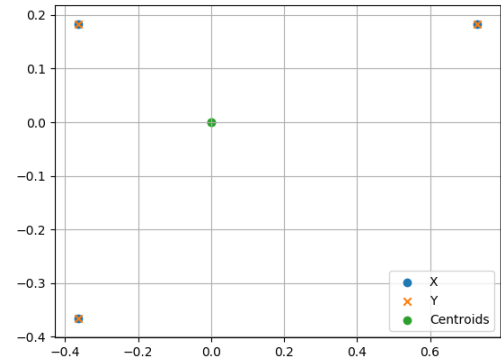
(a) Original coordinates
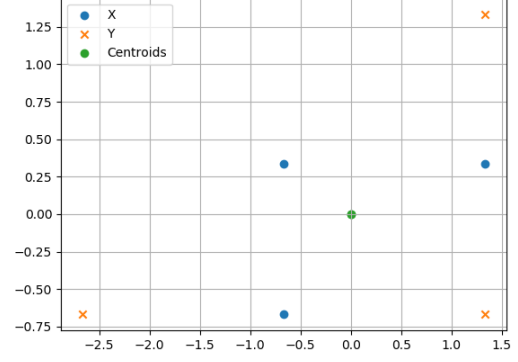
(b) Translation
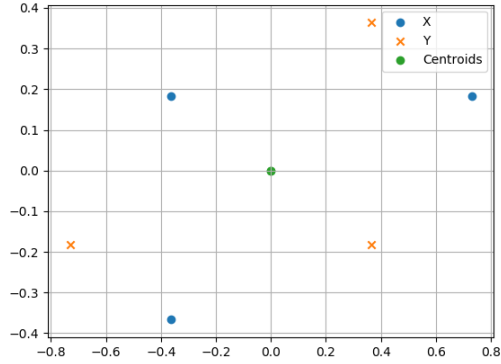
(c) Scaling

(d) Rotation

Figure 1: Application of Procrustes Analysis

points which do not contain any movement due to speech or facial expression, variation in these points can be assumed to be only from head motion. Such positions include the nose, corners of the eyes and cheekbones. If the facial scans capture 360 degrees, then points on the back of the subjects head can also be used. From this subset of points, the translation, scaling and rotational matrix can be found which aligns these points using the steps described in section 1.1 but then applied to the entire object.

References

[1] W. J Krzanowski. *Principles of multivariate analysis : a user's perspective.* Oxford statistical science series. Clarendon Press, Oxford, revised edition. edition, 2000. pages 1