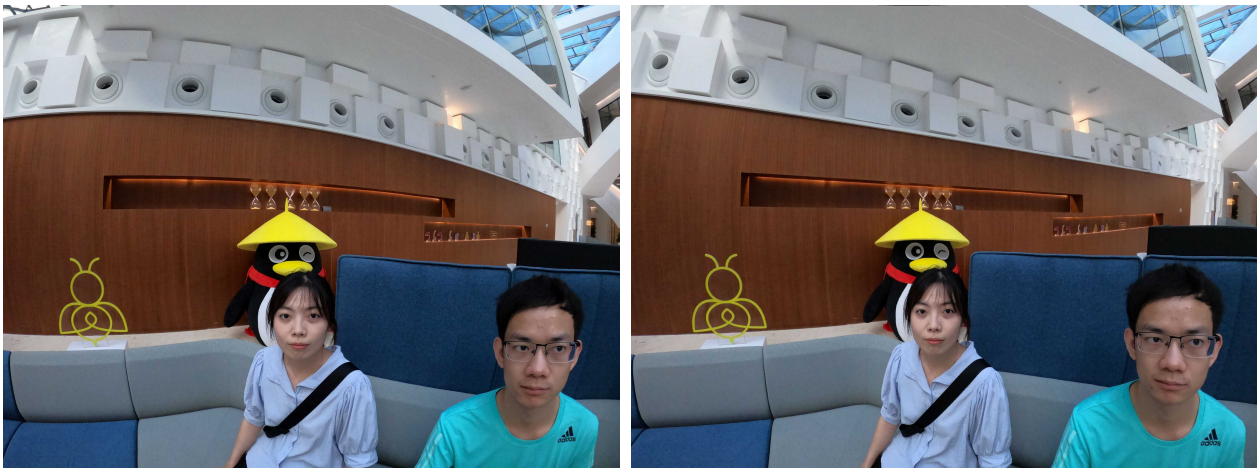
Wide-Angle Rectification via Content-Aware Conformal Mapping
Supplementary Material

Qi Zhang[1] Hongdong Li[2] Qing Wang[3]

[1] Tencent AI Lab [2] Australian National University [3] Northwestern Polytechnical University



(a) Wide-angle image with distortions (b) Rectified image by our method

Figure S1. (a) An image captured by a wide-angle 149° field-of-view camera with different distortions (*e.g.* curved ceiling lines, and skewed face). (b) The proposed method not only eliminates most noticeable distortions in the scene, but also keeps the wide field-of-view.

0. Overview

The supplementary material provides the implementation details of deep image content analysis, baselines (*e.g.* content-aware methods proposed by Carroll *et al*. [2] and Shih *et al*. [8]), and derivation of polar-form Cauchy-Riemann condition. Additional experiments are also presented, which further demonstrates the superior performance of the proposed method compared with all previous methods, including global projections and content-aware baselines.

1. Implement Details

We provide additional implementation details for deep image content analysis in Sec. 4.1 of our submission. Besides, we compare the proposed method to state-of-the-art content-aware methods as well as image warping using global projections. Here we provide additional implementation details regarding the baseline comparison methods described in Sec. 5.1 of our submission.

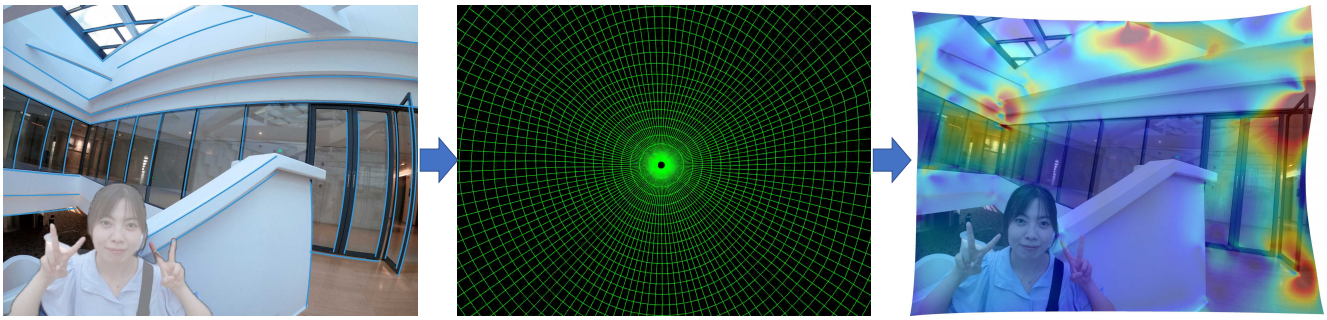
1.1. Deep Image Content Analysis

We present tow components, including curvilinear line perception network (CLP-Net) and boundary-aware salient object detection network (BAS-Net), to detect the results of image content analysis.

For network architecture of the CLP-Net, stacked hourglass network [6] is utilized to learn the segment map $h \in \mathbb{R}^{H\times W}$ of the curvilinear lines from the wide-angle image whose size is $H \times W \times 3$. Three pyramid residual modules are used to extract the feature map with size of $\frac{H}{4} \times \frac{H}{4} \times 512$ from the wide-angle image and then input to five stacked hourglass modules. The feature map is then upscaled by two deconvolution layers to get the feature with size of $H \times W \times 16$. Finally, a convolution layer with 1×1 kernal size to predict the segment map h of curvilinear lines. After the Batch-Normalization and ReLU, the CLP-Net outputs the predicted heat map, which is pixelwisely defined,

$$h(\boldsymbol{u}) = \begin{cases} S(d) & \boldsymbol{u} \in \boldsymbol{l} \\ 0 & \boldsymbol{u} \notin \boldsymbol{l} \end{cases}, \tag{S1}$$

where $S(d) = (1 + e^{-d/(4D)})^{-1}$ refers to a *radial sigmoid*



(a) Input (b) Optimized Mesh (c) Output

Figure S2. Overview of the proposed method: (a) a wide-angle image with visual features (curved lines and prominent regions) as the input; (b) optimizing mesh with polar parameterization by minimizing the energy function; (c) rectified results with conformality, where most lines are corrected, local shapes and original wide field-of-view are preserved.

function, d is the distance from the line to the image center. D is the half image diagonal distance. Considering the fact that the distortion of linear structure is severe far away from the image center, based on the radial sigmoid function, the heat map implicitly contains the radial distortion strength which aids the perception of curved lines farther away from the image center. Besides, $d(\boldsymbol{l})$ which is the length of a distorted line attenuates perception of short lines and local edges in the heat map. Let $\hat{h}(\boldsymbol{u})$ be the perception heat map value from the CLP-Net, the pixel-wise $L2$ loss function is defined,

$$L_h = \sum_{\boldsymbol{u}} \left\| \hat{h}(\boldsymbol{u}) - h(\boldsymbol{u}) \right\|_2^2. \tag{S2}$$

Another crucial problem is the well-annotated distorted lines under the distortion model of the wide-angle lens for training. We adopt the wireframe dataset [3] which already marks the straight lines in perspective images. We use the distortion parameters of different wide-angle lenses to remap the perspective image to the distorted image. The same process is performed on marked straight lines. The heat map is then generated according to Eq. (S1) and distorted lines. Considering the distorted line has complex geometry compared straight line, we randomly flip the perspective images in the horizontal or vertical direction to extend the scale of the training dataset, and each FoV corresponds to one network for accuracy. In summary, we generate 10924 wide-angle images (10000 for training and 924 for testing) with distorted lines for each wide-angle lens. The network is optimized with the Stochastic Gradient Descent (SGD) method on an NVIDIA 1080Ti. The learning rate is set to 0.01. Convergence is reached at 200 epochs.

For BAS-Net, we directly use the pre-trained BAS model provided by Qin *et al*. [7] to extract the salient weights w^{Sa} of wide-angle images. Considering it worked very well on wide-angle images, we did not fine-tune the network on distorted images.

1.2. Spatially-varying Shape-preserving Weight

We compute the spatially-varying weight $\omega^C_{m,n}$ for each polar vertex, which is initialized as 1 (*i.e.*, baseline weight). Image content analysis allows us to adaptively minimize the violation of conformality in places where it is most likely to be noticed. Our spatially-varying weight includes three parts: the baseline weight, the salient weight, and the curvilinear line endpoint weight.

Salient weights. We use BAS-Net to extract salient weight and normalize them in range $[0, 1]$ before adding them to local weights to improve shape preservation in those salient regions. Specially, we assign a salient weight $\omega^{Sa}_{m,n}$ to each polar vertex, allowing us to change the strength of shape-preserving and smoothness terms spatially and preserve the local shape in locations where it is most noticeable.

Curvilinear line endpoint weight. Considering the discontinuous nature between the line-preserving term and shape-preserving term, extremely stretched areas might occur near the endpoints of the curvilinear line. We also add weight $\omega^{Ep}_{m,n}$ to the vertices of the polar cell around the endpoints of curvilinear lines to minimize excessive distortion and extreme straining. Given that this discontinuous distortion is severe far away from the center of the image, we apply the radial sigmoid function to compute the endpoint weight $\omega^{Ep}_{m,n} = (1 + e^{-d/(4D)})^{-1}$, where d is the distance from line segments on the polar cell to the image center, and D is the half image diagonal distance.

We finally combine the baseline weight, the salient weight, and the endpoint weight given to all polar vertices to define the total spatially-varying weight $\omega^C_{m,n}$ as,

$$\omega^C_{m,n} = 1 + 4\omega^{Sa}_{m,n} + 2\omega^{Ep}_{m,n}. \tag{S3}$$

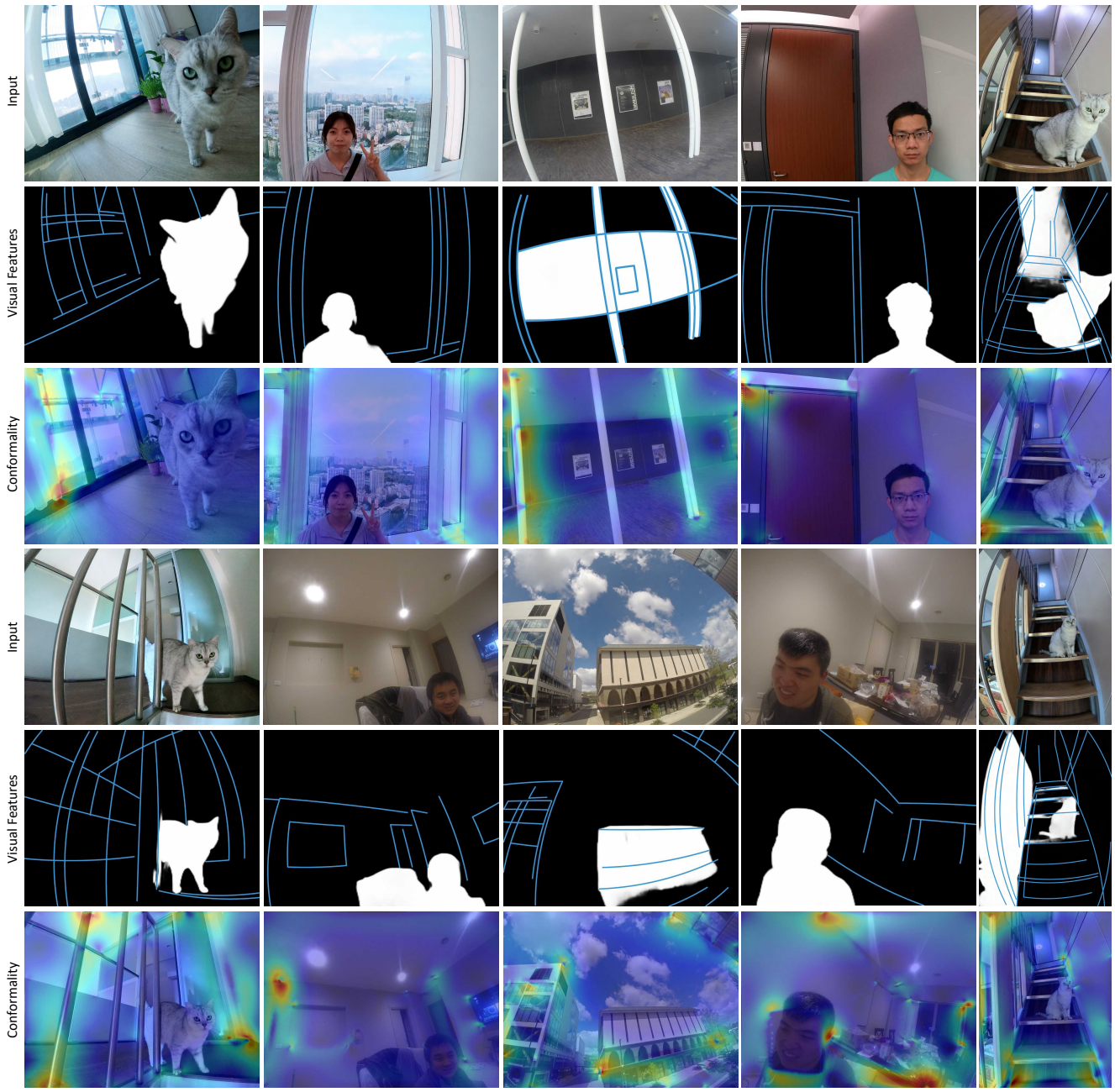


Figure S3. Image content analysis and rectification conformality. The first and fourth rows are wide-angle images with 149° diagonal FoVs except for the second and fourth images in the first rows, whose diagonal FoVs are 117°. The second and fifth rows are visual features, including the detected salient regions by BAS-Net (white) as well as curved line segments by CLP-Net (blue). After the optimization, the third and sixth rows visualize the rectification conformality (*i.e.*, measurement of the shape preservation). Areas with better shape preservation are labeled with colder colors. It can be seen that, after optimization, the value of conformality is small, especially on the salient regions. Besides, the value near the endpoint of curved lines is higher with the distance between the line and the image center.

1.3. Carroll's Content-aware Method

This baseline proposed by Carroll *et al*. [2] is representative of an optimized content-aware method that applies local optimization to correct curved linear structures while preserving the natural shape of objects. We implemented this algorithm from the description provided in the original paper based on Ceres solver [1] and C++, since no open-source code is currently available. We first transfer the image coordinates (u, v) of the raw input to longitude and lat-

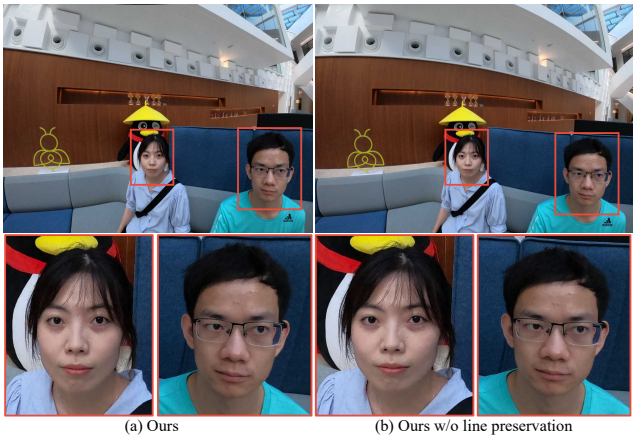
itude coordinates under the Mercator projection, and then establish the mesh grid based on longitude and latitude coordinates. We then implement their energy terms including conformality and smoothness with the same weights. Note that, the preservation term on straight lines is different from their original paper. We use the same line-preserving term (Eqs. (6) and (7) in our submission) provided in our method without the fixed orientations or scales. The setting of spatially-varying constraint weighting is the same as their paper, where salience regions and facial regions are detected by image content analysis and [5] respectively. We finally apply an LM optimizer to find an optimal mesh for image rectification via Ceres solver. Consequently, the comparisons between Carroll's method and our method demonstrate that the proposed content-aware least-squares conformal mapping in polar form outperforms Carroll's local constraints.

1.4. Shih's Content-aware Method

This baseline proposed by Shih *et al*. [8] is representative of the content-aware method specifically addressing portrait photos that rectify distortions on facial regions from a *perspective image* based on stereographic projection. Considering there is no open-source code released by [8], this algorithm was also implemented via the description provided in the original paper based on Ceres solver [1] and C++. We first retrieve the facial regions using a person segmentation. Since the subject mask segmentation [9] used in the original paper [8] does not release the code, we use perspective projection to map the detected salient regions by BAS-Net [7] as the subject mask segmentation. To concentrate the correction on faces and hair, we generate a face mask by intersecting the subject mask with rectangular face bounds returned from face detectors [5]. To include hair, we empirically extend the box height by 2 times along the top direction, and the width by half on both sides. Since both the perspective and stereographic images are mapped from the raw input based on the field-of-view (FoV) [4], so do the focal length for each global projection, these focal lengths (*e.g.* f_p for perspective projection and f_s for stereographic projection) are used to establish stereographic mesh (Eq.(2) in their original paper) and correct distorted facial regions. We then implement the energy terms mentioned in their original paper with the same parameter setting. We finally apply an LM optimizer to search for an optimal mesh for image rectification via Ceres solver.

2. Additional Experiments

We demonstrate the results of the proposed method on a number of examples. **All the results are better to view on the screen with zooming in.**



(a) Ours (b) Ours w/o line preservation

Figure S4. Ablation study of the line-preserving term. (a) shows the rectification of the proposed method. (b) shows the results without the line-preserving term. We can see that, without the line-preserving term, our rectification tends to look globally very similar to the stereographic projections. It also demonstrates that the proposed shape-preserving term preserves the natural shape locally.

2.1. Image Content Analysis

Fig. S3 shows the image content analysis and conformality after optimization. Areas with better shape preservation are labeled with colder colors. It can be seen that, after optimization, the value of conformality is small, especially on the salient regions. It means the proposed method could locally preserve the natural shape, which demonstrates the effectiveness of the LSCM-based shape-preserving term. Besides, the value near the endpoint of curved lines is higher with the distance between the line and image center, which is caused by the line-preserving term. It also explains the reason for the weights setting in the line-preserving term (see Sec. 4.2).

2.2. Ablation Study

To demonstrate the performance of the proposed method, we conduct the ablation study of line-preserving, smoothness and boundary-preserving terms.

Line Preserving. Fig. S4 qualitatively shows the ablation study of our method with/without the line-preserving term. We can see that, without the line-preserving term, our rectification tends to look globally very similar to the stereographic projections, which certifies the performance of the LSCM-based formulation on shape preservation. With the introduction of the line-preserving term, the curved lines become straight, while at the same time the local shape is preserved.

(a) Ours (b) Ours w/o smoothness

Figure S5. Ablation study of smoothness term. (a) shows the rectification of the proposed method. (b) shows the results without smoothness term. It demonstrates the superior performance of our method for smoothing the local shape preservation and global linear structure on the image plane.

(a) Ours (b) Ours w/o boundary preservation

Figure S6. Ablation study of the boundary-preserving term. Red rectangles indicate the cropping frames. It can be seen that, compared with ignoring the boundary-preserving term, our method retains the original wide and rectangular field of view.

Smoothness. Fig. S5 qualitatively illustrates the results of our method with/without the smoothness term. We can see that Fig. S5(a) preserve both local shape and global lines, while Fig. S5(b) causes the "seams' between the neighboring transformations, especially on visual features. It certifies the effectiveness of the proposed smoothness term.

Boundary Preserving. Fig. S6 qualitatively shows the ablation study of our method with/without the boundary-preserving term. It can be seen that, compared with ignoring the boundary-preserving term, our method retains the original wide and rectangular field of view, which demonstrates the effectiveness of the proposed boundary-preserving term.

In summary, our method strikes an excellent balance between local shape-preserving (*e.g.* faces) and global linear-structure-preserving (*e.g.* "straight lines must be straight"), making the rectified images look both real, natural, and visually pleasing, while at the same time enjoying the immersive wide-angle visual experience by retaining the original wide field of view.

2.3. Comparisons of Conformal Mappings

To demonstrate the benefits of our polar-form LSCM, we also perform our method without the boundary-preserving term compared with Carroll's method [2], as shown in Fig. S7. It could be considered as the ablation study of conformal mapping in polar and Cartesian domains. Specifically, Carroll's approach stretches the scene heavily near the poles, even destroying the line preservation. For example, the linear structure of the building of the third scene in Fig. S7 become crooked. The reason for this phenomenon is that the conformal map based on Mercator projection is an axisymmetric projection. Interestingly, it also results in the deflection of the orientation of local objects in Carroll's method, such as the portraits and window frames in Fig. S7. That is why the original Carroll's method requires manually fixing some lines' orientation and adds the line-orientation constraints. In contrast, our results are free from those residual distortions automatically, giving rise to natural looking and pleasing images. Since camera distortions are consensus to be radial symmetry, polar coordinates are indeed a better fit for parameterizing them, as verified by comparisons above.

2.4. Qualitative Evaluation in Details

To further validate the performance of the proposed method, we compare our method to two baselines in detail. Figs. S8 and S9 illustrate the results by our method and the state-of-the-art content-aware methods provided by Carroll *et al*. [2] and Shih *et al*. [8]. We show some details in red boxes with zooming in for better view.

In particular, Carroll's method modifies the ceiling orientation, such as the window frame in the first image of Fig. S8. Besides, Carroll's method suffers the severe loss of image contents, especially at the boundaries. Shih's method performs well with smaller FoVs (*e.g.*, the third and fourth images of Fig. S8) but fails with increasing FoVs, such as the shoulder of the girl in the first scene of Fig. S8. In addition, Shih's method is a content-aware method that handles explicitly the facial regions in portrait photos, which leads the disharmony between portraits and other places (*e.g.* the big head with cube and small body in the third scene. As shown in the second scene of Fig. S9, the face of the girl corrects very well, but the cube is stretched. Finally, in Shih's results, the linear structure of an edge can not be preserved when it is very close to face regions, as shown in the first scene of Fig. S9. Compared with them, our method achieves superior results that preserve both global linear structure and local natural shape.

2.5. Comparison with Baseline

Fig. S12 provides the comparisons against global projections (*e.g.* Mercator, perspective, stereographic projections) as well as content-aware methods proposed by Carroll *et*

(a) Input (b) Carroll's (c) Ours w/o boundary preservation

Figure S7. The comparison between our method without boundary-preserving term and the method proposed by Carroll *et al*. [2]. It shows the differences of conformal mappings in the polar domain (our method) and Cartesian domain (Carroll's method). Considering the conformal mapping derived from axisymmetry Cartesian coordinates and Mercator projection, Carroll's method stretches the boundaries near the poles, destroys the linear structure, and alters the shape-orientations. In contrast, our method preserves all these parts very well, which demonstrates the benefits of our polar-form LSCM constraint.

al. [2] and Shih *et al*. [8]) without cropping. We can see that others suffer severe loss of image contents especially at the peripheral, but our method preserves a nearly perfect rectangular field of view. Specifically, Carroll's method [2] stretches the scene at the pole, while image boundary lines become crooked. The perspective projection only keeps the central part of a wide FoV, and this is the case for Shih's method [8].

Fig. S13 shows manually cropped results of our rectification. We compare them to three global projections and other content-aware baseline methods [2] and [8]. Global projections, while being able to reduce certain types of distortions, often leave other residual distortions. In particular, The perspective projection shows severe stretching here, so does Shih's method, which uses perspective images as inputs to correct facial regions only. The Mercator and stereographic results both bend lines. Carroll's results look globally similar to the Mercator projection, but the lines are straight. However, as shown in Figs. S12 and S13, Carroll's results miss border contents and stretch the lines near the boundaries. Compared with them, our method achieves a decent balance between distortion-minimization and FoV-retaining.

Figs. S14 and S15 demonstrate the rectification results under perspective projection, our method and Carroll's method [2]. Neither perspective projection nor Carroll's method could preserve the boundary. Since the perspective projection is not conformal and leads to perspective distortion, objects near the periphery of wide-angle perspective images can appear unnaturally stretched and distorted. While both our method and Carroll's method successfully preserve most of the straight lines and local shapes, Carroll's approach fails to keep the orientations of linear structure and distorts near the boundaries, such as top boundary in the fifth scene of Fig. S14 and the pipes in the third scene of Fig. S15.

Fig. S15 also shows fisheye rectification examples of the dataset released by Carroll *et al*. [2] (first and second rows). It is worth noting that the only difference between our and Carroll's methods in Figs. S15 is the shape-preservation. Both methods use the same lines, line-preservation and without boundary-preservation. It is interesting to see that Carroll's method modifies the ceiling orientation. In contrast, our results are free from those residual distortions, giving rise to natural looking and pleasing images. Note that the results in Fig. S15 are different with rectifications in [2] due to the different inputs of marked lines, as described in Sec. 1.3. Perspective projection exhibits severe stretching, as perspective stretches to infinity for a 180° FoV. In contrast, our results are free from those residual distortions,

Figure S8. The wide-angle input images (a) at the top, which may be more easily viewed by zooming in. The rectification results (d) of our method are compared with state-of-the-art content-aware methods proposed by Carroll *et al*. [2] (b) and Shih *et al*. [8] (c). Compared with them, our method achieves superior results that preserve both global linear structure and local natural shape.

giving rise to natural looking and pleasing images.

In summary, our method has effectively removed the wide-angle distortions while at the same time maintaining an ultra-wide field of view, without sacrificing salient image contents.

2.6. Failure Cases

As mentioned in Sec. 6 of the submission, our method does not always compute a satisfactory result. We found two typical scenarios for which the result is often less than we desired.

1) The visual quality of the rectified images is influenced by image content analysis, especially the curved line detection. Fig. S11 shows a failure case that a curved line is separated into two curved line segments (*e.g.* 18-th line segment and 19-th line segment labeled in Fig. S11(a)) due to the boundary or occlusions. The local shape of objects (*e.g.* the girl) could be preserved, but the global structure of the window frame is broken.

2) There is a trade-off between shape-preserving and line-

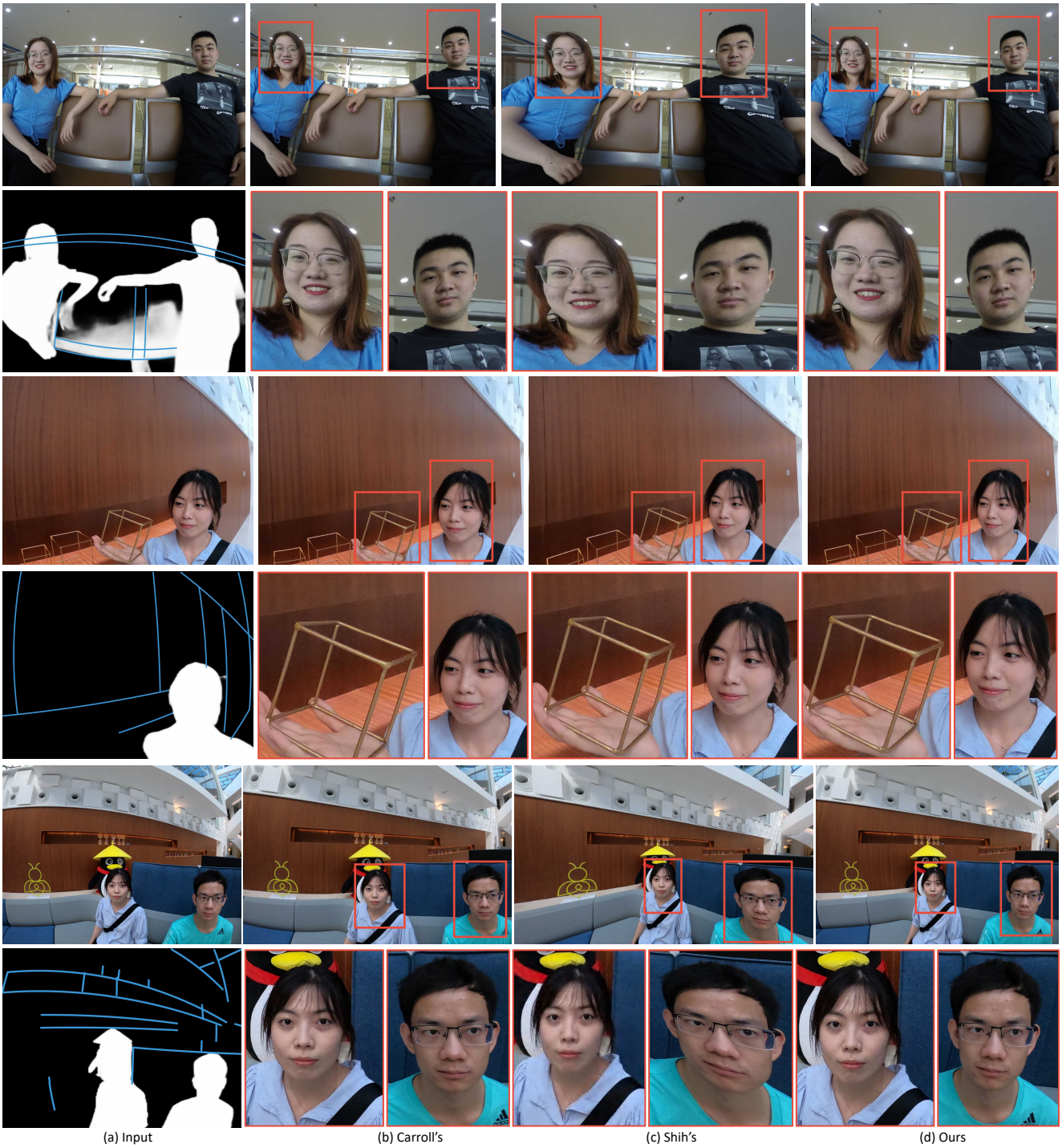
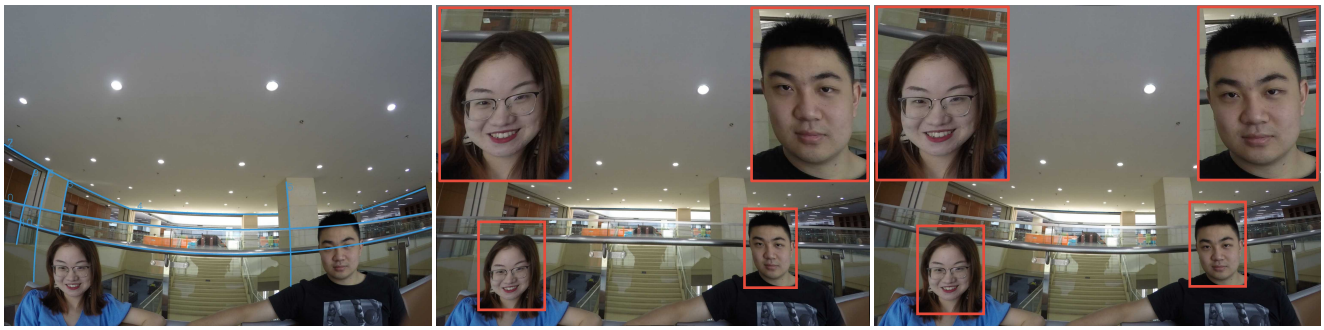


Figure S9. Comparison in detail. (a) shows the input of the proposed method, including the wide-angle image and visual features of image content analysis. (b-d) show the rectification results of Carroll's, Shih's and our methods. Compared with state-of-the-art methods, our method has effectively removed the wide-angle distortions while at the same time maintaining an ultra-wide field of view, without sacrificing salient image contents.

bending. When a curved line cuts across one local shape (*e.g.* face), the shape will be stretched, because the weight applied to line-preserving energy is heavier, as shown in Fig. S10. For example, the faces of the boy and girl are



(a) Input (b) Ours (c) Ours w/o line-preserving term

Figure S10. A failure case of the trade-off between shape-preserving and line-bending. (a) shows the wide-angle input image with detected curved lines (blue). (b) shows the rectification of our method where the faces are stretched because of the correction of the curved lines. In contrast, (c) shows the results without the line-preserving term. It can be seen that the facial regions could be locally preserved.

(a) Input (b) Ours

Figure S11. A failure case belong to the same curved line.

stretched because of the correction of the first curved line segment. In contrast, we also demonstrate the optimized result without the line-preserving term. It can be seen that the facial regions could be locally preserved.

3. Derivation of Eq. (3)

Formally, the polar transform between the polar domain and Cartesian domain (image space (u, v)) is defined in complex form:

$$\mathcal{U} = u + iv = \rho(\theta)\cos\phi + i\rho(\theta)\sin\phi, \tag{S4}$$

where $\rho(\theta)$ and ϕ indicate the radial and angular coordinates, θ is the angle between the principal axis and the incoming ray for specialization of radially symmetry projections. $\mathcal{X} = \theta + i\phi$ is the complex form of polar coordinates. According to [4], the stereographic projection of a wide-angle lens can be described by the following formula,

$$\rho(\theta) = 2f\tan\left(\frac{\theta}{2}\right). \tag{S5}$$

Substituting Eq. (S5) into (S4), we obtain the polar transformation under stereographic projection,

$$\mathcal{U} = 2f\tan\left(\frac{\theta}{2}\right)\cos\phi + i2f\tan\left(\frac{\theta}{2}\right)\sin\phi. \tag{S6}$$

Based on the Cauchy-Riemann condition, the partial equations of polar transformation are then derived,

$$\frac{\partial\mathcal{U}}{\partial\theta} = f\frac{1}{\cos^2\frac{\theta}{2}}\cos\phi + if\frac{1}{\cos^2\frac{\theta}{2}}\sin\phi \tag{S7}$$

$$\begin{aligned}\frac{\partial\mathcal{U}}{\partial\phi} &= -2f\tan\frac{\theta}{2}\sin\phi + i2f\tan\frac{\theta}{2}\cos\phi \\ &= -f\frac{2\sin\frac{\theta}{2}\cos\frac{\theta}{2}}{\cos^2\frac{\theta}{2}}\sin\phi + if\frac{2\sin\frac{\theta}{2}\cos\frac{\theta}{2}}{\cos^2\frac{\theta}{2}}\cos\phi \\ &= i2\sin\frac{\theta}{2}\cos\frac{\theta}{2}\frac{\partial\mathcal{U}}{\partial\theta}\end{aligned} \tag{S8}$$

Considering $\sin\theta = 2\sin\frac{\theta}{2}\cos\frac{\theta}{2}$, according to Eqs. (S7) and (S8), we finally obtain the polar-form Cauchy-Riemann condition:

$$\frac{\partial\mathcal{U}}{\partial\phi} - i\sin\theta\frac{\partial\mathcal{U}}{\partial\theta} = 0. \tag{S9}$$

Acknowledgements. The authors would like to thank Ying Feng, Xiaoyu Li, Guangcan Xu, and Xuying Liu for the dataset capture. The work is partially supported by NSFC under Grant 62031023.

References

[1] Sameer Agarwal, Keir Mierle, and Others. Ceres solver. http://ceres-solver.org, 2010. 3, 4

[2] Robert Carroll, Maneesh Agrawal, and Aseem Agarwala. Optimizing content-preserving projections for wide-angle images. *ACM Trans. Graph.*, 28(3):43, 2009. 1, 3, 5, 6, 7, 11, 12, 13, 14

[3] Kun Huang, Yifan Wang, Zihan Zhou, Tianjiao Ding, Shenghua Gao, and Yi Ma. Learning to parse wireframes in images of man-made environments. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 626–635, 2018. 2

[4] Juho Kannala and Sami S Brandt. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(8):1335–1340, 2006. 4, 9

[5] Jian Li, Yabiao Wang, Changan Wang, Ying Tai, Jianjun Qian, Jian Yang, Chengjie Wang, Jilin Li, and Feiyue Huang. Dsfd: Dual shot face detector. In *IEEE Conf. Comput. Vis. Pattern Recog.*, 2019. 4

[6] Alejandro Newell, Kaiyu Yang, and Jia Deng. Stacked hourglass networks for human pose estimation. In *Eur. Conf. Comput. Vis.*, pages 483–499. Springer, 2016. 1

[7] Xuebin Qin, Zichen Zhang, Chenyang Huang, Chao Gao, Masood Dehghan, and Martin Jagersand. Basnet: Boundary-aware salient object detection. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 7479–7489, 2019. 2, 4

[8] YiChang Shih, Wei-Sheng Lai, and Chia-Kai Liang. Distortion-free wide-angle portraits on camera phones. *ACM Trans. Graph.*, 38(4):1–12, 2019. 1, 4, 5, 6, 7, 11, 12

[9] Neal Wadhwa, Rahul Garg, David E Jacobs, Bryan E Feldman, Nori Kanazawa, Robert Carroll, Yair Movshovitz-Attias, Jonathan T Barron, Yael Pritch, and Marc Levoy. Synthetic depth-of-field with a single-camera mobile phone. *ACM Trans. Graph.*, 37(4):1–13, 2018. 4

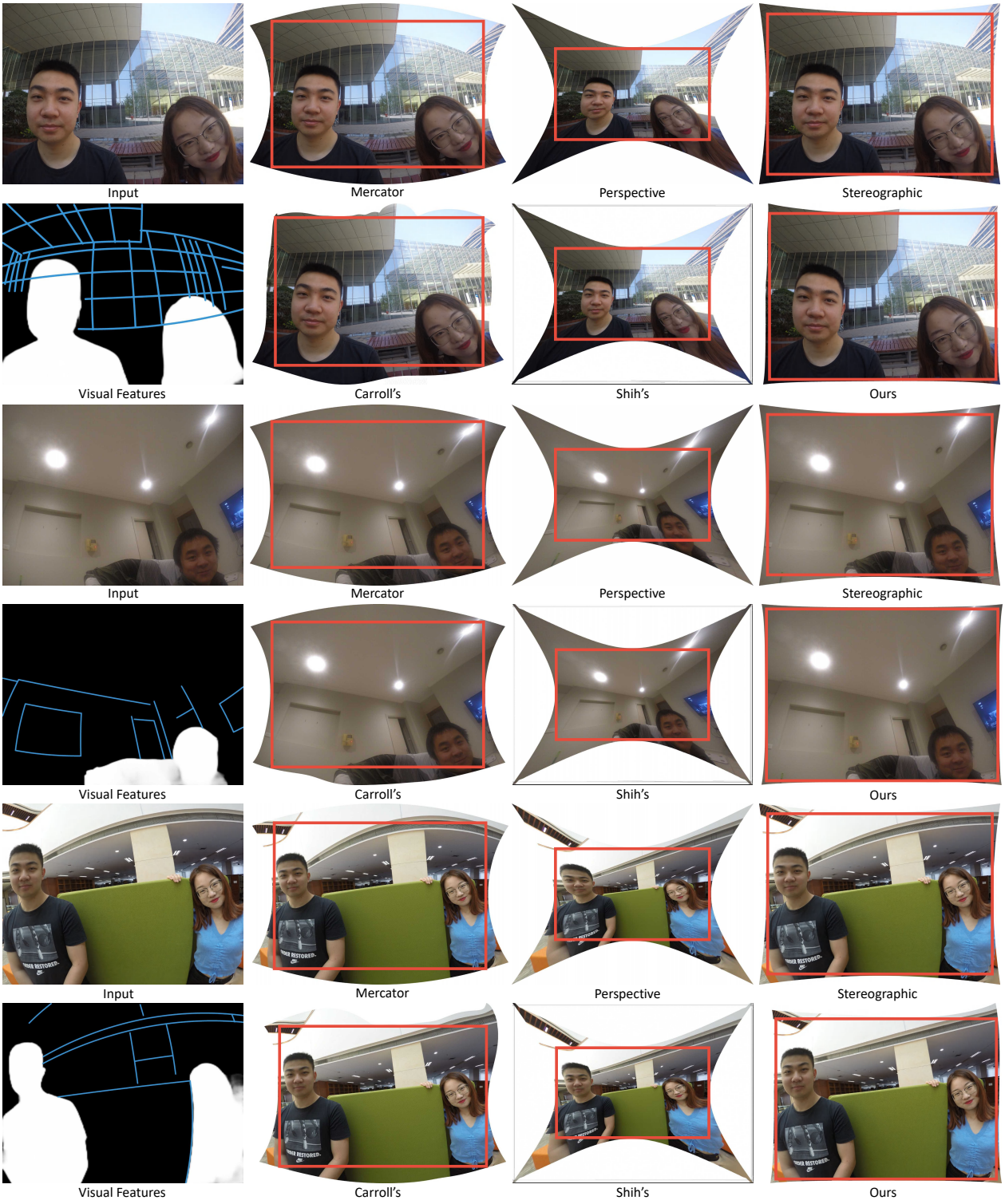


Figure S12. Uncropped comparisons. FoV of each input is 149°. Red rectangles indicate the cropping frames. It can be seen that our method retains the original wide and rectangular field of view. Carroll's results [2] stretch the scene at the pole, while image boundary lines become crooked. The perspective projection only keeps the central part of a wide FoV, and this is the case for Shih's results [8].

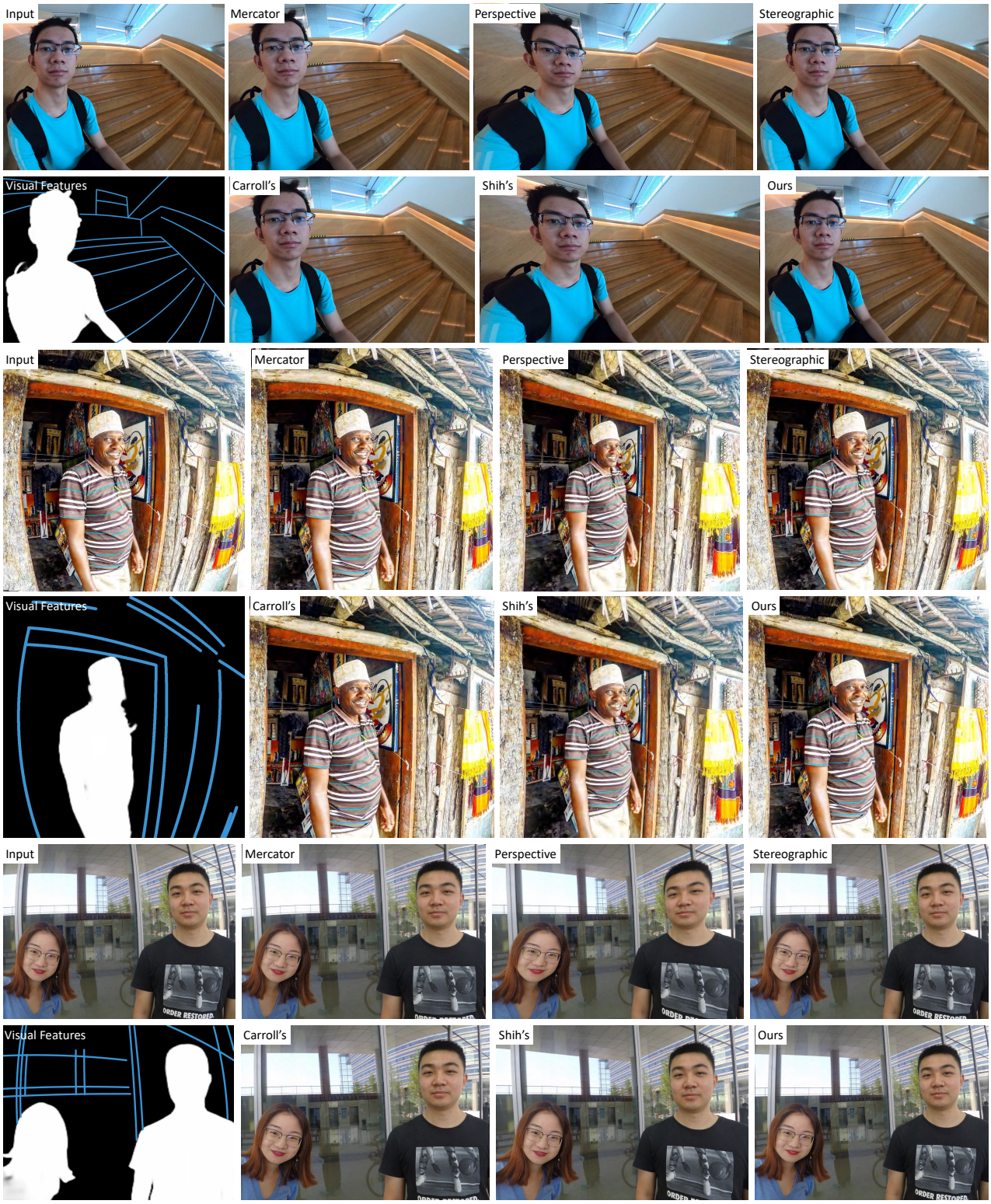


Figure S13. Cropped results of our rectification with global projection methods and local methods, by Carroll *et al*. [2] and Shih *et al*. [8]. FoVs of each input are 149°, 123°, 117°. Our method achieves a decent balance between distortion-minimization and FoV-retaining.

Figure S14. Uncropped rectification results of our method compared with perspective projection and Carroll's method [2]. FoV of each input is 149°. We can see that others suffer severe loss of image contents especially at the peripheral, but our method preserves a nearly perfect rectangular field of view. Carroll's results [2] stretch the scene at the pole, while image boundary lines become crooked. The perspective projection only keeps the local shape near the image center. Compared with them, our method has effectively preserve both local shape and linear structure while at the same time maintaining an ultra-wide field of view.

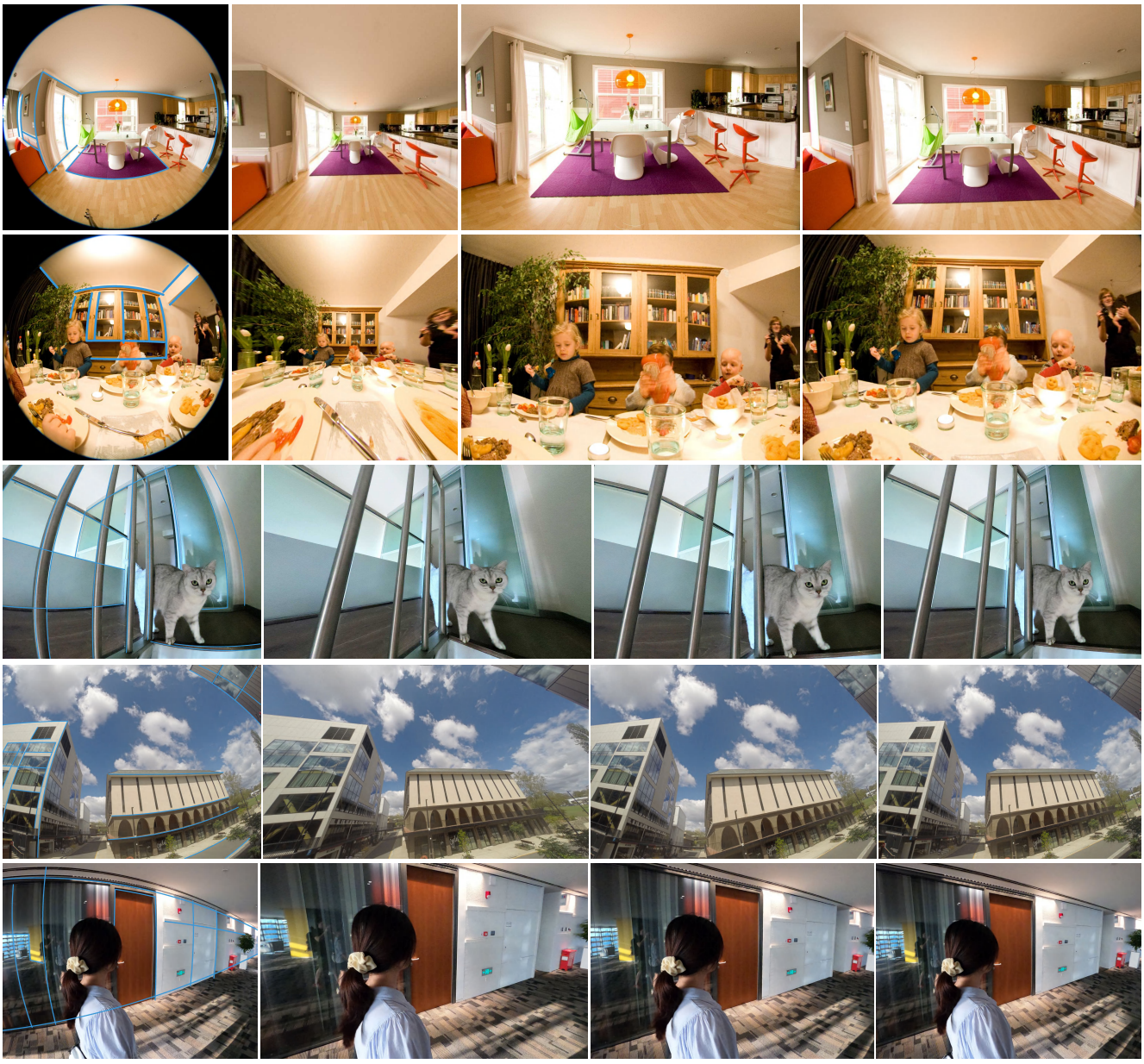


Figure S15. Cropped rectification results of our method compared with perspective projection and Carroll's method [2]. FoVs of inputs are 180°, 180°, 149°, 149°, and 117° . We can see that others suffer severe loss of image contents especially at the peripheral, but our method preserves a nearly perfect rectangular field of view. Carroll's results [2] stretch the scene at the pole. The perspective projection only keeps the local shape near the image center. It is also noting that the Carroll's results are different from that of his paper. The main reason is that the distorted lines detected by CLP-Net are different from manually marked lines in his paper. In addition, according to Sec. 1.3, the line-preservation term of Carroll's method is different from our implemented code, wherein the orientation term is ignored.