# Distortion-Free Wide-Angle Portraits on Camera

YiChang Shih, Wei-Sheng Lai, Chia-Kai Liang

Google

thrive SIGGRAPH2019 LOS ANGELES • 28 JULY – 1 AUGUST

## Problem

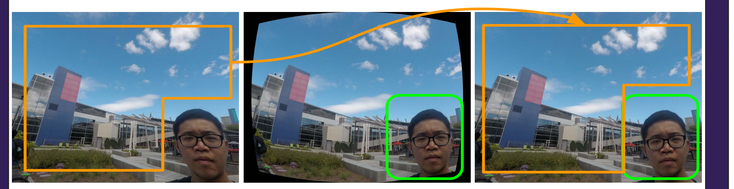
Photographers take wide-angle shots to enjoy expanding views. A wider field-of-view (FOV) introduces a stronger perceived distortion when subjects are near the corners and edges. Faces are stretched, squished, and skewed. We introduce a new algorithm to undistort faces without affecting other parts of the photo. Given a portrait as an input, we formulate an optimization problem to create a content-aware mesh warp for distortion correction.

Input: 97° FOV | Our result

## Observations

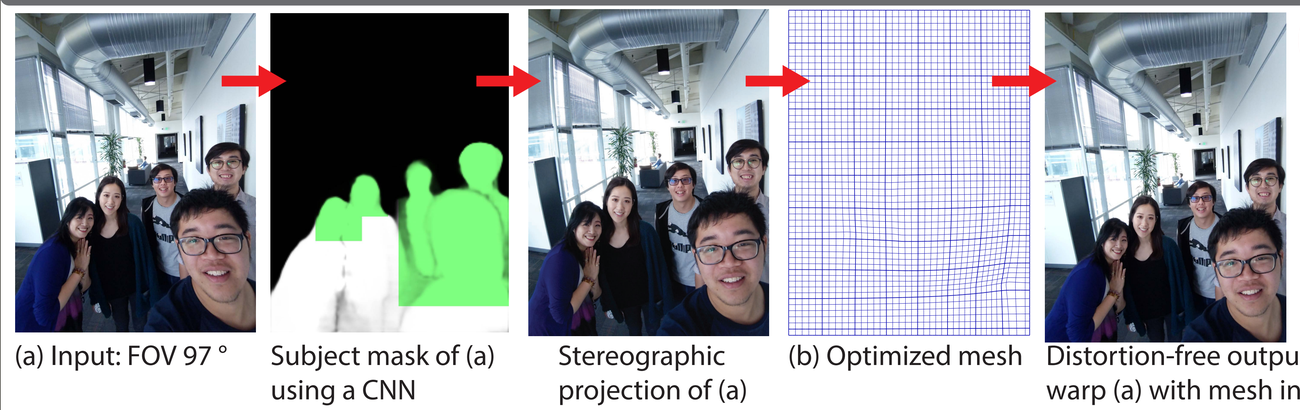
*Perspective projection* preserves straight edges from the 3D world but stretches faces at the image corners. *Stereographic projection* preserves face shape but distorts straight edges in the background.

**Key idea:** combines both projections into a single warp. *Locally* applies stereographic projection on face regions, and perspective projection on the background.



*Perspective projection. Face looks squished. Background looks good.* *Stereographic projection. Face looks good. Background looks distorted.* *Optimal projection by local warp.*

## Algorithm Overview



(a) Input: FOV 97 ° | Subject mask of (a) using a CNN | Stereographic projection of (a) | (b) Optimized mesh | Distortion-free output: warp (a) with mesh in (b)

Tech-talk: Tue 30 Jul
2pm - 3:30pm, Photo Science
Room 150/151

**Mesh Optimization (b)**

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

**A novel per-face energy term:**

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

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

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

**Line-bending term**

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

**Regularization term**

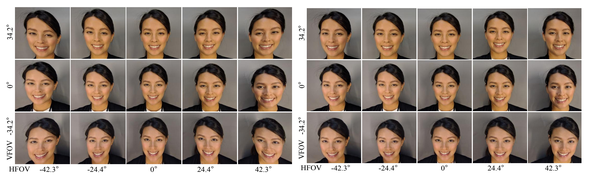
$$E_r = \sum_i \sum_{j \in N(i)} \|\mathbf{v}_i - \mathbf{v}_j\|_2^2$$

## Results

- Validates on 4k+ photos. FOV ranges from 70°-120°. Group size from 1-20 people.
- Efficient on mobile phone: 920ms on SDM845 chipset.


Input | Our result

Left/right: input (79° FOV) and output




Validation over 5x3 grid on 97 ° FOV camera.

## References

[1] Thomas K Sharpless, Bruno Postle, and Daniel M German. 2010. Pannini: a new projection for rendering wide angle perspective images. In Proc. 6th international conference on Computational Aesthetics in Graphics, Visualization and Imaging. 9–16.

[2] Denis Zorin and Alan H Barr. 1995. Correction of geometric perceptual distortions in pictures. In SIGGRAPH. 257–264.

## Comparisons

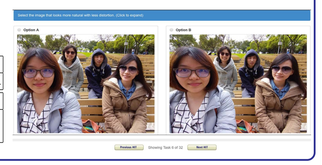
Input | Adobe Photoshop Perspective Warp | Pannini projection [Sharpless et al. 2010] | Our result

Input | [Zorin et al. 1995] | Our result

**User study on Amazon Mechanical Turker**

5235 questions on 1047 testing images from 117 testers.

| Compared to input (perspective projection). | | | | | | | |
|---|---|---|---|---|---|---|---|
| | 5 | 4 | 3 | 2 | 1 | 0 | Total |
| Percentage (%) | 25.0 | 39.2 | 28.2 | 6.6 | 0.9 | 0.0 | 100 |
| Cumulative percentage | 25.0 | 64.2 | **92.4** | 99.1 | 100 | 100 | N/A |



**Limitations**

- Body (torso) are missing from correction and may look unnatural overall.
- Fails when missing face detection.
- Long edges very close to face may be distorted.

Input | Our result

## Acknowledgements

We thank the reviewers for numerous suggestions on user study and exposition. We also thank valuable inputs from Ming-Hsuan Yang, Marc Levoy, Timothy Knight, Fuhao Shi, and Robert Carroll. We thank Yael Pritch, David Jacobs, Neal Wadhwa, Juhyun Lee, and Alan Yang for supports on subject segmenter and face detector integration, Kevin Chen and Sung-fang Tsai for GPU acceleration. We thank Sam Hasinoff, Eino-Ville Talvala, Gabriel Nava, Wei Hong, Lawrence Huang, Chien-Yu Chen, Zhijun He, Paul Rohde, Ian Atkinson, and Jimi Chen for supports on mobile platform implementations. We thank Weber Tang, Jill Hsu, Bob Hung, Kevin Lien, Joy Hsu, Blade Chiu, Charlie Wang, and Joy Tsai for image quality feedbacks, Karl Rasche and Rahul Garg for proofreading. Finally, we give special thanks to Denis Zorin and all the photography models in this work for photo usage permissions and supports on data collection.