






# Online Avatar Motion Adaptation to Morphologically-similar Spaces

Soojin Choi<sup>1</sup> , Seokpyo Hong<sup>2</sup> , Kyungmin Cho<sup>3</sup> , Chaelin Kim<sup>1</sup> , Junyong Noh<sup>1</sup> 

<sup>1</sup>Visual Media Lab, KAIST    <sup>2</sup>Samsung Advanced Institute of Technology    <sup>3</sup>Anigma Technologies

## 1. Motion Update Procedure

Upon the completion of the procedural optimization, we employ an element shuffle function,  $\mathcal{S}(\mathbf{x})$ , in the implementation level, to extract and rearrange the elements of the displacement  $\mathbf{x}$  according to the optimization order. When we update the avatar motion, we convert the rearranged displacements,  $(\mathbf{x}_R^t, \mathbf{x}_S^t, \mathbf{x}_L^t)$ , back to the original form of  $\mathbf{x}$  using the inverse form of  $\mathcal{S}()$ ,  $\mathcal{S}^{-1}(\mathbf{x}_R^t, \mathbf{x}_S^t, \mathbf{x}_L^t) = \mathbf{x}$ .

The resulting avatar motion  $\mathcal{M}_{tar}$  is obtained by performing the procedural updates of the user motion as follows:

$$\begin{aligned} \mathcal{M}_{init} &= \mathcal{M}_{src} \oplus (\hat{\mathbf{p}}_0^p - \mathbf{p}_0, \ln(\hat{\mathbf{q}}_0^p \mathbf{q}_0^{-1}), \mathbf{0}_{1 \times (3 \times (N_f - 1))}), \\ \mathcal{M}_{tar} &= ((\mathcal{M}_{init} \oplus \mathcal{S}_R^{-1}(\mathbf{x}_R)) \oplus \mathcal{S}_S^{-1}(\mathbf{x}_S)) \oplus \mathcal{S}_L^{-1}(\mathbf{x}_L), \end{aligned}$$

where  $\mathbf{x}_R = \mathcal{S}_R(\mathbf{x})$ ,  $\mathbf{x}_S = \mathcal{S}_S(\mathbf{x})$ , and  $\mathbf{x}_L = \mathcal{S}_L(\mathbf{x})$  are the element shuffle functions for each optimization variable in the implementation level. In the only example of the inverse form,  $\mathcal{S}_R^{-1}(\mathbf{x}_R) = \mathbf{x}$ ,  $\mathbf{x}$  has valid values of  $\mathbf{x}_R$  only in the original displacement form, while the remaining elements are zero.

## 2. Sub-matrices in Eq (1)

In this section, we explain the sub-matrices of the matrix used in Eq (1) of the main paper in detail.

$$\begin{bmatrix} \alpha & r_{1,2} & \cdots & r_{1,N_C} & 1 & \mathbf{s}_1^T \\ r_{2,1} & \alpha & \cdots & r_{2,N_C} & 1 & \mathbf{s}_2^T \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ r_{N_C,1} & r_{N_C,2} & \cdots & \alpha & 1 & \mathbf{s}_{N_C}^T \\ 1 & 1 & \cdots & 1 & 0 & \mathbf{0}_{1 \times 3} \\ \mathbf{s}_1 & \mathbf{s}_2 & \cdots & \mathbf{s}_{N_C} & \mathbf{0}_{3 \times 1} & \mathbf{0}_{3 \times 3} \end{bmatrix} \begin{bmatrix} \mathbf{w}_1^T \\ \mathbf{w}_2^T \\ \vdots \\ \mathbf{w}_{N_C}^T \\ \mathbf{w}_{N_C+1}^T \\ \mathbf{w}_{N_C+2}^T \\ \mathbf{w}_{N_C+3}^T \\ \mathbf{w}_{N_C+4}^T \end{bmatrix} = \begin{bmatrix} \mathbf{t}_1^T \\ \mathbf{t}_2^T \\ \vdots \\ \mathbf{t}_{N_C}^T \\ \mathbf{0}_{1 \times 3} \\ \mathbf{0}_{1 \times 3} \\ \mathbf{0}_{1 \times 3} \\ \mathbf{0}_{1 \times 3} \end{bmatrix} \quad (1)$$

For the full derivation of Eq (1), please refer to Eberly [Ebe18] and Keller and Borkowski [KB19]. Thin plate spline minimizes a smoothed bending energy [KB19]. The minimization can be formulated as a form of Euler-Lagrange differential equation [Ebe18], and it is generally solved by exploiting the Green's function that has a radial basis kernel,  $r_{i,j} = \|\mathbf{s}_i - \mathbf{s}_j\|^2 \ln \|\mathbf{s}_i - \mathbf{s}_j\|$ , in our case. Given source and target points, a matrix form of the differential equation can be written as follows:

$$\begin{aligned} \mathbf{T} &= (\mathbf{G} + \alpha \mathbf{I}) \mathbf{W}_1 + \mathbf{S}' \mathbf{W}_2, \\ \text{where } \mathbf{G} &= \begin{bmatrix} 0 & r_{1,2} & \cdots & r_{1,N_C} \\ r_{2,1} & 0 & \cdots & r_{2,N_C} \\ \vdots & \vdots & \ddots & \vdots \\ r_{N_C,1} & r_{N_C,2} & \cdots & 0 \end{bmatrix}_{N_C \times N_C}, \\ \mathbf{W}_1 &= \begin{bmatrix} \mathbf{w}_1^T \\ \vdots \\ \mathbf{w}_{N_C}^T \end{bmatrix}_{N_C \times 3}, \quad \mathbf{W}_2 = \begin{bmatrix} \mathbf{w}_{N_C+1}^T \\ \vdots \\ \mathbf{w}_{N_C+4}^T \end{bmatrix}_{4 \times 3}. \end{aligned} \quad (2)$$

Here,  $\mathbf{G}$  is composed of the basis functions with given source points.  $\alpha$  is a smoothing factor that is set to 0 in our case.  $\mathbf{W}_1$  and  $\mathbf{W}_2$  are unknown matrices.  $\mathbf{S}' \in \mathbb{R}^{N_C \times 4}$  is a matrix consisting of row vectors,  $[1 \quad \mathbf{s}_i^T]$ ,  $i = 1, 2, \dots, N_C$ . The additional dimension by 1 in  $\mathbf{S}'$  makes each row a basis of the null space satisfying the condition of orthogonality [KB19] as shown below:

$$\mathbf{S}'^T \mathbf{W}_1 = \mathbf{0} \quad (3)$$

Simply concatenating Eq (2) and Eq (3) leads to the same linear form as Eq (1) as follows:

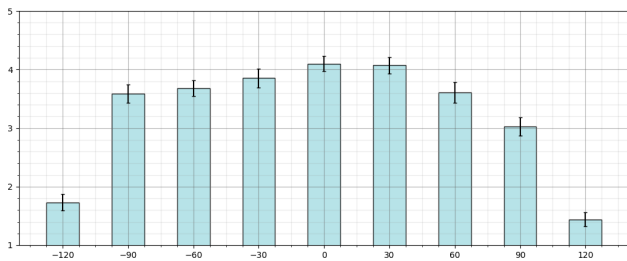
$$\begin{bmatrix} \mathbf{G} + \alpha \mathbf{I} & \mathbf{S}' \\ \mathbf{S}'^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{T} \\ \mathbf{0} \end{bmatrix} \quad (4)$$

## 3. Qualitative Evaluation

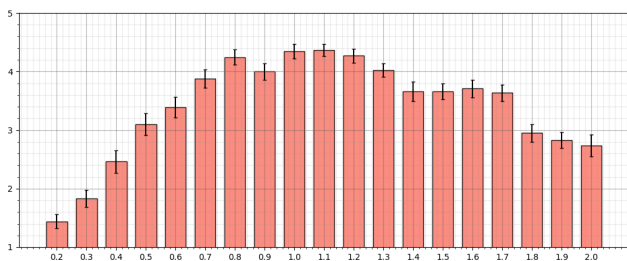
We conducted two experiments to find out the operable range of our method. The first experiment was to identify the allowed orientation difference between the corresponding remote and local object before the resulting motions become unacceptable. We created 9 scenarios by rotating the remote chair with respect to the corresponding local chair from 30° to 120° both in clockwise and counter-clockwise directions with an increment of 30°. The second experiment was to identify the allowed size difference between two spaces. We created 19 remote spaces by scaling the size of the remote space to be -80% to +100% with an increment of 10%, where 0% has the same size as the user space. The remote space of -90% was too small to include a single object and therefore excluded from the experiment.

Forty-one participants were recruited (27 males and 14 females) with an age range from 21 to 42 (average age: 29.2). The users were instructed to watch the videos that show the resulting avatar motion for each scene and asked to score on a 5-Likert Scale about the naturalness of the overall motion, where 1, 3, and 5 represent 'very unnatural', 'neutral', and 'very natural', respectively.

For statistical analysis, we applied the Kruskal-Wallis rank sum test to both experiments. For the first experiment, the average rating scores became smaller as the orientation difference between two objects becomes larger as shown in Figure 1a. There existed significant differences at  $-120^\circ$  and  $+90^\circ$  and above ( $p < 0.05$ ). For the second experiment, the score dropped more sharply when the remote room becomes smaller than when it becomes larger as shown in Figure 1b. There existed significant differences at  $-50\%$  and below and  $+80\%$  and above ( $p < 0.05$ ). Based on this analysis, we concluded that our method synthesizes visually-plausible motions within the ranges of  $-90^\circ$  to  $+60^\circ$  in terms of the orientation difference between two objects and  $-40\%$  to  $+70\%$  in terms of the size difference between two spaces.



(a) Average ratings on naturalness with respect to the orientation differences between two objects. The error bar indicates the standard error of the mean.



(b) Average ratings on naturalness with respect to the size difference between two spaces. The error bar indicates the standard error of the mean.

**Figure 1:** Average rating scores for the orientation and size differences between two objects and spaces, respectively.

## References

- [Ebe18] EBERLY: Thin-plate spline. geometric tools. <https://www.geometrictools.com/>, 2018. 1
- [KB19] KELLER W., BORKOWSKI A.: Thin plate spline interpolation. *Journal of Geodesy* 93, 9 (2019), 1251–1269. 1