

Corrective 3D Reconstruction of Lips from Monocular Video

Supplemental Document

Pablo Garrido¹ Michael Zollhöfer¹ Chenglei Wu² Derek Bradley³
 Patrick Pérez⁴ Thabo Beeler³ Christian Theobalt¹

¹Max Planck Institute for Informatics

²ETH Zurich

³Disney Research

⁴Technicolor

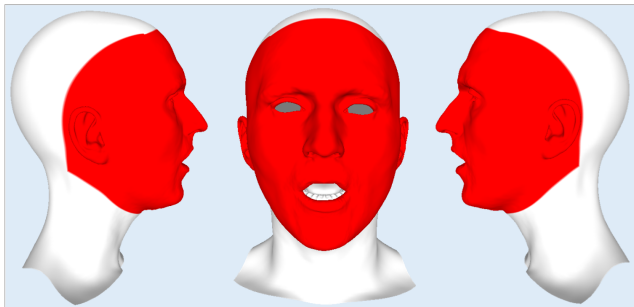


Figure 1: *Mask used for correspondence association - The red region is used for computing the correspondences between the topology of the high-quality and monocular reconstructions.*

1 Dense Correspondences

We establish dense correspondences between the topology of the high-quality and monocular reconstructions using non-rigid registration. Registration is performed using a subregion of the face, see Fig. 1 (red). Note, the subregion does not include the teeth, gums and mouth cavity.

2 Training Data

All lip motions captured for training our RBF network are given in Table 1 (S1), Table 2 (S2), Table 3 (S3) and Tab. 4 (S4). The motions were performed sequentially, as listed in the tables (please refer to the accompanying additional video for more details). To train the generalization regressor (**GR**), we stacked the captured data of three or four of the subjects (three if the test subject is one out of the four) in the following order: S2, S3, S1, S4. The multiple subject regressor **MS** uses the same order.

3 Additional Results

Fig. 2 shows additional results on our test subjects. Note that our regression approach better captures bending, stretching, and rolling than the state-of-the-art monocular face tracker of Garrido et al. [2016]. We can even reconstruct extreme lip deformations if state-of-the-art monocular tracking fails. For the complete sequences we refer to the accompanying videos.

References

GARRIDO, P., ZOLLHÖFER, M., CASAS, D., VALGAERTS, L., VARANASI, K., PÉREZ, P., AND THEOBALT, C. 2016. Reconstruction of personalized 3D face rigs from monocular video. *ACM TOG* 35, 3, 28:1–28:15.

Table 1: *Performed Lip Motions - Subject S1.*

Action	Features	# Frames
Roll lips	Outwards; mouth half-open	48
Roll lips	Outwards; mouth open	28
Move lips	Both up; mouth closed	14
Move lip	Lower lip down; mouth closed	24
Move lip	Upper lip up; mouth closed	20
Pull corners	Mouth closed	43
Pull corner	Left; mouth closed	29
Pull corner	Right; mouth closed	19
Open	Mouth closed to half-open	34
Open	Mouth half-open to wide-open	54
Kiss	Mouth closed	37
Smile	Mouth closed	23
Smile	Mouth half-open	38
Smile	Mouth open	23
Move left	Mouth closed	31
Move left	Mouth half-open	21
Move left	Mouth open	32
Move right	Mouth closed	34
Move right	Mouth half-open	26
Move right	Mouth open	34
Pout	Mouth closed	29
Roll lips	Inwards; mouth closed	25
Roll lips	Inwards; mouth half-open	36
Roll lips	Inwards; mouth open	34
Roll lips	Outwards; mouth closed	31
Viseme TH	Exaggerated; mouth half-open	17
Viseme I:	Exaggerated; mouth closed to half-open	18
Viseme OU	Exaggerated; mouth closed to open	32
Viseme F	Exaggerated; mouth closed	15
Viseme SH	Exaggerated; mouth closed	16
Viseme M	Exaggerated; mouth half-open to closed	32
Viseme OO	Exaggerated; mouth closed to half-open	21
Viseme AE	Exaggerated; mouth half-open to open	24

Table 2: Performed Lip Motions - Subject S2.

Action	Features	# Frames
Roll lips	Outwards to inwards; mouth closed	51
Roll lips	Outwards to inwards; mouth half-open	47
Roll lips	Inwards to outwards; mouth open	39
Smile	Mouth closed	17
Smile	Mouth half-open	49
Smile	Mouth open	51
Move left	Mouth closed	30
Move right	Mouth closed	32
Move left	Mouth half-open	30
Move right	Mouth half-open	26
Move left	Mouth open	28
Move right	Mouth open	21
Bite	Lower lip; mouth closed	26
Open	Mouth closed to half-open	76
Kiss	Mouth closed	33
Round lips	Mouth closed	27
Round lips	Mouth half-open to open	74
Move lip	Lower lip down; mouth closed	30
Move lip	Upper lip up; mouth closed	31
Pull corners	Mouth closed	33

Table 3: Performed Lip Motions - Subject S3.

Action	Features	# Frames
Kiss	Mouth closed	25
Smile	Mouth closed	32
Smile	Mouth half-open	25
Smile	Mouth open	27
Move left	Mouth closed	18
Move left	Mouth half-open	17
Move left	Mouth open	27
Move right	Mouth closed	17
Move right	Mouth half-open	24
Move right	Mouth open	18
Pout	Mouth closed	25
Roll lips	Inwards; mouth closed	15
Roll lips	Inwards; mouth half-open	20
Roll lips	Inwards; mouth open	25
Roll lips	Outwards; mouth closed	20
Roll lips	Outwards; mouth half-open to open	76
Move lips	Both up; mouth closed	13
Move lip	Lower lip down; mouth closed	22
Move lip	Upper lip up; mouth closed	21
Pull corners	Mouth closed	21
Pull corner	Right; mouth closed	7
Pull corner	Left; mouth closed	18
Open	Mouth half-open to open	34
Open	Mouth open to wide open	30
Open	Mouth closed to half-open	42
Viseme F	Exaggerated; mouth half-open to closed	22
Viseme SH	Exaggerated; mouth half-open to closed	18
Viseme M	Exaggerated; mouth half-open to closed	18
Viseme TH	Exaggerated; mouth half-open to open	12
Viseme OO	Exaggerated; mouth half-open	13
Viseme I:	Exaggerated; mouth half-open	14
Viseme AE	Exaggerated; mouth half-open to open	19
Viseme OU	Exaggerated; mouth half-open to open	16

Table 4: Performed Lip Motions - Subject S4.

Action	Features	# Frames
Kiss	Mouth closed	18
Smile	Mouth closed	22
Smile	Mouth half-open	31
Smile	Mouth open	31
Move left	Mouth closed	18
Move left	Mouth half-open	22
Move left	Mouth open	26
Move right	Mouth closed	18
Move right	Mouth half-open	31
Move right	Mouth open	24
Pout	Mouth closed	22
Roll lips	Inwards; mouth closed	28
Roll lips	Inwards; mouth half-open	25
Roll lips	Inwards; mouth open	25
Roll lips	Outwards; mouth closed	29
Roll lips	Outwards; mouth half-open	20
Roll lips	Outwards; mouth open	37
Move lips	Both up; mouth closed	24
Move lip	Lower lip down; mouth closed	20
Move lip	Upper lip up; mouth closed	17
Pull corners	Mouth closed	19
Pull corner	Left; mouth closed	19
Pull corner	Right; mouth closed	25
Open	Mouth closed to half-open	50
Open	Mouth half-open to wide-open	48
Open	Sticky lips; mouth closed to half-open	24
Viseme F	Exaggerated; mouth half-open to closed	21
Viseme SH	Exaggerated; mouth closed	23
Viseme M	Exaggerated; mouth half-open to closed	22
Viseme TH	Exaggerated; mouth half-open	21
Viseme OO	Exaggerated; mouth closed to half-open	23
Viseme I:	Exaggerated; mouth closed	17
Viseme AE	Exaggerated; mouth half-open to open	27
Viseme OU	Exaggerated; mouth closed to open	18

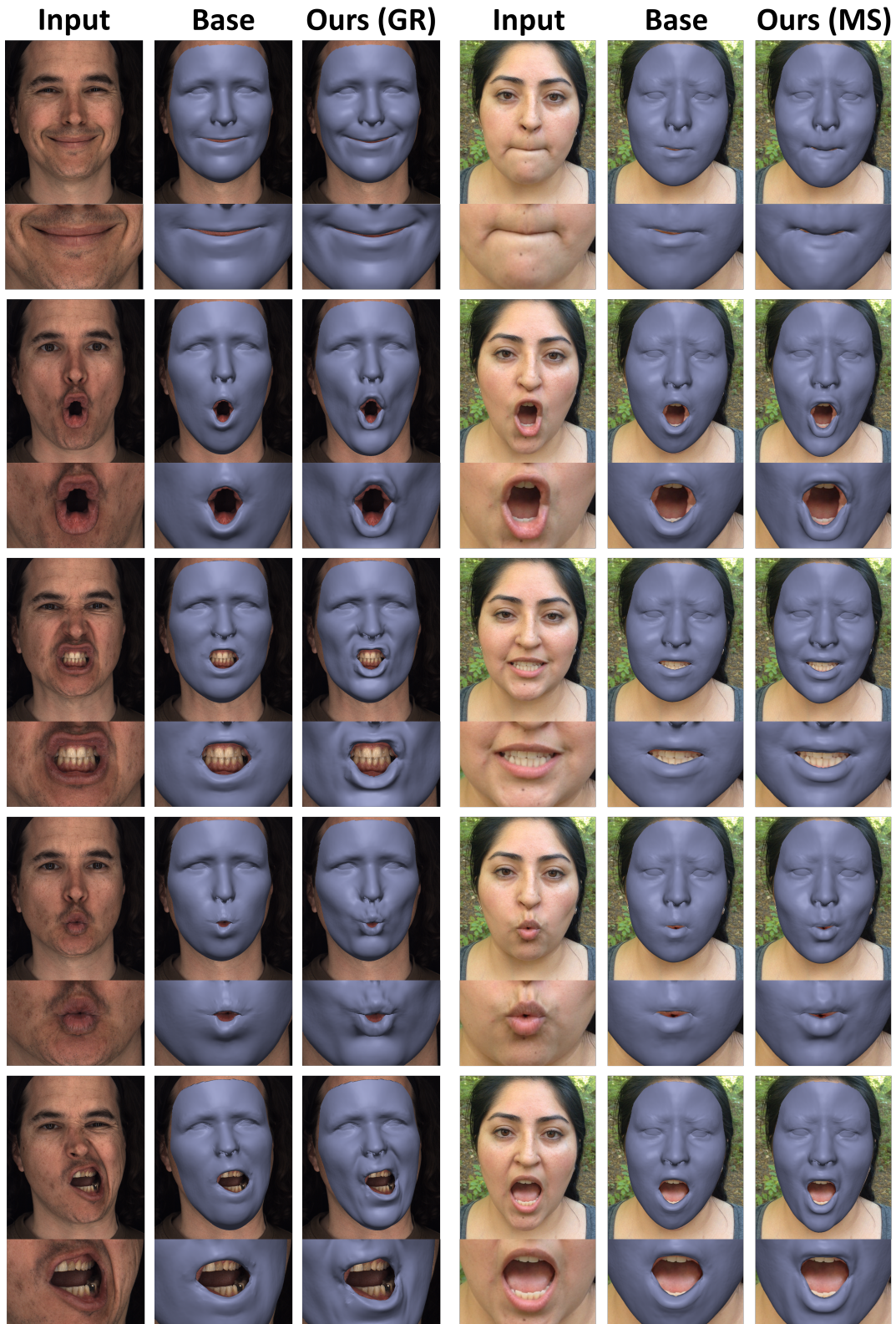


Figure 2: More indoor and outdoor results - Our proposed RBF network (**GR**) generalizes well to new subjects and arbitrary performances (left). In addition, novel scenes under general illumination can be handled (right). Compared to the state-of-the-art approach of Garrido et al. [2016], our approach better captures bending, stretching, and rolling of the lips. Note that the reconstruction of areas surrounding the lips, such as nasolabial folds, dimples and the chin are also significantly improved. (Subjects: S5 (left), S3 (right))