https://doi.org/10.2352/lim.2024.5.1.10

A True Panoramic Camera for Smartphone Applications

Jian Ma [1], Shenge Wang [1], Yunwen Li [2], Adrian Giura [1], Chris Miclea [1], and Sergio Goma [1]
[1] Qualcomm Technologies, Inc. 5775 Morehouse Dr. San Diego, CA 92121, USA
[2] Qualcomm Semiconductor Limited, No. 16-1 Zhanye 2nd Rd. East District, Hsinchu City, 300091, Taiwan

Abstract

We introduce our cutting-edge panoramic camera – a true panoramic camera (TPC), designed for mobile smartphone applications. Leveraging prism optics and well-known imaging processing algorithms, our camera achieves parallax-free seamless stitching of images captured by dual cameras pointing in two different directions opposite to the normal. The result is an ultra-wide (140°x53°) panoramic field-of-view (FOV) without the optical distortions typically associated with ultra-wide-angle lenses.

Packed into a compact camera module measuring 22 mm (length) x 11 mm (width) x 9 mm (height) and integrated into a mobile testing platform featuring the Qualcomm Snapdragon® 8 Gen 1 processor, the TPC demonstrates unprecedented capabilities of capturing panoramic pictures in a single shot and recording panoramic videos.

Introduction

Capturing panoramic pictures and videos is a delightful experience with smartphone cameras. However, the current landscape presents challenges: no commercially available cameras can seamlessly capture panoramic images in a single shot or record panoramic videos. The following is a brief recap:

Cumbersome process and artifacts: Taking a panoramic picture using a smartphone camera involves steady rotation while continuously capturing frames. This process can be cumbersome for users. Despite efforts, artifacts often plague panoramic images. These include curved horizons, ghosting, stitching errors, and vertical banding.

Large optical distortion: Previous attempts at ultra-wide-angle cameras relied on using ultra-wide lenses, necessitating digital distortion correction [1-2]. Unfortunately, this approach falls short in delivering the high-quality results comparable to today's smartphone cameras.

Bulky optics arrangement: Some experiments involved bulky optics arrangements with multiple cameras [3-4]. However, such solutions are impractical for mobile devices.

Modern smartphone ultra-wide cameras achieve a diagonal field of view (FOV) of approximately 120°. While they excel in capturing wide scenes, optical and perspective distortions become pronounced, especially with close-up subjects. Additionally, their horizontal FOV, crucial for panoramic views, is often limited to less than 100°.

Thus, the quest for high image quality panoramic photography continues and advancements are needed to enhance the user experience while maintaining image quality.

Our Approach

To create a true panoramic camera (TPC), we utilize two conventional cameras, each positioned to capture different horizontal views opposite to the normal (approximately ±35°). Through seamless image stitching from these two cameras, we achieve a 2X increase in the horizontal field of view (HFOV). The use of conventional lenses in each camera ensures minimal distortions. Consequently, the resulting panoramic view remains free from distortions.

However, when the two cameras are spatially separated, seamless image stitching becomes problematic due to parallax. Parallax occurs because the disparity, which represents the relative position of objects between the images captured by the two cameras, changes laterally with distances. This distance-dependent variation leads to misalignment during the stitching process. To mitigate parallax, we must ensure that the entrance pupils (the effective apertures through which light enters the cameras) overlap spatially. This concept is illustrated in Figure 1. Unfortunately, achieving perfect spatial overlap is physically impossible.

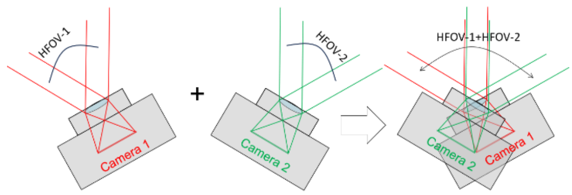


Figure 1. *The concept of achieving 2X increase of horizontal FOV (HFOV) with two cameras*

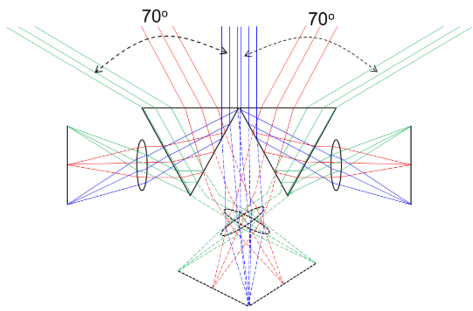


Figure 2. *The folded optics design for the true panoramic imaging*

Our solution involves a folded optical design. We employ two equilateral prisms arranged as shown in Figure 2. The key feature is that one surface of both prisms that faces each other is coated with a high-reflectance material. This reflective surface forms a virtual image of the optical aperture located at the center behind the prisms. By ensuring that the centers of the two virtual apertures align precisely, we achieve parallax-free seamless stitching. In the figure, the solid lines represent the actual optical paths, while the dotted lines depict the virtual paths.

Optical Design and Precision Assembly

Virtual images are generated through reflective surfaces. The deliberate choice to utilize prisms, rather than relying solely on mirrors, serves two essential purposes: reducing vignetting and expanding the field of view (FOV).

Prisms made from materials with a high refractive index significantly benefit the FOV. This property is demonstrated in Figure 3, which illustrates the relationship between relative illumination of the TPC and the object field angle for prisms with a side length of 7mm. Specifically, we compare two materials: BK-7 (with a refractive index of 1.5) and N-SF11 (with a refractive index of 1.8). The N-SF11 prism exhibits higher relative illumination at field angles greater than ±55 degrees. Furthermore, by increasing the size of the prism, it can also effectively reduce vignetting, resulting in a broader FOV.

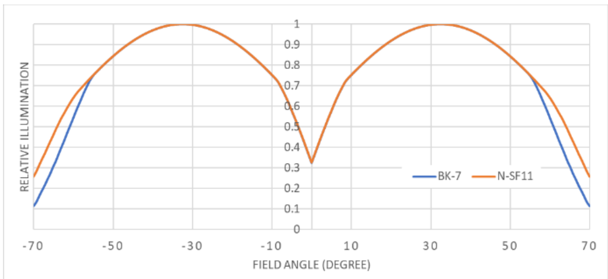
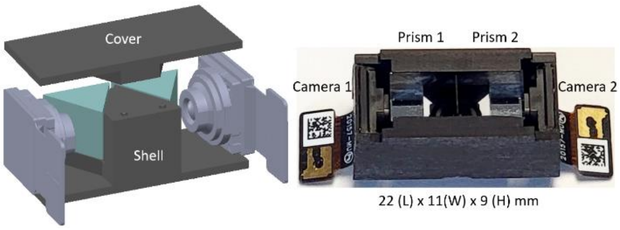


Figure 3. *TPC relative illumination with BK-7 and N-F11 prisms.*

Precision optical alignment is required to achieve true parallax-free imaging. The lateral alignment accuracy of the two cameras (specifically the accuracy of center overlap) is calculated as $\Delta x = (d \times p)/f$, where d represents the closest object distance, p is the allowable disparity at the sensor plane, and f denotes the focal length. For instance, assuming d is 10 cm, p is 1 µm, and f is 3 mm, we obtain Δx = 33 um, which is achievable. However, the angular alignment accuracy is less forgiving, given by $\Delta\theta = p/f$. Using the same parameters, we find $\Delta\theta = 3.3 \times 10^{-4}$ radians. Achieving such precise angle alignment is feasible with an active alignment scheme; nevertheless, maintaining this accuracy over the device's lifespan poses challenges, particularly due to angular shifts caused by thermal expansion of adhesives and components. Although such misalignment is small (e.g., several pixels), it must be addressed. We have developed a real-time dynamic stitching algorithm to mitigate such small angular misalignment, which is discussed in a later section.

The device components are illustrated in Figure 4(a). It includes a shell, two 7mm prisms, two off-the-shelf fixed-focus smartphone front-facing cameras, and a shell cover. The alignment process includes the following: (1) Align and fix the prisms inside the shell. (2) Align and fix the cameras. (3) Glue the shell cover. The active alignment procedure was used to ensure parallax free operation by ensuring that the perfect alignment of targets is achieved at both near field and far field simultaneously. A finished TPC module is shown in Figure 4(b).



(a) *(b)*

Figure 4. *(a) TPC components 3D model. (b) A working prototype.*

Imaging Perspective Correction

The implementation of the TPC uses the images captured by two cameras, each pointing in different directions. These images require perspective correction to create a merged or stitched image that matches the perspective of a single camera pointing directly forward (as shown in Figure 5). For instance, consider a flat surface with a checkerboard pattern placed in front of a TPC. The left and right cameras capture separate images, depicted in Figure 6(a). By applying proper perspective transformations and stitching these images together, we obtain the result shown in Figure 6(b). This merged image demonstrates the correct aspect ratio and faithfully represents the shape of the checkerboard pattern.

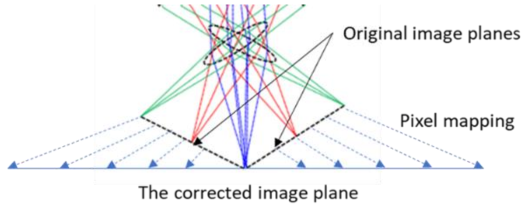


Figure 5. *TPC perspective correction pixel mapping*

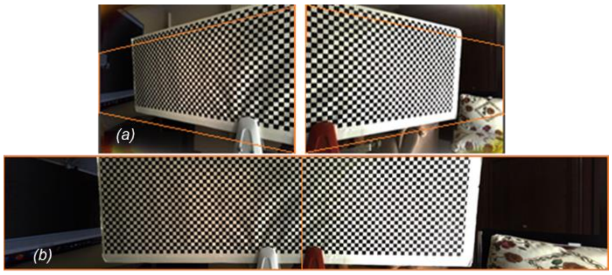


Figure 6. *(a) Images captured by the two cameras in a TPC. (b) Stitched and perspective corrected TPC image.*

The perspective transformation involves mapping image content from two trapezoids shown in Figure 6(a) to corresponding rectangles in Figure 6(b) for the left and right cameras. To achieve this, we derive the necessary perspective transform for each camera using a specified quadrilateral within the captured image area as the mapping source and another quadrilateral within the TPC image area as the mapping destination. Specifically, the image content within the quadrilateral defined by the four corners, captured by the left camera, is transformed using a perspective transformation function into the left half of the TPC image area. Similarly, the image content within the quadrilateral defined by the four corners, captured by the right camera, undergoes another perspective transformation to map it into the right half of the TPC image area.

Perspective transformations can be mathematically described using these expressions: X=Fx(U,V) and Y=Fy(U,V), where, (X, Y) represents the panoramic image coordinates, (U, V) corresponds to the camera pixel coordinates, Fx and Fy denote the transfer functions for the X and Y dimensions, respectively.

In Figure 6, we observe that when a camera captures a scene from an angle, it includes more content compared to viewing the scene directly forward. Consequently, certain image content—such as areas outside the trapezoidal field of view—captured by the two cameras is not utilized in creating the TPC image. Additionally, to achieve seamless stitching, we intentionally introduced a small overlapping area between the left and right cameras. We will delve into this critical aspect further in the upcoming sections.

Image Stitch

The TPC system combines two images captured by separate cameras to create a wide-angle panoramic image. This process involves two stitching techniques: static stitch (SS) and dynamic stitch (DS). SS occurs during the assembly and calibration of the TPC module. It aligns the images precisely, and a stable base panorama is established. DS corrects small residual misalignments that may occur over time due to environmental factors (such as temperature changes or vibrations). To enable DS, a small overlap between the captured images by the two cameras is necessary. This overlap is achieved by slightly tilting the cameras—say, by approximately 0.5 degrees (equivalent to about 30 pixels). By combining SS and DS, the TPC system ensures high-quality wide-angle panoramas, even in challenging conditions. Let's delve into the details:

Imagine a scene with capital alphabet letters arranged continuously from green-letters A to Z and then red-letters A to Z. Initially, without any overlap, the left camera captures only the entire green letters, while the right camera captures only the entire red letters.

Figure 7. (a) Images see by the TPC two cameras with overlapping areas (shaded). (b) A stitched TPC image

By adjusting the camera angles, we ensure that the left camera sees the red-letter A at the right sensor edge, and similarly, the right camera sees the green-letter Z at the left sensor edge. As depicted in Figure 7(a), both cameras simultaneously capture the green Z and red A within the overlapping sensor areas (highlighted in shaded regions). The two dotted lines represent the stitch lines.

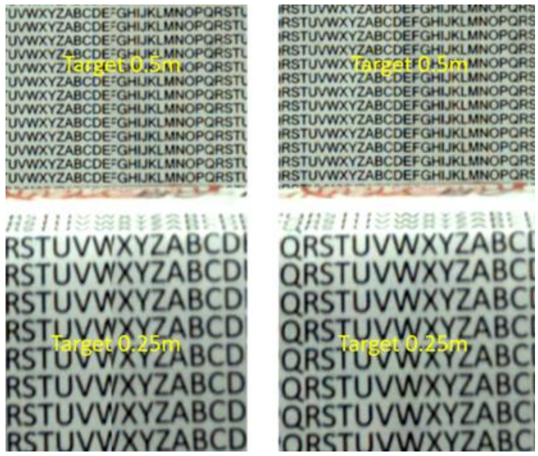
In Figure 7(b), we show the outcome of the final stitched image. To achieve seamless stitching, we made a trade-off by sacrificing 0.5 degrees on each side. This adjustment is visually represented by the green-letter A and red-letter Z, which are subtly shifted beyond the sensor areas, as depicted in Figure 7(a).

The overlapping areas between two images play a critical role in achieving seamless image stitching. For a flawless stitch, the two overlapping images should be identical, as demonstrated in Figure 7(a)—specifically, the green-letter Z and the red-letter A. However, misalignments can occur, resulting in differing content between the left and right images in the overlapping area. To address misalignment, a process of digital realignment is executed. The goal is to eliminate differences between the overlapping portions, ultimately achieving a perfect stitch.

The amount of stitch misalignment can be distance dependent when there is a disparity. Even though it is a weak dependence since the disparity is normally very small, a global shift of the images won't necessarily yield a perfect stitch because scene content along the stitch line may vary at different distances. We employed pattern matching techniques, including normalized cross-correlation (NCC). The narrow overlap between the two images minimizes computational overhead, facilitating efficient frame-to-frame comparison and correction.

Seamless correction need not encompass the entire image. Instead, it can be selectively applied to a small zone near the stitch line. As the horizontal position moves away from the stitching line, the correction gradually fades to zero.

By iteratively refining the alignment dynamically, we achieve a seamless and visually coherent stitched image in live situations. This was demonstrated experimentally in Figure 8. Two targets were strategically placed at different distances: 0.25 meters and 0.5 meters. These targets were simultaneously captured by the TPC. The 0.25m target exhibits a misalignment of 6 pixels (the letter W). In contrast, the 0.5m target experiences a smaller misalignment of only 3 pixels (the letter F). Attempting a global lateral digital shift won't achieve perfect alignment. Remarkably, the DS algorithm successfully aligns both targets in real-time.



(a) *(b)*

Figure 8. Pictures captured by TPC: (a) without DS, (b) with DS

Static and Dynamic Gain Correction

Gain correction plays a crucial role in producing high quality TPC imaging. It comprises two components: static gain correction (SGC) and dynamic gain correction (DGC). SGC is akin to the traditional lens shading correction (LSC) used in regular cameras, but it requires higher spatial resolution due to the presence of prism vignetting in addition to lens shading effects. A correction lookup table (LUT) is derived by uniformly illuminating the panoramic camera with diffused white light. Figure 9 visually demonstrates a TPC image before and after SGC.

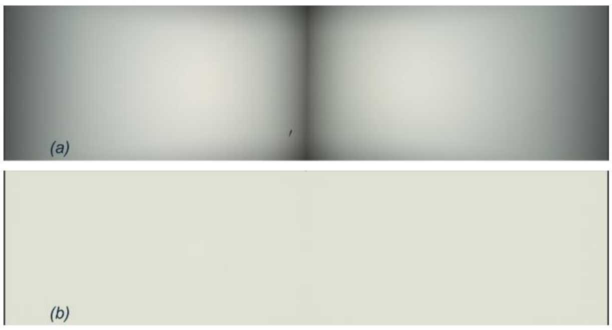


Figure 9. Pictures captured by TPC with the illumination of a uniform white light: (a) without SGC, (b) with SGC.

When dealing with environmental illumination that exhibits strong nonuniformity, relying solely on SGC falls short. Imperfections in the prism and vignetting disparities between the two sides introduce subtle variations in brightness as lighting conditions change. Even the tiniest difference in brightness becomes apparent in the stitch line between two images. While adjusting exposure time can address overall brightness, localized discrepancies persist along the stitching boundary.

Consider a scenario where a powerful point light source exists on either side of the camera. This results in uneven brightness near the stitching boundary. Detecting these variations involves calculating the average brightness for each vertical position in the overlapping area along the stitch line in both left and right images. Subsequently, real-time, frame-by-frame gain correction is applied along the stitch line.

To avoid false detection caused by brightness differences due to real-world objects (such as black and white strips) in the middle stitch area, the DGC method targets a specific correction zone and employs thresholding. Furthermore, a blending algorithm [5] seamlessly integrates the correction area with the rest of the image.

In Figure 10(a), we observe strong window light, resulting in a slightly brighter right side and a visible stitch line. Figure 10(b) demonstrates that DGC effectively conceals the stitch line.

Figure 10. *Pictures captured by TPC: (a) without DGC, (b) with DGC.*

Implementation on a Mobile Test Platform

The panoramic camera module was installed in a mobile testing platform (MTP) featuring the Snapdragon® 8 Gen 1 processor (Figure 11). The TPC image processing pipeline is illustrated in Figure 12 and described below:

Figure 11. *A TPC module integrated in an MTP.*

The objective of the TPC imaging processing pipeline is to stitch the output from two sensors and to perform image blending. Its primary goal is to minimize mismatches across the stitch line. This pipeline operates in real time on the MTP hardware. The computational tasks are handled by both the ISP and the GPU.

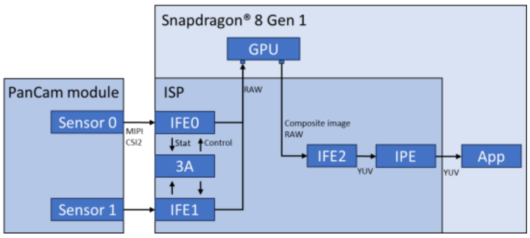


Figure 12. *TPC image processing pipeline.*

As illustrated in Figure 12, initially, each sensor produces an MIPI CSI2 stream. On the receiving end, the stream is decoded and processed by the IFE block. The static gain correction (SGC) is implemented within the lens shading correction (LSC) module inside the IFE block, which is fine-tuned based on the characteristics of the TPC LUT to account for prism vignetting. Furthermore, the 3A (AE, AF, and AWB) of the two sensors are synchronized to minimize any stitching mismatches between the two images. Subsequently, the GPU works on Bayer raw images and executes the TPC dynamic correction algorithms (DS and DGC), ultimately stitching the two images. Additionally, the SGC may be included in this stage for flexibility.

After stitching, the composite image undergoes processing by the IFE and IPE blocks in the ISP. During this stage, the two blocks operate in a conventional setup, but with a configuration optimized for the wide FOV resolution.

The result of the image processing pipeline is a single output image that provides a panoramic view. The key features are:
Field of View (FOV): 140° (horizontal) x 53° (vertical).
Panoramic Preview: Instantly preview the panoramas before capturing them.
Single-Shot Panoramas: Capture panoramic pictures at a resolution of 15360 x 3880 pixels in a single shot.
Panoramic Video: Record panoramic videos at 7297 x 1940 pixels, with a frame rate of 30 fps.

Figure 13 demonstrates TPC technology in a real-world scenario. Figure 14 compares the same scene captured by TPC and a cutting-edge smartphone ultra-wide camera, highlighting TPC's advantages: a broader HFOV and no noticible distortion.

Figure 13. *Picture captured by TPC.*

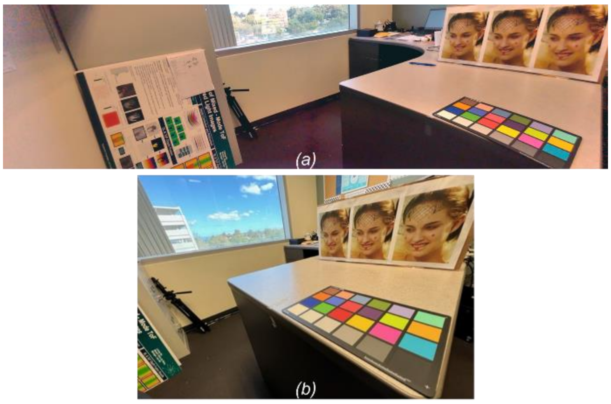


Figure 14*. Pictures captured by: (a) TPC, (b) smartphone UW camera.*

Future Works

Our forthcoming efforts will center around the following key objectives:

1. Reducing TPC module z-height to less than 7mm: Achieving this goal involves downsizing the camera module package.
2. Incorporating auto-focus capabilities: Auto-focus functionality can play a crucial role in enhancing TPC image quality. However, for tilted cameras, like those in the TPC, the optimal auto-focus approach should adhere to the Scheimpflug condition by incorporating a distance-dependent tilt [6].

References

[1] G. Krishnan and S. K. Nayar, "Cata-Fisheye Camera for Panoramic Imaging", 2008 IEEE Workshop on Applications of Computer Vision, January 2008

[2] H. Kim, er al, "Fisheye lens camera based surveillance system for wide field of view monitoring", Optik, Vol.127, Issue 14, 2016

[3] K.H. Tan, et al, "Multiview Panoramic Cameras Using a Mirror Pyramid", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26, Issue: 7, July 2004

[4] T. Nakao and A. Kashitani, "Panoramic camera using a mirror rotation mechanism and a fast image mosaicking", Proceedings 2001 International Conference on Image Processing, Oct. 7, 2001

[5] P. J. Burt and E. H. Adelson, "A multiresolution spline with application to image mosaics", ACM Transactions on Graphics, Vol. 2, No. 4, October 1983, Pages 217-236.

[6] https://en.wikipedia.org/wiki/Scheimpflug_principle