Table 1: The list of video sources

Jennifer Lawrence Shares Her Most Embarrassing Moments https://www.youtube.com/watch?v=4OTQV48qBoY
President Barack Obama: Between Two Ferns with Zach Galifianakis https://www.youtube.com/watch?v=UnW3xkHxIEQ
Moon Jae-in and Kim Jong-un address 150,000 crowd | Al Jazeera English https://www.youtube.com/watch?v=NXyQSvbKGIM
Mark Zuckerberg : How to Build the Future https://www.youtube.com/watch?v=Lb4IcGF5iTQ
How To Flirt Via TEXT Message | 10 Texting Tips https://www.youtube.com/watch?v=fWVqk7FTcnw
Sam Altman : How to Build the Future https://www.youtube.com/watch?v=sYMqVwsewSg
Broadsheet Owner's Exclusive Interview | Why General Malek Wanted Corrupt Commission From Him https://www.youtube.com/watch?v=c55-hlzbc8k
WATCH: Joe Biden gives first speech as president https://www.youtube.com/watch?v=cTtKDN4LgL8
Putin wishes Russians brighter New Year, 'return to normal' https://www.youtube.com/watch?v=1kn5JBEHzD4
5 Simple Journalist Techniques for Effective Interviews https://www.youtube.com/watch?v=NWDL_UYMc7Q
To Be Honest 2.0 | Mathira| Tabish Hashmi | Full Episode | NashpatiPrime https://www.youtube.com/watch?v=nhMaY17m11w
Elon Musk : How to Build the Future https://www.youtube.com/watch?v=tnBQmEqBCY0
Dwight's Acceptance Speech -The Office https://www.youtube.com/watch?v=ilE4lr9Qb3A

A APPENDIX

A.1 Sparse control points warping

Figure 10(b) shows the case of only using a set of sparse grid points to represent the face boundary. In order to obtain reasonable results, besides the line-bending energy term E_l and regularization energy term E_r in Section 5.2, a distortion term E_d is further added. This energy term makes the triangle formed by the contour points and their surrounding points as identical as possible in shape and area with respect to the corresponding triangle formed by the grid points before reshaping, thus makes the boundary to be more coherent. Let Γ and γ be the greatest and the least singular value of the triangle deformation matrix, respectively. The energy term proposed first by Sander et al. [19] can be described as:

$$E_d = ||\Gamma|| + \sqrt{\frac{1}{2}(\Gamma^2 + \gamma^2)}. \tag{1}$$

This equation requires that the triangular area and shape do not change significantly. To further constrain the rotation and shape change of the triangle, we add a similarity energy term E_s [6] as:

$$E_s = \sum_{e \in M_u} ||(\mathbf{v}_0' - \mathbf{v}_1') - \frac{||\mathbf{v}_0 - \mathbf{v}_1||}{||\mathbf{v}_2 - \mathbf{v}_1||} \mathbf{R}_\theta (\mathbf{v}_2' - \mathbf{v}_1')||^2, \tag{2}$$

where $e = (\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2)$ is the triangle in the grid M_u before reshaping, and $e' = (\mathbf{v}_0', \mathbf{v}_1', \mathbf{v}_2')$ is the corresponding triangle of e in the grid M_u with $\mathbf{R}_\theta$ being the 2×2 rotation matrix which rotates the edge $(\mathbf{v}_2 - \mathbf{v}_1)$ to $(\mathbf{v}_0 - \mathbf{v}_1)$. In this case, we performed a two-step optimization of the face grid. The first step uses only the same energy function as Eqn. [17], which makes the deformed grid after reshaping uniform and regular. The second optimization adds the distortion term and similarity energy term with the following energy function:

$$E = w_l E_l + w_r E_r + w_d E_d + w_s E_s. \tag{3}$$

A.2 List of video sources

The used video sources are listed in Table 1.