Wide-angle Rectification via Content-aware Conformal Mapping

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

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

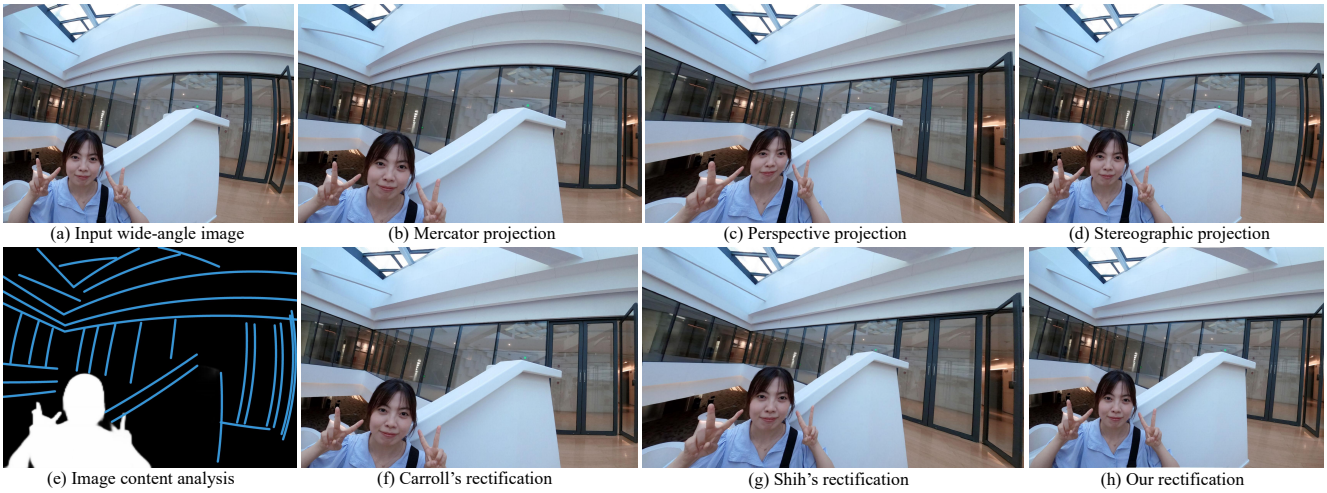


Figure 1. Given a wide-angle image (as shown in (a) *e.g.*), this paper proposes a new image rectification method which automatically corrects evident wide-angle lens distortions (*e.g.* curved ceiling lines, and skewed face), and obtains a rectified image (*e.g.* (h)). This is achieved by deep-learning based image content analysis ((e)), as well as *content-aware conformal mapping*. Our method effectively reduces wide-angle distortions, and at the same time maintains the original ultra-wide field of view, without sacrificing salient image contents. For comparison, we also show results by previous methods in (b,c,d,f,g). Our method achieves the best results. For more comparisons please see the paper as well as the supplemental material. **Better viewed on screen with zooming in.**

Abstract

Despite the proliferation of ultra wide-angle lenses on smartphone cameras, such lenses often come with severe image distortion (e.g. curved linear structure, unnaturally skewed faces). Most existing rectification methods adopt a global warping transformation to undistort the input wide-angle image, yet their performances are not entirely satisfactory, leaving many unwanted residue distortions uncorrected or at the sacrifice of the intended wide FoV (field-of-view). This paper proposes a new method to tackle these challenges. Specifically, we derive a locally-adaptive polar-domain conformal mapping to rectify a wide-angle image. Parameters of the mapping are found automatically by analyzing image contents via deep neural networks. Experiments on a large number of photos have confirmed the superior performance of the proposed method compared with all available previous methods.

1. Introduction

It has become trendy to equip modern smartphone cameras with ultra wide-angle lenses, to allow the user to shoot photographs of natural landscapes or buildings with a wide field-of-view (FoV), or capture a group-selfie in a tight space. This trend can be easily seen on high-end phones for example iPhone 13 which features a rear camera with 120° FoV, or Samsung S21 that is 123°.

While such lenses provide the user with an immersive visual experience, they also induce apparent and unavoidable image distortions, resulting in *e.g.* curved straight lines or sheared human faces. Traditional methods for lens distortion removal solve this problem by finding a global parametric geometric transformation to warp the input image. Their performances are however far from satisfactory, with either obvious residual distortions on linear structure or local shapes, or missing image contents at the image boundaries due to a much compromised field of view.

In contrast, human eyes, while enjoying a wide field of

view (about 120° in monocular case [24]), are capable of perceiving a wide environment without obvious distortions. In our mind's eye, lines appear to be straight, and objects preserve their natural shapes. Our brain seems to be able to intelligently "undistort" different parts of the view-field by applying different content-aware transformations. Moreover, it is also recognized that human vision is most sensitive to global linear structures as well as perceptually salient regions in a scene. Given that distortions are unavoidable when one projects a view-sphere onto a flat image plane, our goal of this paper is not to eliminate all distortions (which is impossible), but to *minimize* those most visually salient distortions such that they become unnoticeable or tolerable.

To this end, this paper develops a content-aware image projection method which focuses on correcting the most salient distortions (*e.g.* visual features), while preserving local shapes in the scene as much as possible. Specifically, our method searches for an optimal content-aware *conformal mapping* which warps a wide-angle input image to a rectified one, in a locally adaptive manner by respecting local image contents. This way, it not only eliminates most noticeable distortions in the scene, but also retains the wide FoV, offering the user the intended immersive experience endowed by the wide-angle lens.

Specifically, key contributions of the paper are:

- An automatic content-aware wide-angle image rectification method which preserves both local shapes and global structures by analyzing image contents via deep-learning.
- A new formulation for Least-Squares Conformal Mapping (LSCM) in the polar domain to achieve locally adaptive shape-preserving transformation.
- A new optimization procedure which incorporates multiple energy terms, each encodes a different prior on local shapes, linear structures, smoothness and image boundaries, respectively.

Our method strikes an excellent balance between local-shape-preserving (*e.g.*, "circles remain circular") 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. Our method evidently outperforms all previous methods for wide-angle rectification, including both global warp based methods (*e.g.* perspective correction, Mercator projection), and local optimization method [4] that alters the orientation of local shapes, and method that is restricted to portrait photo only [22]. Our method is fully automatic, without the need of human intervention. It also runs fast, taking about 1-2 seconds per image, and can be easily optimized on a mobile-GPU to reach sub-second processing time, sufficient for real-world photography applications.

2. Related Work

This section gives a brief review of existing methods for wide-angle image rectification task. Although the task is closely related to radial lens calibration (*e.g.*, fisheye undistortion [6,14,25,26] or learning-based methods [15,27,28]), their purposes and objectives are different. Radial lens calibration aims to recover a *perspectively correct* image from a radially distorted one. The resulted images still contain obvious distortions (*i.e.*, foreshortening) which twist shapes in the scene, especially for wide-angle lens. Moreover, if image cropping is applied to the result in order to obtain a rectangular image boundary, there will be significant reduction in the field of view. In contrast, wide-angle image rectification, as studied in this paper, aims to remove (or minimize) prominent distortions including but not limited to perspective foreshortening, while at the same time retaining the original wide field of view.

Existing methods for wide-angle image rectification can be roughly classified in two groups: (i) global rectification, and (ii) locally adaptive rectification.

Global Rectification. Global rectification applies a single global geometric mapping to the input image aiming to *undo* the lens distortion. Examples of global methods include perspective projection, stereographic map, Mercator map, and Pannini projection [21]. These global methods, while are able to remove some degree of distortions, often suffer from various residual distortions and at the loss of FoV. Many unwanted distortions remain un-corrected, as illustrated in Fig. 1. Taking the perspective rectification for example, while a perspective mapping always keeps straight lines straight, it unnaturally stretches shapes in the scene especially at the peripheral of an image [19]. For camera lenses whose FoV exceed 60°, stereographic, Mercator, or Pannini projections [21, 23] are popular choices to rectify distortion, as shown in Fig. 1. These projections only partially preserve shapes along certain directions, but cause stretches in other directions. Stereographic projection, as a special conformal mapping, preserves circular shapes but often bends long straight lines in the scene.

Local Adaptive Rectification. In a classic work, Carroll *et al*. [4] applied local optimization to correct curved linear structures while preserving the natural shape of objects. However, visual contents on the boundary are missing. More importantly, it alters orientations of local shape (*e.g.*, the girl's head in Fig. 1(h)), which requires much user interaction and fixed line-orientation constraints. Given the obvious radial symmetry of camera distortion, the conformal mapping based on Cartesian coordinates and axisymmetric Mercator projection, we believe, is the reason. It

also leads their method to stretch scene at poles and even destroy linear structure. In contrast, the proposed automatic method based on polar coordinates rectifies the image visually pleasing. Recently, Shih *et al*. [22] (at Google AI) rectify distortions on facial regions starting from a perspective image (Fig. 1(d)). However, noticeable perspective distortions on other parts of the image remain, and become more evident with wider FoV, as shown in Fig. 1(e). As such, their method fails on ultra wide-angle lens (FoV above 120°), whereas the proposed method achieves a good balance among different distortions. Locally adaptive projection idea has also been used in other tasks of image manipulation, for example for content-aware image resizing [3, 5, 16] and panorama squaring [7, 13, 29]. The goal is to manipulate the image while keeping those perceptually salient image contents as intact as possible. With simple modification, the framework of our method can be adapted to solve those tasks as well.

3. Conformal Mapping in Polar Domain

To solve the wide-angle rectification problem, we optimize a geometric transformation which *warps* the image such that its visually salient features are undistorted, while keeping its wide field of view. In this paper, we employ *Conformal Mapping* as the particular form of the geometric transformation. Mathematically, a conformal mapping is defined as a *holomorphic* function between two domains that preserves both angles and local shapes.

Conformal maps are typically formulated in the Cartesian domain. As discussed in Sec. 2, Carroll's method based on Cartesian domain alters the shape-orientations, stretches scene at poles, and destroy linear structure, (also see in Fig. 7). However, due to the obvious radial symmetry of camera lenses (also see in Fig. 3), it is more convenient to derive wide-angle rectification methods in the polar domain. Polar domain is a two dimensional coordinates in which each position is determined by an angle and a radial distance with respect to a center. Formally, the polar transform between the polar domain and Cartesian domain (image space (u, v)) is defined [9, 31] in complex form:

$$\mathcal{U} = u + iv = \rho(\theta)\cos\phi + i\rho(\theta)\sin\phi, \tag{1}$$

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.

3.1. Polar-form Cauchy-Riemann Condition

To take advantages of polar domain in our optimization, we first derive a polar-form conformal mapping. According to Eq. (1) and [2], $\mathcal{U}$ could be considered as a mapping from a polar domain (θ, ϕ) to a image space, and $\mathcal{U}$ is conformal

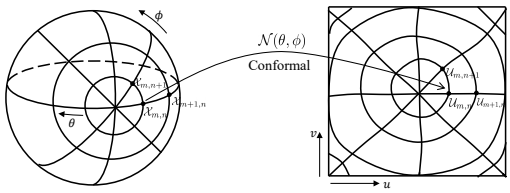
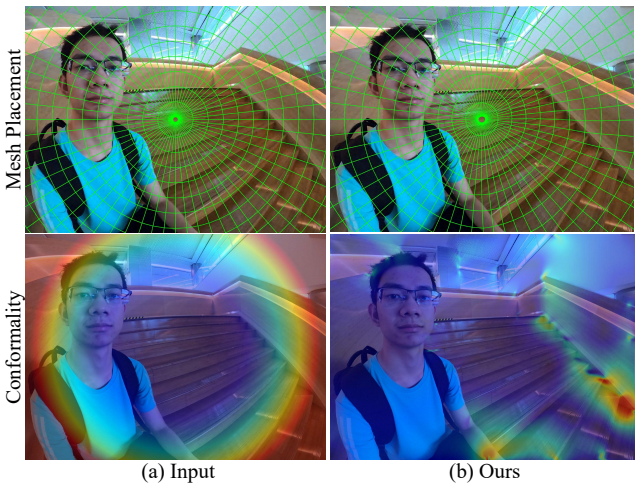


Figure 2. By a polar-form conformal mapping $\mathcal{N}$, angles and shapes are locally preserved in the polar gradient directions.



(a) Input (b) Ours

Figure 3. This figure (top row) shows the polar meshes before and after our optimization. The bottom row visualizes the conformalities (*i.e.* shape-preserve-ness) before and after the optimization.

if it satisfies Cauchy-Riemann condition in polar domain,

$$\mathcal{N}(\theta,\phi)\frac{\partial \mathcal{U}}{\partial \theta} = \frac{\partial \mathcal{U}}{\partial \phi}, \tag{2}$$

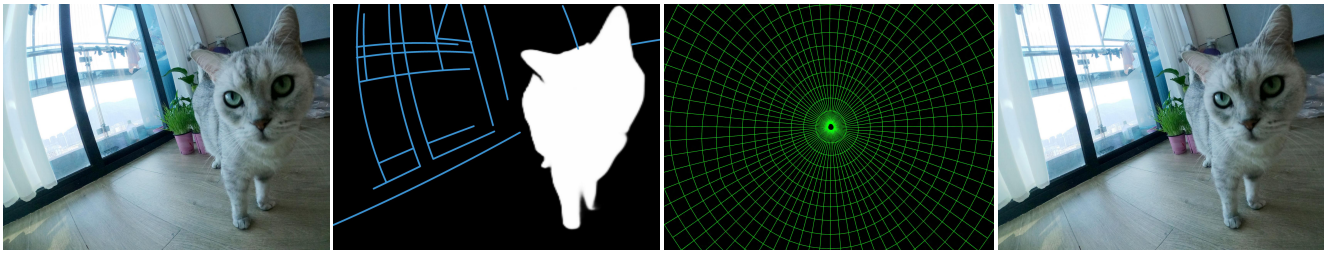
where $\mathcal{N}(\theta, \phi)$ represents the orthogonal preservation of polar gradient directions under mapping $\mathcal{U}$ in the complex form, as shown in Fig. 2.

Stereographic and Mercator projections are popular conformal maps for local shape preservation [21, 31]. Compared with axisymmetric Mercator projection used in [4], stereographic projection is radially symmetric and more suitable for polar parameterization [9]. We then obtain its polar-form Cauchy-Riemann condition:

$$\frac{\partial \mathcal{U}}{\partial \phi} - i\sin\theta\frac{\partial \mathcal{U}}{\partial \theta} = 0. \tag{3}$$

3.2. Optimal Mesh Placement

Since Eq. (3) cannot be strictly enforced in general, we first formulate the task as an optimization defined on the polar mesh. Vertices on the polar mesh are uniformly sampled in radial and angular directions as shown in Fig. 3(a). We then parameterize the polar mesh $\boldsymbol{M}$ as $\{\boldsymbol{u}_{m,n}\}$ in the



(a) Input (b) Visual features (c) Optimized mesh (d) Output

Figure 4. Pipeline: (a) given an input wide-angle image with obvious wide-angle distortions; (b) we first extract salient visual features including curved lines and prominent regions, using deep neural nets (see Sec. 4.1); (c) we then optimize the polar mesh by minimizing the energy function defined in Sec. 4.2; (d) Finally, our obtain our rectification result, where most lines are rectified and local shapes preserved, and the original wide field of view (given by the rectangular image frame) maintained.

image plane. $\boldsymbol{u}_{m,n}=(u_{m,n}, v_{m,n})$ is the Cartesian coordinates of the polar vertex, *as per* Eq. (1) and polar coordinates $(\theta_{m,n}, \phi_{m,n})$, where m and n denote the indices of polar vertices in the radial and angular directions, respectively. Fig. 3 shows an example of mesh placement before and after our optimization.

3.3. Least-Squares Polar Conformal Map

To introduce Eq. (3) to polar mesh placement, we afterward derive a least-squares approximation to the Cauchy-Riemann equation in the polar domain. This can be viewed as a natural extension of the LSCM (least-squares conformal mapping) as reported in [12]. Strictly speaking, an LSCM is not a mathematically valid conformal mapping, but an approximate, or quasi-conformal mapping.

Within each of the quad-cells in the polar domain, angle-preserving Eq. (3) is enforced approximately in the least-squares sense. Suppose that the polar cell is small enough, we discretize the domain such that the conformal mapping is locally linear. Denote $\mathcal{U}_{m,n}$ as a vertex of a polar cell, $\mathcal{U}_{m+1,n}, \mathcal{U}_{m,n+1}$ denote the vertices in the polar gradient directions, as shown in Fig. 2, we derive a linear transformation on the polar cell according to Eq. (3). Conformality criterion C in the least-squares sense is therefore given as:

$$C(\mathcal{U}_{m,n}) = \frac{\partial \mathcal{U}}{\partial \phi} - i \sin\theta \frac{\partial \mathcal{U}}{\partial \theta} = \begin{bmatrix} 1 - i\sin\theta \\ i\sin\theta \\ -1 \end{bmatrix}^{\top} \begin{bmatrix} \mathcal{U}_{m,n} \\ \mathcal{U}_{m+1,n} \\ \mathcal{U}_{m,n+1} \end{bmatrix}. \quad (4)$$

Fig. 3 illustrates an example of the computed conformality before and after the optimization. Warmer (red) colors indicate higher distortions. It can be seen that, before optimization, the violations of conformality increase radially as they get closer to image boundaries. In contrast, after optimization those values are significantly reduced, suggesting more local shapes are preserved.

4. Automatic Content-Aware Rectification

In a nutshell, our goal aims to search for an optimal content-aware approximate conformal mapping on a polar mesh, to correct a wide-angle image so that global lines become more straight and local shapes are kept. The overall pipeline is illustrated in Fig. 4.

4.1. Deep Image Content Analysis

To develop a fully-automatic content-aware rectification method, the first stage of our pipeline is automated image contents analysis. By "image content", in this paper, we only attend to global (long) linear structures (*e.g.* vertical pillars, ceilings, window frames) and visually salient regions (*e.g.* foreground objects). This is because human eyes pay more attention to these salient features when observing a wide field of view, as shown in Fig. 5. Other visual features can also be handled similarly by our framework.

For this purpose, we simply employ existing state-of-the-art deep neural networks with modification. Specifically, for linear structure detection, we adopt and modify the Wireframe-Net proposed in [8]. Their network is based on Pyramid Residual Modules and Stacked Hourglass network [18]. However, [8] is designed for detecting straight lines only, which is unfit for our task at hand. Therefore, we present a curvilinear line perception network (CLP-Net) for wide-angle images. The curvilinear heat map tailored for our task is defined as $h(\boldsymbol{u}) = \begin{cases} S(d) & \boldsymbol{u} \in \boldsymbol{l} \\ 0 & \boldsymbol{u} \notin \boldsymbol{l} \end{cases}$, where $S(d) = (1 + e^{-d/(4D)})^{-1}$ refers to a *radial sigmoid function*, d is the distance from the line to the image center. D is the half image diagonal distance. Using 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. We simulate a wide-angle distorted version of the wireframe dataset used in [8] under different lens models and re-train the network to detect curved lines.

For automatic salient region detection, we directly apply a pre-trained model using a boundary-aware salient object detection network (BAS-Net) proposed by Qin *et al.* [20]. We then normalize the salient weights w^{Sa} in range $[0, 1]$ before adding them to local weights to improve shape

Figure 5. Image content analysis. The top, and bottom rows are the raw inputs, and the detected salient regions by BAS-Net (white) as well as curved line segments by CLP-Net (blue). From left to right, the fields of view increase from 100° to fisheye 180°.

preservation in those salient regions. It is worth noting that other algorithms for curved line detection and salient region segmentation may also be used. Previously we tested the well-known LSD algorithm for line segment detection, but found it only detects small and short segments hence it is unfit for our purpose.

4.2. Energy Formulation

We then design an energy function to search for the optimal shape-preserving LSCM transformation, satisfying a set of additional structure-preserving conditions (*i.e.*, straight lines, and the rectangular FoV boundaries).

Local Shape Preserving. Our shape-preserving term E_C encourages each polar cell to undergo a linear LSCM transformation, as discussed in Sec. 3.3. It is defined as

$$E_C = \sum_{\mathcal{U}_{m,n} \in \boldsymbol{M}} \left\| \omega^C_{m,n} C(\mathcal{U}_{m,n}) \right\|^2, \tag{5}$$

where C is the conformality criterion (Eq. (4)). $\omega^C_{m,n}$ is the spatially-varying weight of each polar vertex, which is initialized as 1. Image content analysis allows us to adaptively minimize the violation of conformality in places where it is most likely to be noticed. We add the normalized salient weights to $\omega^C_{m,n}$ to strengthen shape preservation on the noticeable regions. To avoid excessive distortion, we also apply the radial sigmoid function $S(d) = (1 + e^{-d/(4D)})^{-1}$ to weights near the endpoints of curved lines, where d is the distance from line segments on the polar cell to the image center.

Line Preserving. Our line-preserving energy term is inspired by previous works [4,26], but is modified to fit for the polar domain. We also discard the orientation-preserving term in [26], for it is already taken care of by the LSCM (see [12]). We subdivide all the curved lines into multiple line segments, one in each of the polar cells. To propagate the line direction between cells, we require that all the line segments belonging to the detected curve share the same or similar direction. To speed up the optimization, we only evaluate the line-preserving term at the midpoint of each segment.

Let $\boldsymbol{u}^{\boldsymbol{l}}_k$ denote the midpoint in the k-th cell-line of $\boldsymbol{l}$. Its position is defined by bilinear interpolation of its four neighboring vertices in polar form, *i.e.* $\boldsymbol{u}^{\boldsymbol{l}}_k = a\boldsymbol{u}_{m,n} + b\boldsymbol{u}_{m+1,n} + c\boldsymbol{u}_{m,n+1} + d\boldsymbol{u}_{m+1,n+1}$. These bilinear coefficients are constant during optimization. We also define the rectified orientation unit vector $\boldsymbol{e}(\boldsymbol{u}^{\boldsymbol{l}}_s, \boldsymbol{u}^{\boldsymbol{l}}_e)$ of $\boldsymbol{l}$ using the difference vector of its two endpoints $\boldsymbol{u}^{\boldsymbol{l}}_s$ and $\boldsymbol{u}^{\boldsymbol{l}}_e$. Consequently, for $\boldsymbol{l}$, the line-preserving term $E_{\boldsymbol{l},n}$ is defined as,

$$E_{\boldsymbol{l},n} = \sum_k \left\| \boldsymbol{e}(\boldsymbol{u}^{\boldsymbol{l}}_s, \boldsymbol{u}^{\boldsymbol{l}}_e)^\top \boldsymbol{u}^{\boldsymbol{l}}_k \right\|^2_2, \tag{6}$$

where k indexes the k-th line segment. The total line-preserving energy E_L is the sum of terms $\{E_{\boldsymbol{l},n}\}$,

$$E_L = \sum_n \omega_{\boldsymbol{l},n} E_{\boldsymbol{l},n} \tag{7}$$

where n indexes detected curved lines. Considering wide-angle images have more severe deformation on the boundary, $\omega_{\boldsymbol{l},n}$ is defined to increase the straightness of the lines close to the boundary. Based on Sec. 4.1, we compute $\omega_{\boldsymbol{l},n} = S(d)$ by defining distances between the image center and the line connecting two endpoints $\boldsymbol{u}^{\boldsymbol{l}}_s$ and $\boldsymbol{u}^{\boldsymbol{l}}_e$.

Unlike the method proposed by Carroll *et al.* [4], there is no need to manually mark the curved lines with fixed orientations in our formulation. It is also helpful to use the endpoints in Eq. (6) instead of orientation, since the locations of the lines are optimized in the process. Besides, they only use the orientation vector from one endpoint to optimize the straight line. The points neighboring the endpoint are more sensitive compared to other sides during optimization.

Smoothness. The line-preserving term and conformality may cause "seams" between the neighboring transformations, especially on visual features. To remedy this, we also

add a smoothness energy term which acts as the regularization term to ensure the obtained geometric warping is smooth everywhere on the image plane.

We directly minimize a second-order energy term, given by the squared norm of the Hessian, as opposed to [4] that enforces a second-order condition on the warping function:

$$E_S = \left\| \frac{\partial^2 \mathcal{U}}{\partial \theta^2} \right\|_2^2 + \left\| \frac{\partial^2 \mathcal{U}}{\partial \phi^2} \right\|_2^2 + 2 \left\| \frac{\partial^2 \mathcal{U}}{\partial \theta \partial \phi} \right\|_2^2. \tag{8}$$

Eq. (8) is then discretized via finite difference approximations to the second derivatives and defined as a smoothness energy term E_S based on polar mesh,

$$\begin{aligned} E_S = & \sum \left\| w^C_{m,n} \left(\boldsymbol{u}_{m+1,n} - 2\boldsymbol{u}_{m,n} + \boldsymbol{u}_{m-1,n} \right) \right\|_2^2 \\ & + \sum \left\| w^C_{m,n} \left(\boldsymbol{u}_{m,n+1} - 2\boldsymbol{u}_{m,n} + \boldsymbol{u}_{m,n-1} \right) \right\|_2^2 \\ & + 2 \sum \left\| w^C_{m,n} \left(\boldsymbol{u}_{m+1,n+1} - \boldsymbol{u}_{m,n+1} - \boldsymbol{u}_{m+1,n} + \boldsymbol{u}_{m,n} \right) \right\|_2^2, \end{aligned} \tag{9}$$

where $w^C_{m,n}$ is the spatially-varying weight that equals to shape-preserving term.

Boundary Preserving. In order to retain the rectangular field of view, we apply force to points on the four sides of the image boundary. Similar to the term used for line preserving, we treat the four boundaries as four straight lines, and interpolate the midpoints for each line segment via bilinear interpolation in the polar domain. $\boldsymbol{u}^b_k = (u^b_k, v^b_k)$ denotes the k-th midpoint of a boundary divided by polar mesh. The boundary preserving energy term E_B is written as:

$$E_B = \sum_{\boldsymbol{u}^b_k \in L} \left\| u^b_k \right\|_2^2 + \sum_{\boldsymbol{u}^b_k \in R} \left\| u^b_k - W \right\|_2^2 + \sum_{\boldsymbol{u}^b_k \in T} \left\| v^b_k \right\|_2^2 + \sum_{\boldsymbol{u}^b_k \in B} \left\| v^b_k - H \right\|_2^2, \tag{10}$$

where $L/R/T/B$ refer to the left/right/top/bottom image boundaries, and W/H are the width/height of the target frame.

4.3. Optimization Procedure

The total energy function is the weighted combination of energy terms from Eqs (5), (7), (9) and (10),

$$E = \lambda_C E_C + \lambda_L E_L + \lambda_S E_S + \lambda_B E_B. \tag{11}$$

In experiments, we find that line preservation (λ_L) is more important than shape preservation (λ_C) because the linear structures are distorted severely compared with local shapes in a wide-angle image. Furthermore, the boundary preservation weight λB is sensitive: a higher value leads to overstretched areas near the boundary, whereas a lower value results in an irregular frame. Empirically, we fix λ_C, λ_L, λ_S and λ_B to 1, 100, 2 and 10 throughout the paper. *Supplemental material* contains more details and ablation study.

To speedup the optimization, we initialize vertices by applying stereographic projection, and ignore the curved lines near the projection center. Furthermore, distortions of short line segments are less perceivable, so only line segments longer than a pre-defined length are used.

Table 1. Quantitative evaluations of our method and baselines.

	Carroll's	Shih's	Ours	Ours$_{\text{w.o.}E_B}$
StraightAcc↑	0.9743	0.9907	0.9864	0.9872
ShapeAcc↑	0.9407	0.8771	0.9683	0.9736
ConformalAcc↑	0.7639	0.5794	0.8259	0.8458

Ours$_{\text{w.o.}E_B}$ indicates our methods without boundary preservation.

5. Experiments

5.1. Implementation Details

We apply Levenberg-Marquardt optimizer [11, 17] to minimize the energy in Eq. (11). Our code is written in C++ with Ceres solver [1], on a standard laptop (single-core). We achieve good convergence usually after 4-10 iterations. Since the core computation involves only constrained Least squares apart from the feed-forward networks, the computation is fast by our un-optimized CPU code. We compare the proposed method with both global methods and local methods. For methods proposed by Carroll *et al*. [4] and Shih *et al*. [22], we set the mesh resolutions at 192×122 and 201×105, resp., while ours is 100×180 in all tests. The implement details could be found in the *supplemental material*. Salient visual features extracted by deep neural networks are fed into the optimization. Note that [22] uses perspectively-corrected images as its input, in which lines are already corrected, but at the cost of losing FoV (after cropping) and amplification of perspective distortion.

5.2. Quantitative Evaluation

To demonstrate the performance of our method, we conduct quantitative comparisons with other methods, as shown in Tab. 1. We introduce three measures, *i.e*. StraightAcc, ShapeAcc, and ConformalAcc, for evaluation. Specifically, according to the line-preserving term of Eq. (7) and straightness in [10], we define the straight line measure as, $\text{StraightAcc} = \min\left(\left\{ \frac{1}{1 + \| \boldsymbol{e}(\boldsymbol{u}^{\boldsymbol{l}}_s, \boldsymbol{u}^{\boldsymbol{l}}_e)^\top \boldsymbol{u}^{\boldsymbol{l}}_k \|_2} \right\}_k \right)$, which is computed from the endpoints $\boldsymbol{u}^{\boldsymbol{l}}_s, \boldsymbol{u}^{\boldsymbol{l}}_e$ of marked lines and the midpoint $\boldsymbol{u}^{\boldsymbol{l}}_k$ of k-th line segment. StraightAcc is close to 1 if the distortion of the line segment is low.

For shape preservation, ShapeAcc [30] is used. Considering ShapeAcc only works on portrait regions, ConformalAcc of salient regions based on the conformality Eq. (4) is also defined,

$$\text{ConformalAcc} = \frac{\min\left(\{ \| \mathbf{u}_{m+1,n} + \mathbf{u}_{m,n+1} - 2\mathbf{u}_{m,n} \|_2 \} \right)}{\max\left(\{ \| \mathbf{u}_{m+1,n} + \mathbf{u}_{m,n+1} - 2\mathbf{u}_{m,n} \|_2 \} \right)}. \tag{12}$$

Being closer to 1 means less distortion on the local shape around the salient region. For fair comparisons, all methods use the same Cartesian meshes for evaluation. As shown in

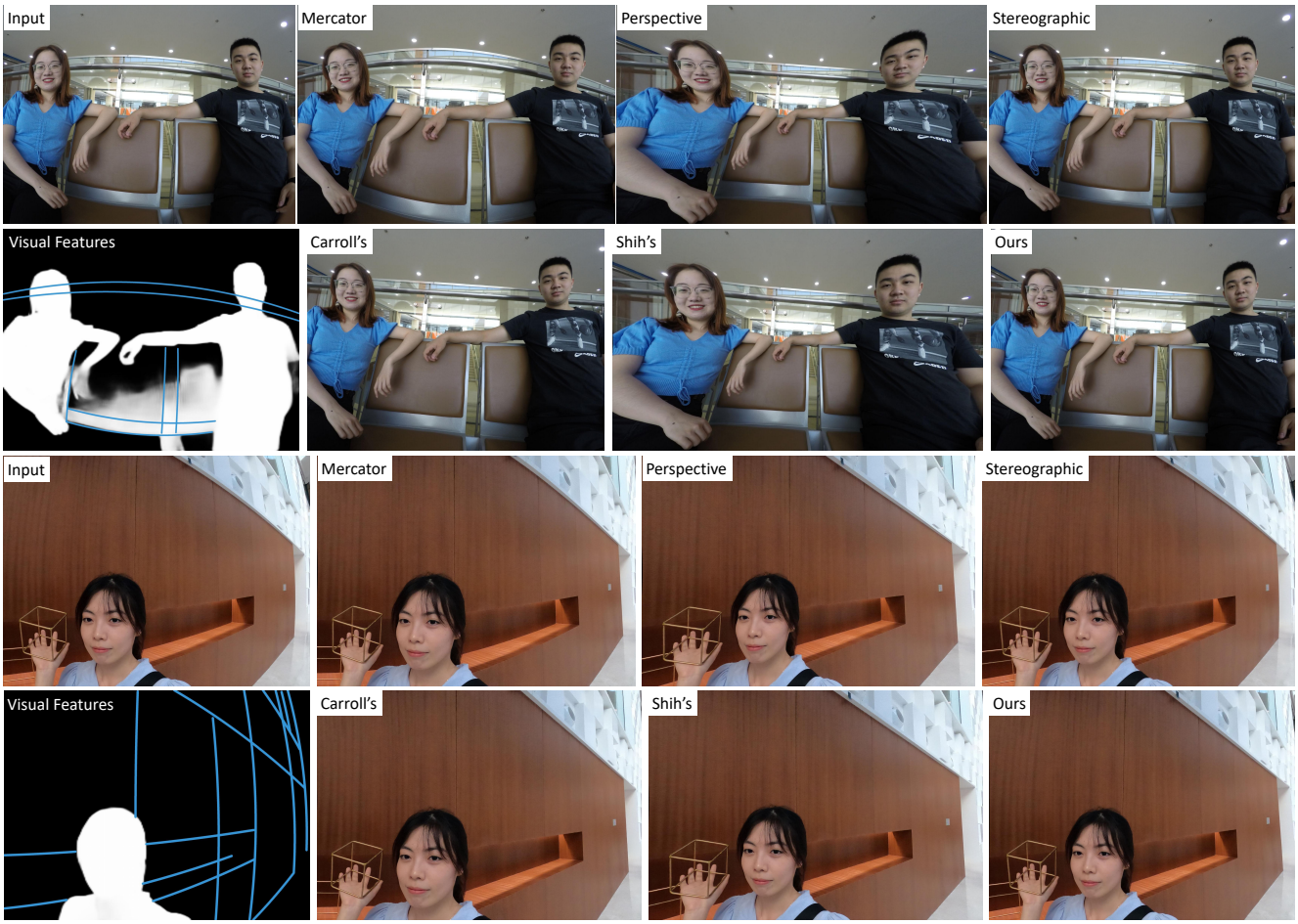


Figure 6. Example results of our rectification compared with global projection methods and local methods, by Carroll *et al*. [4] and Shih *et al*. [22]. FoVs of each input are 149° and 117°. Compared with state-of-the-art methods, our method has effectively removed the wide-angle distortions while at the same time maintaining an ultra-wide field of view, without sacrificing salient image contents.

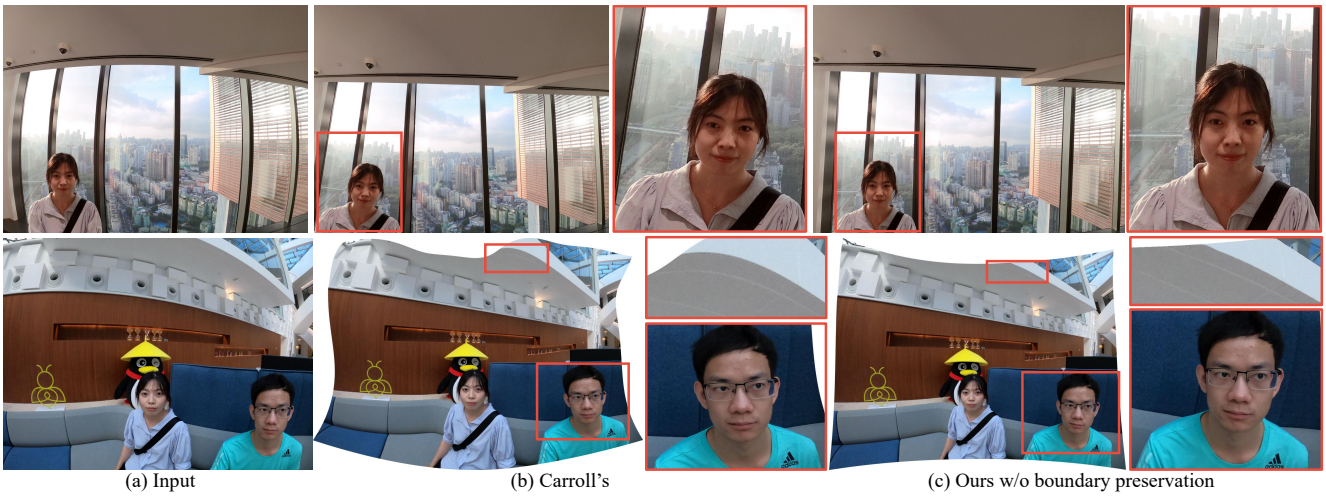
Tab. 1, our method achieves a decent balance between local shapes (ShapeAcc and ConformalAcc) and linear structures (StraightAcc). Meanwhile, our method obviously outperforms baselines in terms of ShapeAcc and ConformalAcc. And there are only slight differences in StraightAcc between ours and baselines, especially Shih's method. Also, with the introduction of boundary preservation, the evaluations of our method slightly decrease.

5.3. Qualitative Evaluation

We demonstrate the performance of the proposed method on test examples. Additional results and ablation studies are included in the *supplemental material*. Our experiments confirm the superior performance of the proposed method on shape, orientation, line and boundary preserving, compared with all previous state-of-the-arts.

Fig. 6 illustrates the comparison with global projections and local methods. Global projections, while are able to reduce certain type of distortions, often leave other residual distortions. In particular, wide FoV is well beyond the typical limit for perspective images, causing severe stretching. Note Shih's method uses perspective images as inputs to correct facial regions only. We can see that wide FoVs lead to the visual disharmony between corrected portraits and perspective regions (*e.g.* body in first scene or cube in second scene) and suffer severe loss of image content. It also bends the linear structures near local shapes, such as guardrail behind the portraits. Carroll's method preserves local shapes and linear structures, but alters the field of view as well as orientation (*e.g.* the heads in first scene). Compared with them, our method achieves a decent balance between distortion-minimization and FoV-retaining.

To demonstrate the benefits of our polar-form LSCM, we also perform our method without the boundary-preserving term, as shown in Fig. 7. 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 preserva-



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

Figure 7. The comparison between our method without boundary-preserving term and the method proposed by Carroll *et al*. [4]. The only difference of two methods is the conformal constraints and the corresponding smoothness in polar and Cartesian domains.

tion. For example, the white floor of the second scene in Fig. 7 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. 7. That is why the original Carroll's method requires manually fixing some lines' orientation. 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.

5.4. User Study

Because our ultimate goal is to create wide-angle visual experiences without noticeable and irritating distortions for a human user, we recruit 60 participants to conduct the user study. Among them, 5 are technical artists, 10 are graduate students working in computer vision and graphics, 15 are research scientists, and the rest are volunteers from the general public. We collect 150 wide-angle images from Flickr, from 100° to 180°. These images contain both local objects and global linear structures. For each participant, we randomly sample 30 images, and compare the results obtained by our method, and by another method side-by-side, and report his/her preferred result. The order of displaying the pair is entirely random and totally blind to the participant. Results of our User Study are given in the table below, which confirms that our method consistently outperforms all competing methods by a large margin. (We did not test [22] because it only works on perspectively-corrected images as input and behaves poorly for Larger FoV).

Table 2. User Preference Study (Ours *Versus* Others)

Method	Perspective	Mercator	Stereographic	Carroll
Ours →	95.2%	91.6%	86.5%	74.3%

6. Closing Remarks

Our method has a few limitations in the following aspects, and showcases could be found in the *supplemental material*. 1). the visual quality of the rectified images is influenced by image content analysis, especially the curved line detection. 2). there is a trade-off between shape-preserving and line-bending. When a curved line cuts across one local shape (*e.g.* face), the shape may be bent, because the weight applied to line-energy is heavier.

In this paper, we combine the advantages of traditional geometric vision and modern learning-based methods for wide-angle image rectification. Specifically, deep networks are used to detect scene contents (*e.g.* curved line segments and salient regions) which are crucial for structural analysis but hard to optimize. To perform the locally adaptable shape-preserving transformation, a least-squares polar conformal mapping is developed. A set of energy terms is introduced to leverage geometry knowledge (*e.g.* shape-preserving conformal map and line-preserving constraint) to achieve content-aware image rectification. Experiments on a large number of images confirm the superiority of the proposed method.

Acknowledgements. The authors would like to thank Ying Feng and Xiaoyu Li 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. 6

[2] Lars Valerian Ahlfors. *Conformal invariants: topics in geometric function theory*, volume 371. American Mathematical Soc., 2010. 3

[3] Moab Arar, Dov Danon, Daniel Cohen-Or, and Ariel Shamir. Image resizing by reconstruction from deep features. *arXiv preprint arXiv:1904.08475*, 2019. 3

[4] Robert Carroll, Maneesh Agrawal, and Aseem Agarwala. Optimizing content-preserving projections for wide-angle images. *ACM Trans. Graph.*, 28(3):43, 2009. 2, 3, 5, 6, 7, 8

[5] Che-Han Chang, Chia-Kai Liang, and Yung-Yu Chuang. Content-aware display adaptation and interactive editing for stereoscopic images. *IEEE Trans. Multimedia*, 13(4):589–601, 2011. 3

[6] Frederic Devernay and Olivier Faugeras. Straight lines have to be straight. *Machine Vision and Applications*, 13(1):14–24, 2001. 2

[7] Kaiming He, Huiwen Chang, and Jian Sun. Rectangling panoramic images via warping. *ACM Trans. Graph.*, 32(4):1–10, 2013. 3

[8] 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. 4

[9] 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. 3

[10] Yeong Won Kim, Chang-Ryeol Lee, Dae-Yong Cho, Yong Hoon Kwon, Hyeok-Jae Choi, and Kuk-Jin Yoon. Automatic content-aware projection for 360 videos. In *Int. Conf. Comput. Vis.*, pages 4753–4761. IEEE, 2017. 6

[11] Kenneth Levenberg. A method for the solution of certain nonlinear problems in least squares. *Quart. Appl. Math*, 2(2):164–168, 1944. 6

[12] Bruno Lévy, Sylvain Petitjean, Nicolas Ray, and Jérome Maillot. Least squares conformal maps for automatic texture atlas generation. *ACM Trans. Graph.*, 21(3):362–371, 2002. 4, 5

[13] Dongping Li, Kaiming He, Jian Sun, and Kun Zhou. A geodesic-preserving method for image warping. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 213–221, 2015. 3

[14] Hongdong Li and Richard Hartley. A non-iterative method for correcting lens distortion from nine point correspondences. In *IEEE International Conference on Computer Vision Workshops*, 2005. 2

[15] Kang Liao, Chunyu Lin, Yao Zhao, and Mai Xu. Model-free distortion rectification framework bridged by distortion distribution map. *IEEE Trans. Image Process.*, 29:3707–3718, 2020. 2

[16] Feng Liu and Michael Gleicher. Automatic image retargeting with fisheye-view warping. In *Proceedings of the 18th annual ACM symposium on User interface software and technology*, pages 153–162, 2005. 3

[17] Donald W Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11(2):431–441, 1963. 6

[18] Alejandro Newell, Kaiyu Yang, and Jia Deng. Stacked hourglass networks for human pose estimation. In *Eur. Conf. Comput. Vis.*, pages 483–499. Springer, 2016. 4

[19] Jon Peddie. The history of visual magic in computers. *Springer, London. https://doi. org/10*, 1007:978–1, 2013. 2

[20] 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. 4

[21] Thomas K Sharpless, Bruno Postle, and Daniel M German. Pannini: A new projection for renderingwide angle perspective images. In *Computational Aesthetics*, pages 9–16, 2010. 2, 3

[22] 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. 2, 3, 6, 7, 8

[23] John P Snyder. *Flattening the earth: two thousand years of map projections*. University of Chicago Press, 1997. 2

[24] Hans Strasburger, Ingo Rentschler, and Martin Jüttner. Peripheral vision and pattern recognition: A review. *Journal of vision*, 11(5):13–13, 2011. 2

[25] Jean-Philippe Tardif, Peter Sturm, and Sébastien Roy. Self-calibration of a general radially symmetric distortion model. In *Eur. Conf. Comput. Vis.*, pages 186–199. Springer, 2006. 2

[26] Jin Wei, Chen-Feng Li, Shi-Min Hu, Ralph R Martin, and Chiew-Lan Tai. Fisheye video correction. *IEEE Trans. Vis. Comput. Graph.*, 18(10):1771–1783, 2011. 2, 5

[27] Zhucun Xue, Nan Xue, Gui-Song Xia, and Weiming Shen. Learning to calibrate straight lines for fisheye image rectification. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 1643–1651, 2019. 2

[28] Xiaoqing Yin, Xinchao Wang, Jun Yu, Maojun Zhang, Pascal Fua, and Dacheng Tao. Fisheyerecnet: A multi-context collaborative deep network for fisheye image rectification. In *Eur. Conf. Comput. Vis.*, pages 469–484, 2018. 2

[29] Lihi Zelnik-Manor, Gabriele Peters, and Pietro Perona. Squaring the circle in panoramas. In *Int. Conf. Comput. Vis.*, volume 2, pages 1292–1299. IEEE, 2005. 3

[30] Fushun Zhu, Shan Zhao, Peng Wang, Hao Wang, Hua Yan, and Shuaicheng Liu. Semi-supervised wide-angle portraits correction by multi-scale transformer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 19689–19698, 2022. 6

[31] Denis Zorin and Alan H Barr. Correction of geometric perceptual distortions in pictures. In *SIGGRAPH*, pages 257–264, 1995. 3