Perspective Projection

97° FOV (19mm)

Perspective Projection

97° FOV (19mm)

Preview: Our Result

97° FOV (19mm)

Preview: Our Result

97° FOV (19mm)

Motivation

- Perspective projection makes **unnatural look** on faces near image corners.

- **Portrait** is important to photography.
  - 44% of photos contains faces.
  - E.g., selfie, group portraits, group, etc.


Group Selfie Cam
Google Pixel 3

Trend: Wide-angle Camera are GROWING!

Make	Samsung	Google	Asus	Huawei	LG	OnePlus	Sony	Xiaomi
Flagship Model (2019)	S10 plus	Pixel 3	Zenfone 6	P30 Pro	V50 Thin Q	7 Pro	Xperia 1	Mi9
Field-of-view (FOV)	**123°**	**97°**	**125°**	**107°**	**107°**	**117°**	**130°**	**104°**
35mm equiv. focal length	12mm	19mm	11mm	16mm	16mm	13mm	10mm	17mm

- Human vision (FOV/35mm-equiv): **60°**/37.5mm

Trend: Wide-angle Camera are GROWING!

Make	Samsung	Google	Asus	Huawei	LG	OnePlus	Sony	Xiaomi
Flagship Model (2019)	S10 plus	Pixel 3	Zenfone 6	P30 Pro	V50 Thin Q	7 Pro	Xperia 1	Mi9
Field-of-view (FOV)	**123°**	**97°**	**125°**	**107°**	**107°**	**117°**	**130°**	**104°**
35mm equiv. focal length	12mm	19mm	11mm	16mm	16mm	13mm	10mm	17mm

- Human vision (FOV/35mm-equiv): **60°/**37.5mm
- Major camera phone prior to 2018 (FOV/35mm-equiv): **80°**/26mm

NOT A Lens Distortion

Lens distorted image

NOT A Lens Distortion

Lens distorted image

Corrected image (perspective projection)

Goal: Distortion-Free Wide-angle Portraits

- Makes faces look **natural** even large at field-of-view (FOV).
  - As if faces are captured from the camera center.
  - Robust to face poses, facial expressions, and occlusions.

Goal: Distortion-Free Wide-angle Portraits

- Makes faces look **natural** at large field-of-view (FOV).
  - As if faces are captured from the camera center.
  - Robust to face poses, facial expressions, and occlusions.

- Works for mobile platform.
  - **Automatic** correction without user intervention.
  - Efficient processing to achieve interactive rate.

Related Work: Global Projection

Perspective Projection

Stereographic Projection

Mercator Projection

Related Work: Content-aware Warping

- Spatially-variant radial correction [Zorin and Barr 1995].

Input

Output

Related Work: Content-aware Warping

- Optimizing global conformality [Carroll et al. 2009].

Input

[Carroll et al. 2009]

Related Work: Content-aware Warping

- Optimizing global conformality [Carroll et al. 2009].
- Selects a different virtual camera viewpoint [Tehrani et al. 2016].

Input

[Carroll et al. 2009]

Input

[Tehrani et al. 2016]

Related Work: Perspective-aware Warping

- Corrects foreshortening [Fried et al. 2016]

Input

Output

Our Idea: *Local* Stereographic Projection on Faces

Perspective projection
- Face is stretched ☹
- Background looks good ☺

Stereographic projection
- Face looks natural ☺
- Background is distorted ☹

Our result by local warp
- Face look natural ☺
- Background look good ☺

Algorithm Overview

Input (perspective projection)

Algorithm Overview

Subject segmentation using CNN
[Wadhwa et al. 2018]

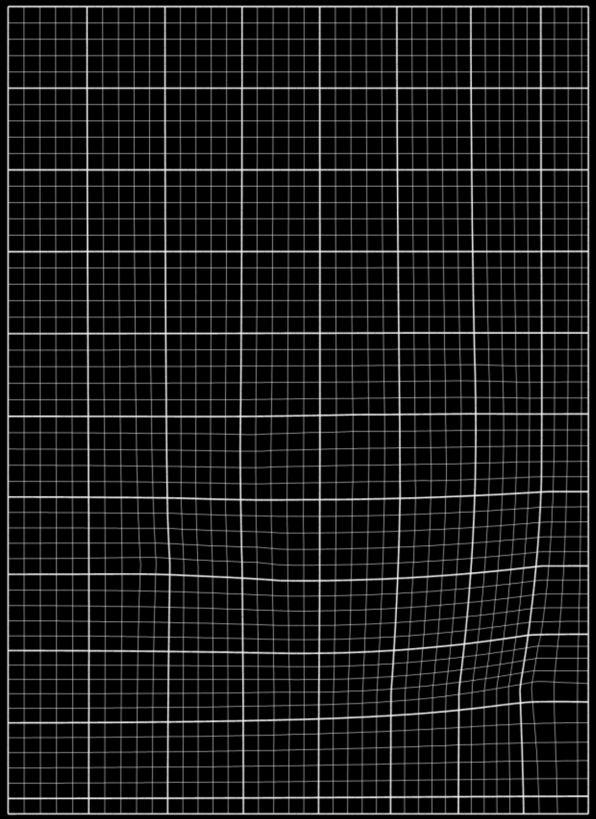
Algorithm Overview

Face detector for face region

Algorithm Overview

Input (perspective projection)

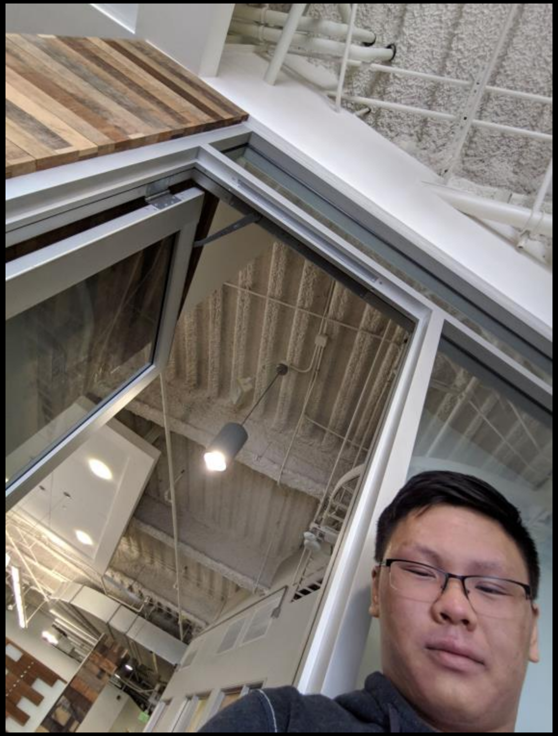
Stereographic projection



Mesh optimization

Output by warping

Naïve Mesh Blending

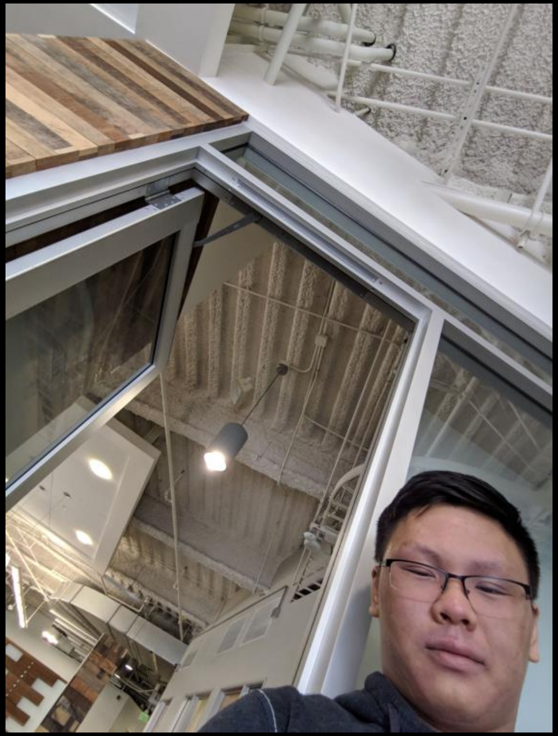


(a) Input (97° FOV)

Inset of (a)

Naïve mesh blending

Naïve Mesh Blending



(a) Input (97° FOV)

Inset of (a)

Naïve mesh blending

Optimized mesh warp

Mesh Optimization

- Defines an energy function over the vertices on the mesh.

$$\{\mathbf{v}_i^*\} = \underset{\{\mathbf{v}_i\}}{\operatorname{argmin}} E_t(\{\mathbf{v}_i\}),$$

Face Objective Term

- Each detected face k associates with a face term $E_{s,k}$,

$$E_f = \sum_k E_{s,k},$$

stereographic projection vertices.

$$E_{s,k} = \sum_{i \in \mathbf{B}_k} w_i m_i \|\mathbf{v}_i - (\mathbf{S}_k \mathbf{u}_i + \mathbf{t}_k)\|_2^2 + \lambda(\mathbf{S}_k),$$

face weight; 1/0 for face/non-face region using segmentation.

Face Objective Term

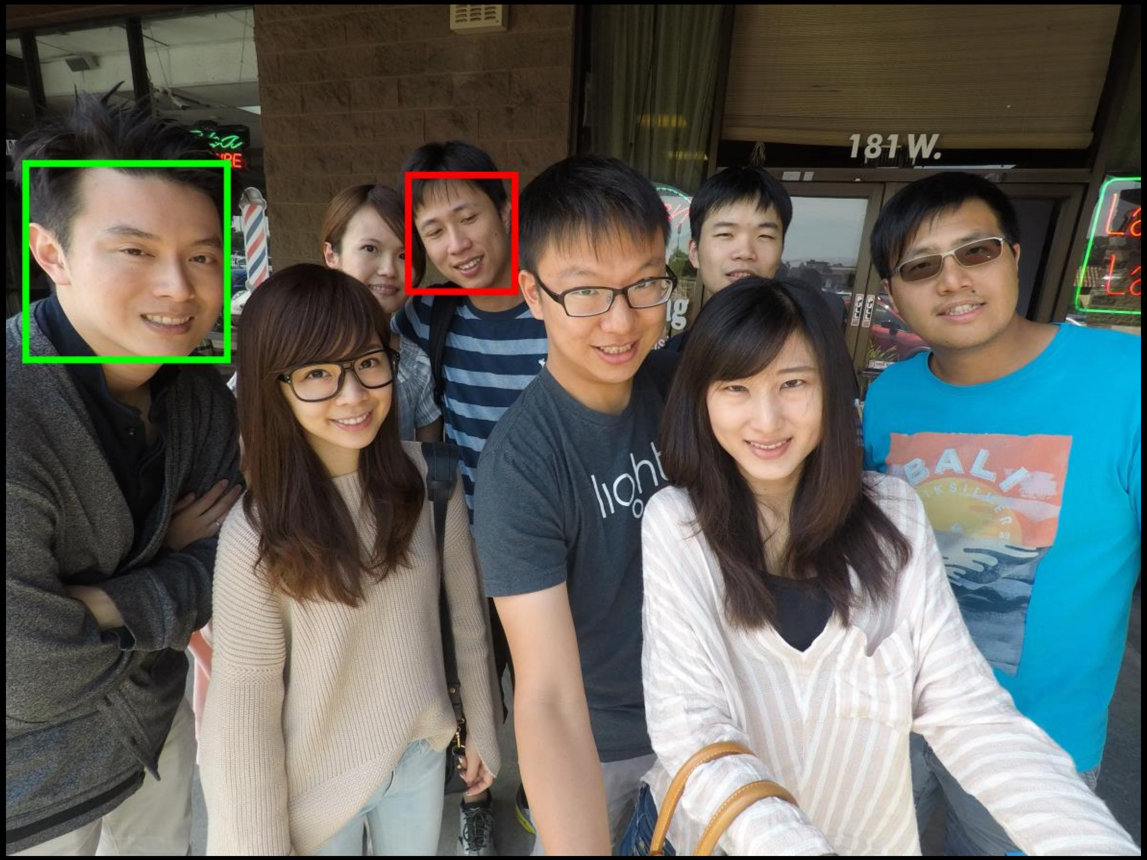
$$(1 + \exp(-(r - r_a)/r_b))^{-1}$$

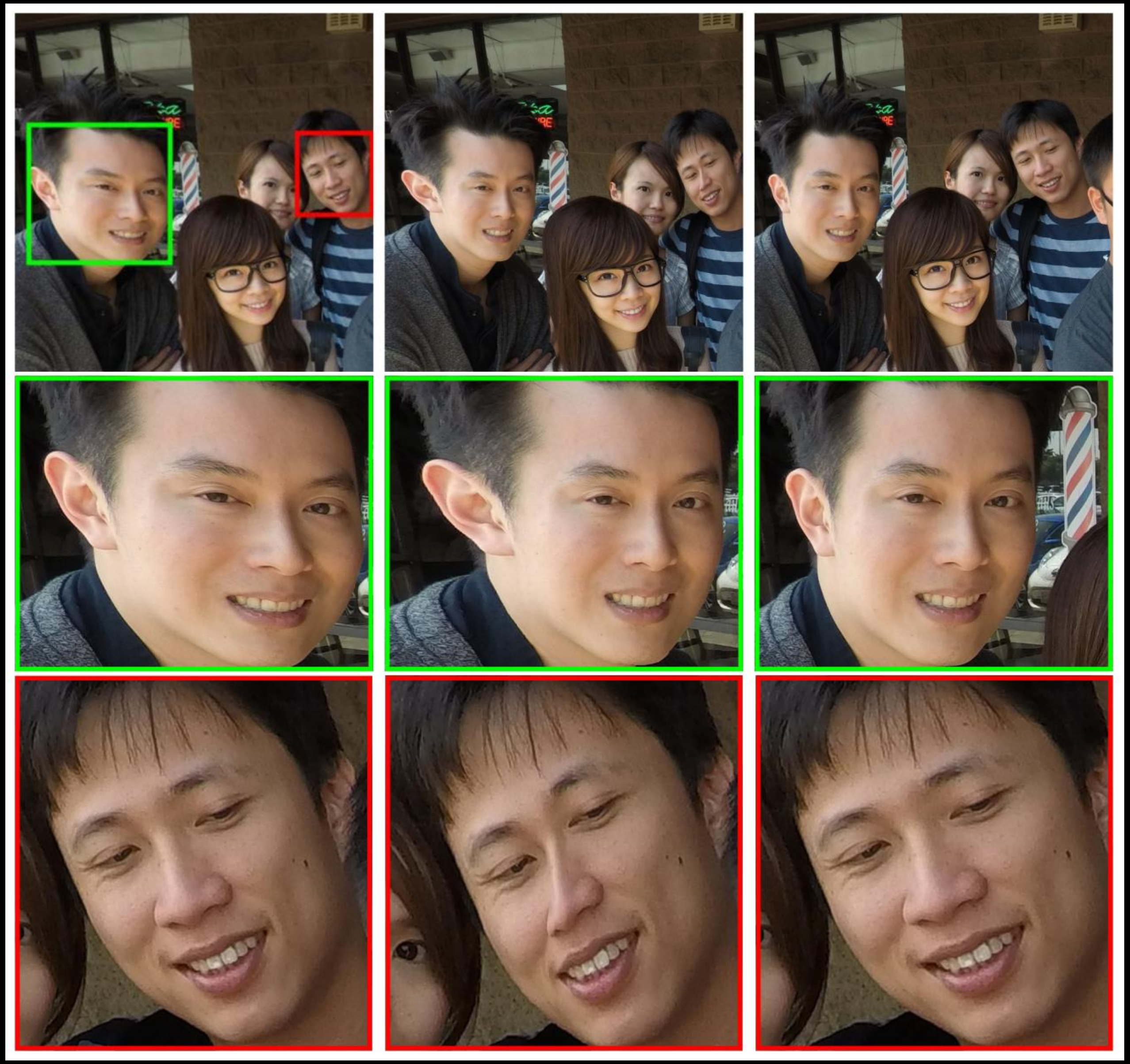
$$\lambda(\mathbf{S}_k) = w_s \|a_k - s_t\|_2^2 ,$$

$$E_{s,k} = \sum_{i \in \mathbf{B}_k} w_i m_i \|\mathbf{v}_i - (\mathbf{S}_k \mathbf{u}_i + \mathbf{t}_k)\|_2^2 + \lambda(\mathbf{S}_k) ,$$

$$\mathbf{S}_k = \begin{bmatrix} a_k & b_k \\ -b_k & a_k \end{bmatrix}$$

Benefits of Face Term

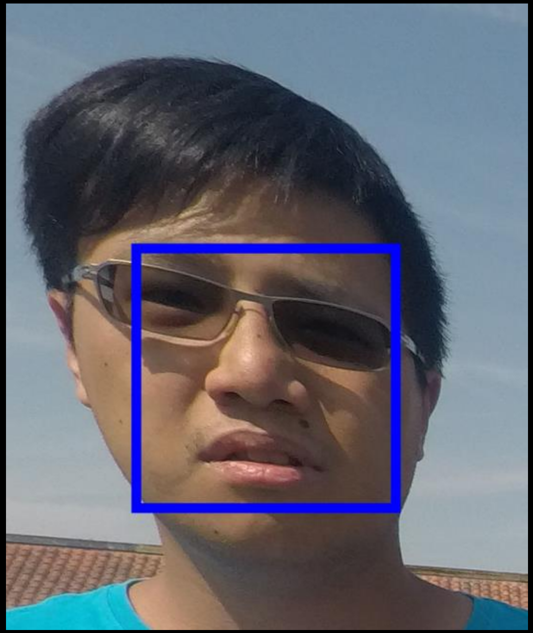




Input Without ***S, t*** Our method

Benefits of Subject Mask

(a) Input and mask



Inset of (a)

Without mask

Our method

Line Bending Terms

$$E_b = \sum_i \sum_{j \in N(i)} \|(\mathbf{v}_i - \mathbf{v}_j) \times \mathbf{e}_{ij}\|_2^2$$

$$E_r = \sum_i \sum_{j \in N(i)} \|\mathbf{v}_i - \mathbf{v}_j\|_2^2 \, .$$

(a) Input

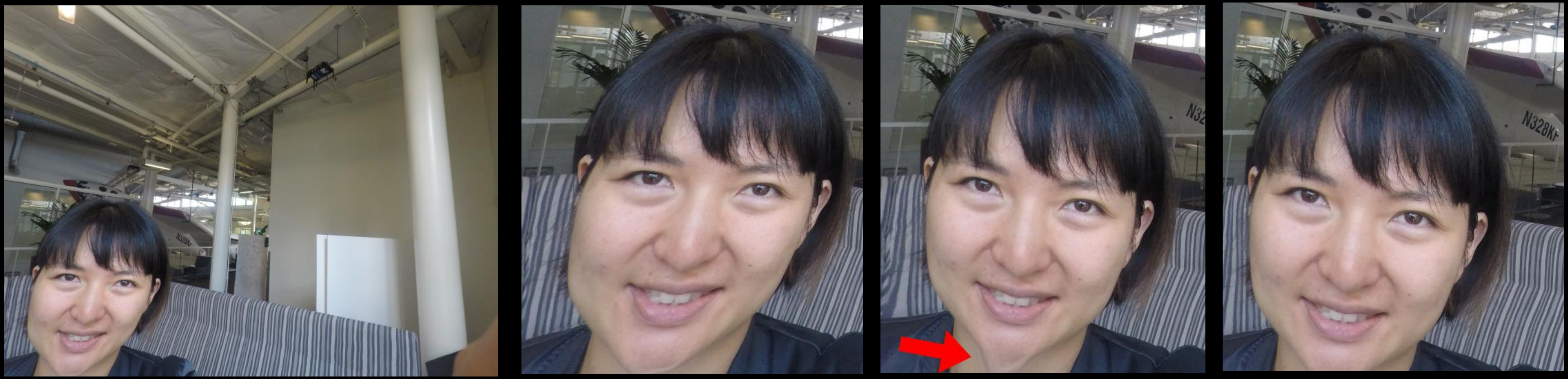
Inset of (a)

Without line bending term

Our method

Mesh Boundary Padding

- Relax boundary by padding additional vertices.



(a) Input | Inset of (a) | Without padding | Our method

Mesh Optimization

- Total energy cost by summing up
  - Face term
  - Line bending term
  - Regularization term
  - Boundary conditions
- Solves the linear energy function with Ceres.

Handling Lens Distortion

- Combines the optimal mesh with lens undistortion warp.

(a) Lens undistorted input

Our method by warping (a)

Results

- 167 photos from Flickr.

- FOVs range from 70° -120°.

- Various features:
  - Different genders, ages, facial expressions, lighting conditions.
  - Glasses, hair styles, accessories.

- Available on our project website.

Input

97°FOV
Pixel 3

Output

Input

104°FOV

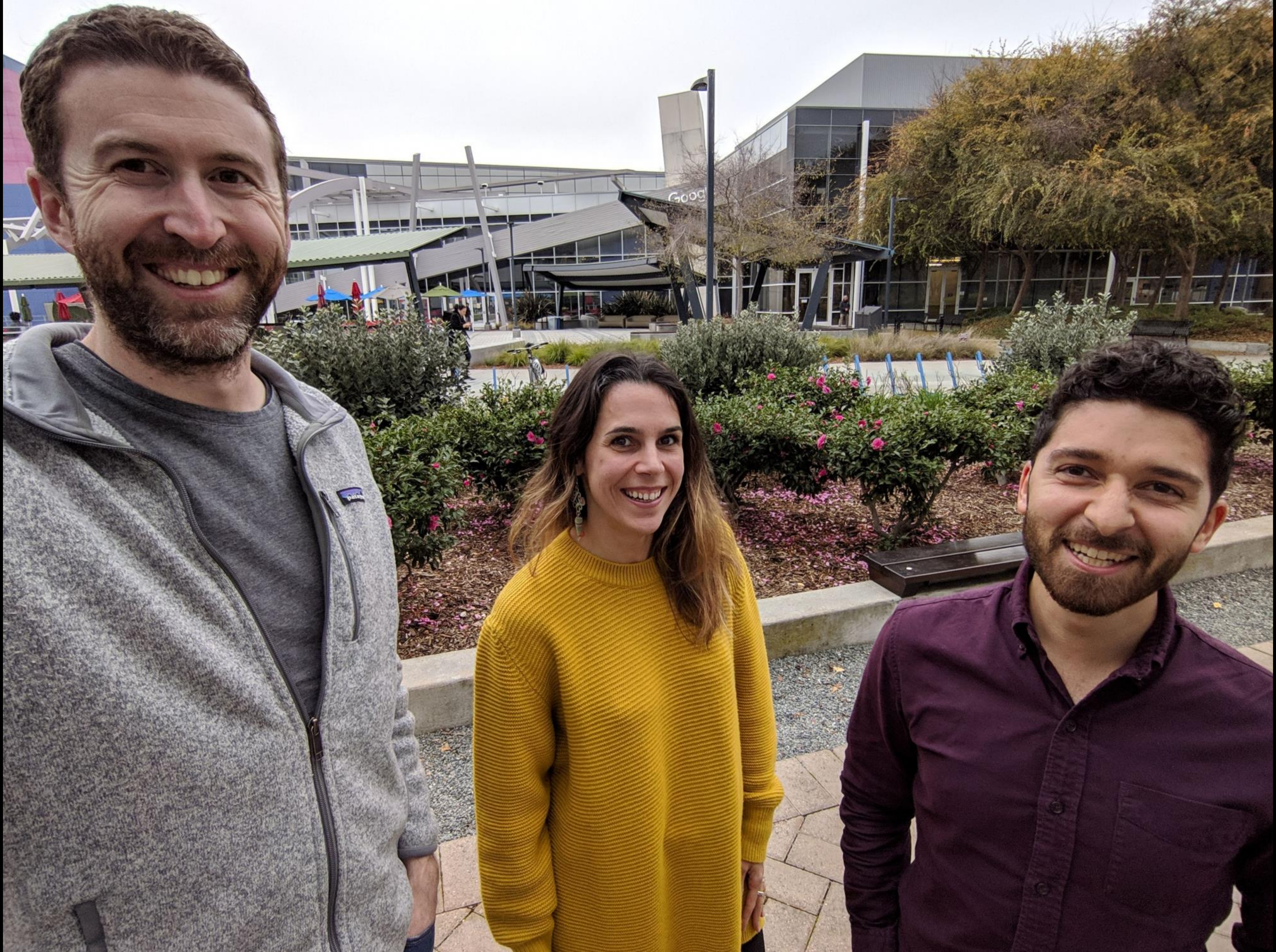
DSLR

Output

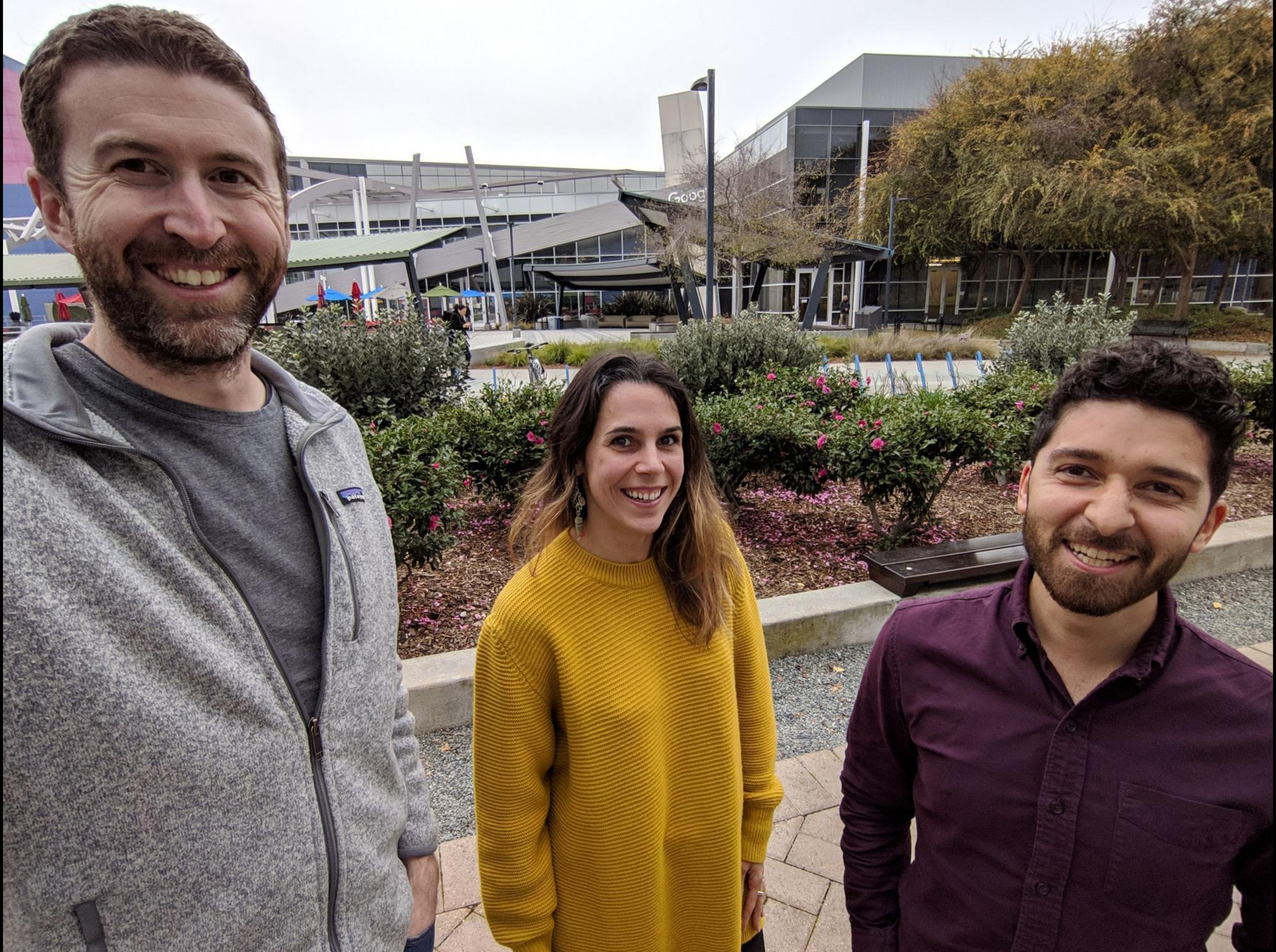
Input

97°FOV

Pixel 3



Output

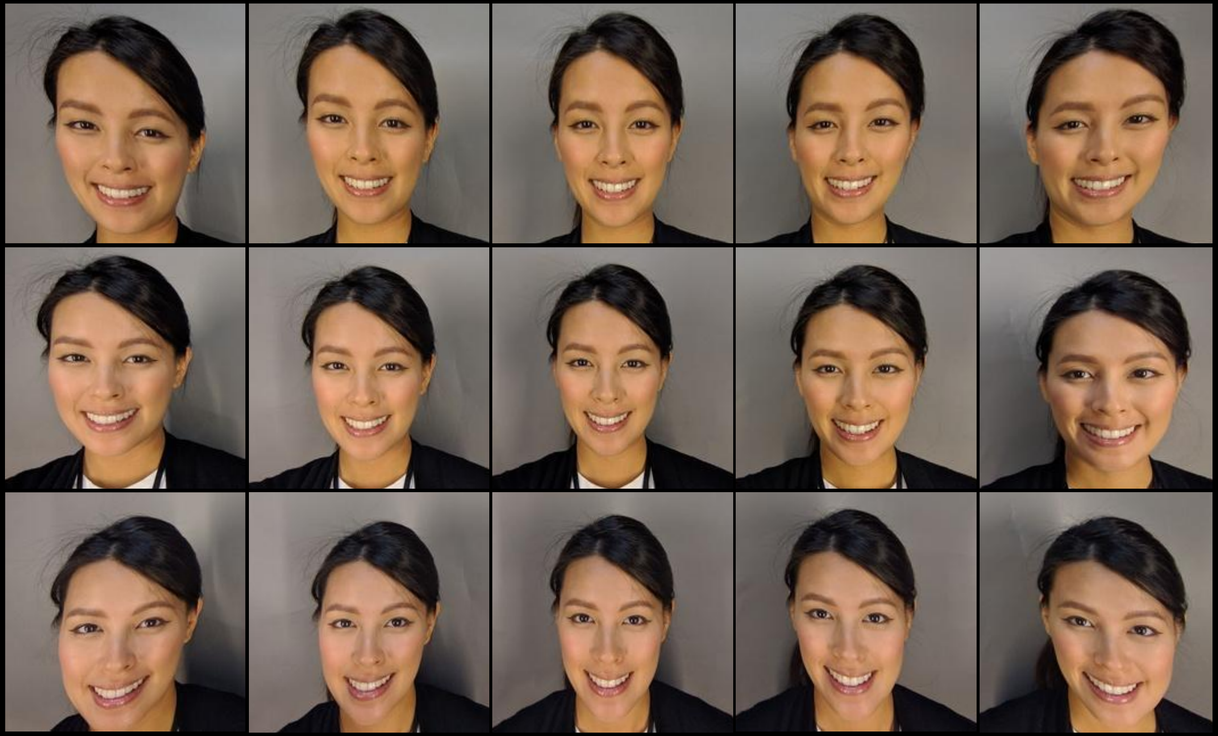



Input
97°FOV (Pixel 3)

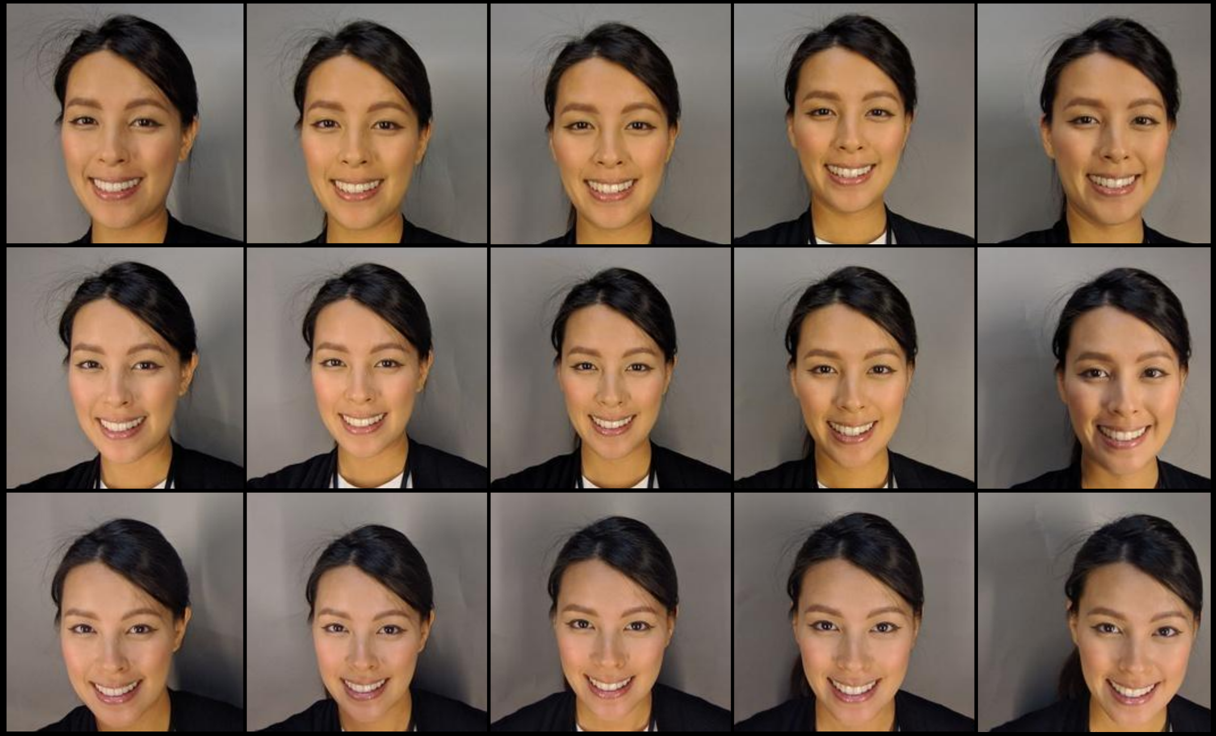

Output

Correction at Various FOVs

- Subjects at a 5x3 grid over a 97°FOV camera.



Input



Output

Normal
FOV

Input 80°FOV

Normal FOV

Output

Large Group

Input 80°FOV

Large Group

Output

Input

Output

Stress Test

Input 103°FOV
(GoPro Hereo5)

Stress Test

Output

Occlusion

Input <u>103°FOV</u>

Occlusion

Output

Comparisons

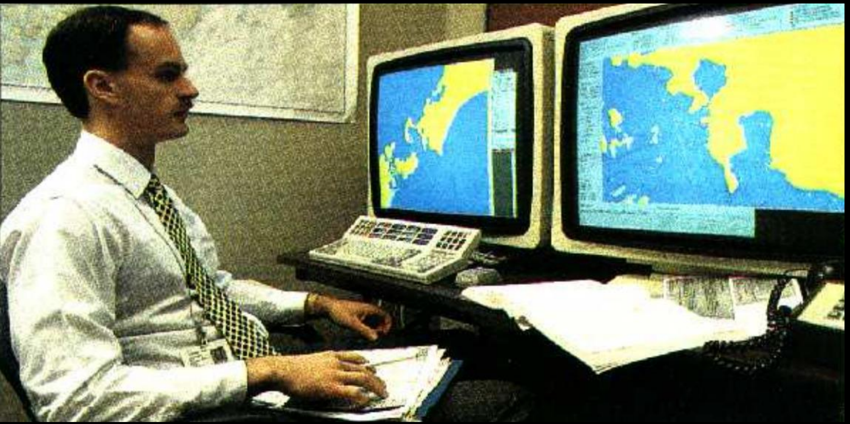
Input (100° FOV)

Pannini
[Sharpless et al. 2010]

Adobe Photoshop
Perspective Warp

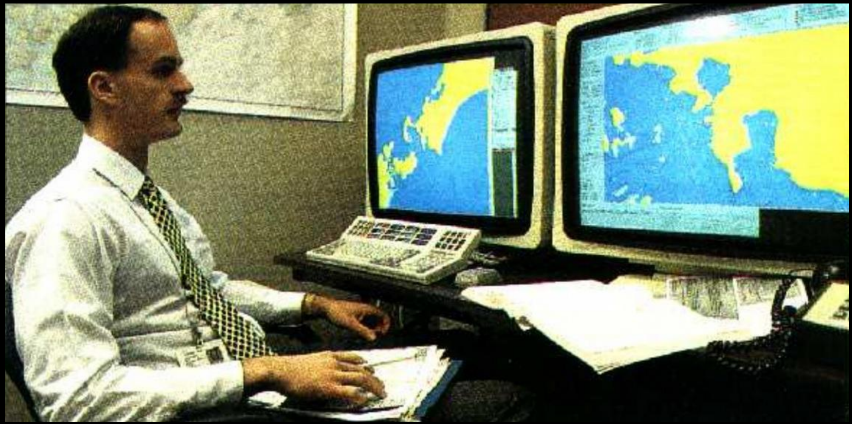
Our method

Comparisons



Input (92° FOV)

[Zorin and Barr 1995]



Our method

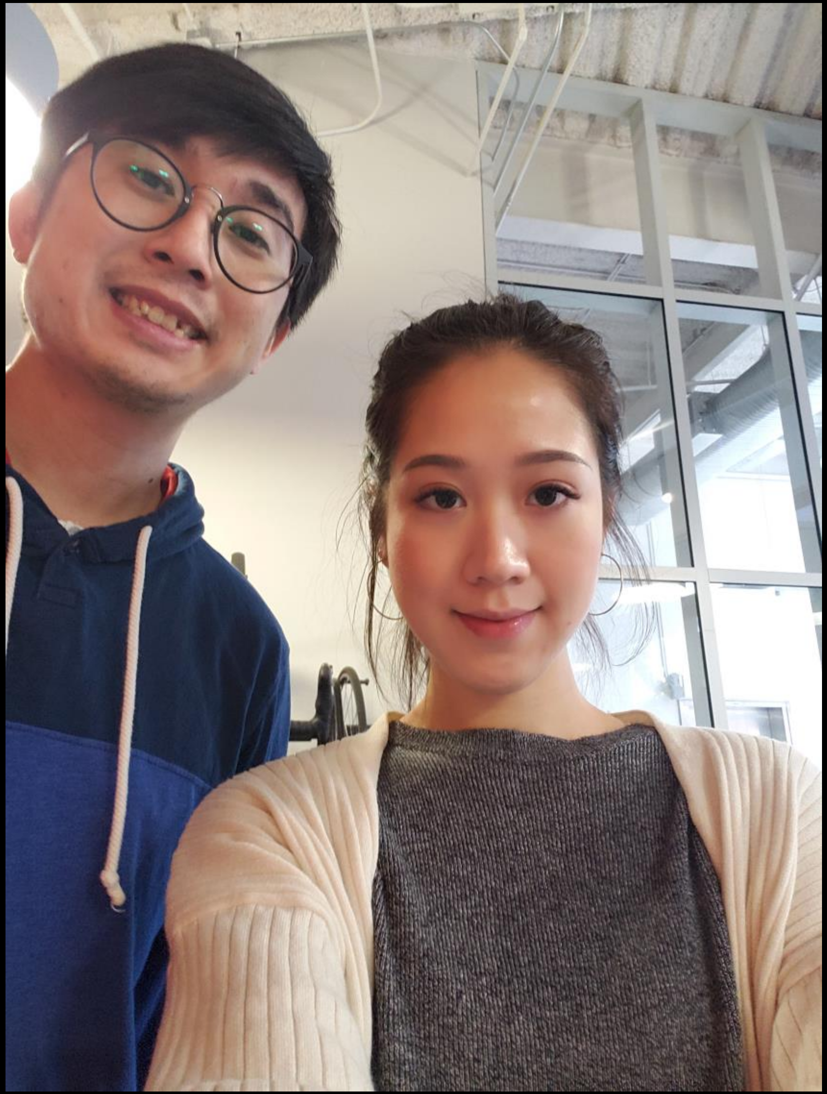
Comparisons: DxO Viewpoint3

Input (97° FOV)

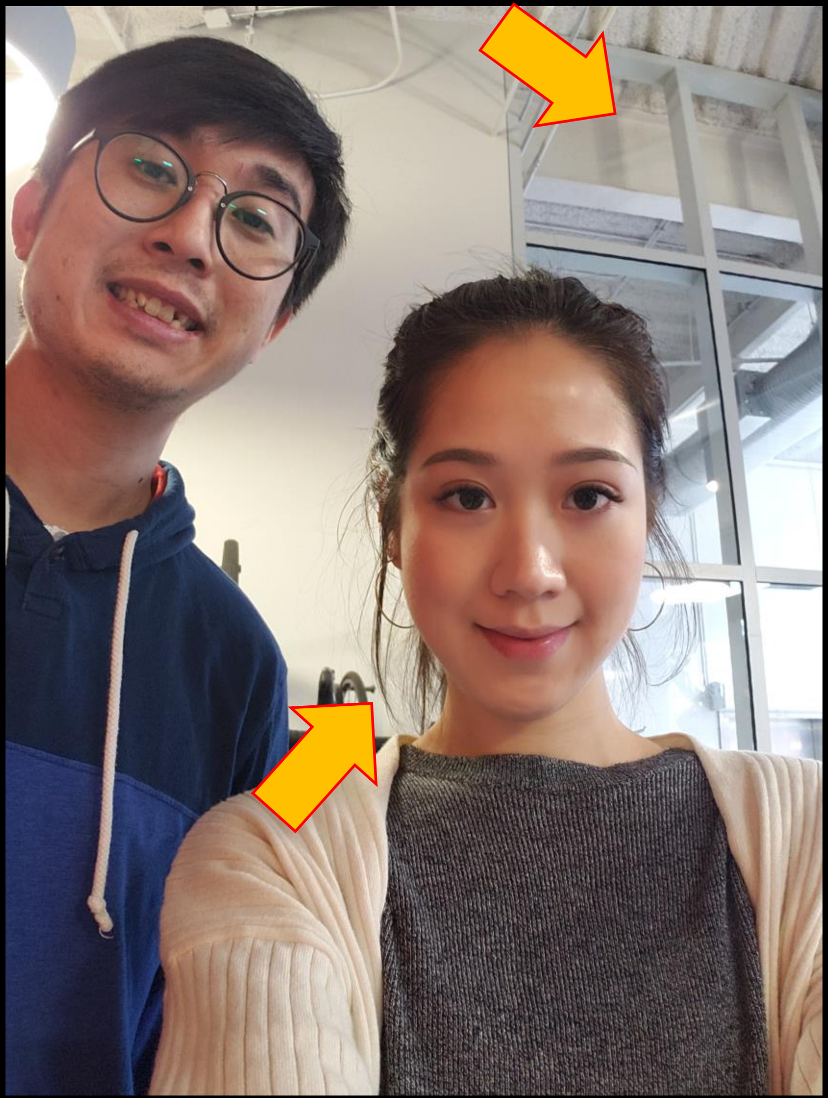
DxO Viewpoint3
Volume Deformation

Our method

Comparisons: Samsung S9+



Input (80° FOV)

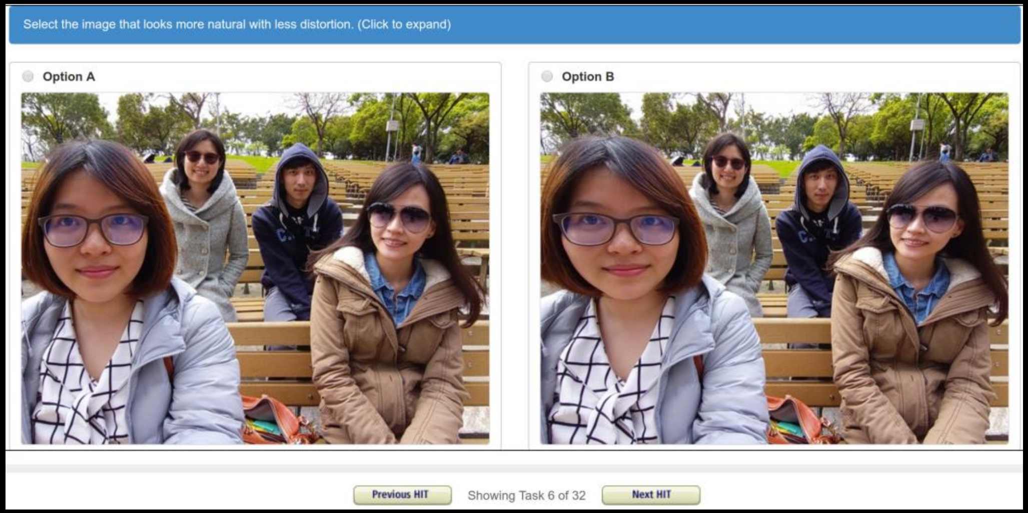


Samsung S9+
“Face Shape Correction”

Our method

User Study on Amazon Mechanic Turk

- 1047 photos from Flickr and collected by us.
- FOVs range from 70°-120°.
- 5 testers on each photo.
  - “looks more natural with less distortion.”



Comparator	Perspective projection	Stereographic projection	Mercator projection	Pannini projection
Favoring our method	92.4%	84%	81%	85%

Runtime

- Takes 920ms for a 12MP image on Qualcomm SDM845
  - Segmentation: 280ms (GPU)
  - Optimization: 340ms (CPU)
  - Warping: 115ms (GPU)

Limitation: missing face detection.

Input (104° FOV)

Output

Limitation: uncorrected body.

Input (90° FOV)

Output

Summary

- Introduce a distortion correction algorithm for wide-angle portraits.
- Based on mesh optimization and subject segmentation.
- Fully automatic and suited for mobile platform.
  - Basis for Wide-angle Group Selfie on Pixel 3.

Input (97° FOV)

Our method

Acknowledgement

Thanks the reviewers for numerous suggestions.

Thanks valuable inputs from Ming-Hsuan Yang, Marc Levoy, Timothy Knight, Fuhao Shi, Robert Carroll, Yael Pritch, David Jacobs, Neal Wadhwa, Juhyun Lee, Alan Yang, Kevin Chen, Sung-fang Tsai, Sam Hasinoff, Eino-Ville Talvala, Gabriel Nava, Wei Hong, Lawrence Huang, Chien-Yu Chen, Zhijun He, Paul Rohde, Ian Atkinson, Jimi Chen, Weber Tang, Jill Hsu, Bob Hung, Kevin Lien, Joy Hsu, Blade Chiu, Charlie Wang, Joy Tsai, Karl Rasche and Rahul Garg.

Thanks to Denis Zorin and all the photography models in this work for photo usage permission.

Thank you

- Find us in Poster session for **Live Demo** on Pixel 3!

  Wednesday, 31 July, 12:15-1:15 pm

Input (97° FOV)

Our method

Poster session for **Live Demo** on Pixel 3: Wed, 31 July, 12:15-1:15 pm

Input (perspective projection)
103° FOV

Our corrected result