Efficient Utility of Dual Cameras in a Mobile Device

Sachin P. Kamat
Samsung India Software
Operations Pvt. Ltd.
Bangalore
India

ABSTRACT

Modern handheld devices like mobile phones and tablets feature two digital cameras - one in the front and one at the back. Though the back camera (which is of higher resolution and quality) is used for general purpose photography, under certain specific conditions the photographs taken are of unacceptable quality. This paper describes an efficient utility of the dual cameras to produce good quality still photographic pictures making use of hardware features provided by the camera controller and the SoC in cases where normal photography through the single camera does not produce good results. In the absence of a sophisticated image processing/editing software in mobile devices, this technique provides an easy way to create those effects in minimal time with less power by processing the foreground and background images obtained from the two cameras separately and then blending the two. Experimental results provided at the end of the paper show promising results.

General Terms

Digital imaging applications.

Keywords

Dual camera, image processing, mobile device.

1. INTRODUCTION

Advances in embedded and manufacturing technology have made it possible for realizing cameras with resolutions of 12 mega pixels (MP) or more to be integrated into small form factor devices like smart phones, tablet PCs, notebooks, etc. Initially a single low resolution camera mostly of VGA resolution (640x480) remained the state of the art. It was either on the front or on the back of the device depending on the utility. Certain manufacturers provided a rotatable camera module on the top of the device that could face the user or the opposite side as required by the user. Over the years, reductions in the costs of these peripherals have made the device manufacturers to provide multiple cameras on the product. Currently gadgets with up to three cameras (one on the front and two on the back) to record stereoscopic data are available in the market.

In a dual camera scenario, one camera is provided in the front of the device generally above the screen (referred to as front camera in this paper) and one at the back (referred to as back camera in this paper). The camera in the front is generally a low resolution camera supporting upto 2 mega pixels. This camera is used for making video calls or taking self portraits. The camera at the back is a high resolution camera of several mega pixels and is used for taking snapshots and recording video. This camera is usually equipped with a flash light and has features like auto focus, optical/digital zoom, etc. Figure 1 shows the line diagram of a mobile phone to show the general placement of cameras. The image captured from the front camera is referred to as foreground image and the image captured from the back camera is referred to as background image in this paper.

With increasing capabilities of mobile cameras, they are being used as competent alternatives to digital still cameras for taking photographs. Several environmental factors affect the quality of photographs like lighting conditions, etc. This paper tries to address the problems caused by such factors in photography by taking advantage of the dual camera architecture present in these devices and the capabilities provided by the camera controller and the SoC.

The rest of the paper is organized as follows. Existing art and the problem statement is described in section 2. Section 3 discusses the proposed architecture. Section 4 describes the experimental details and conclusions are provided in section 5.

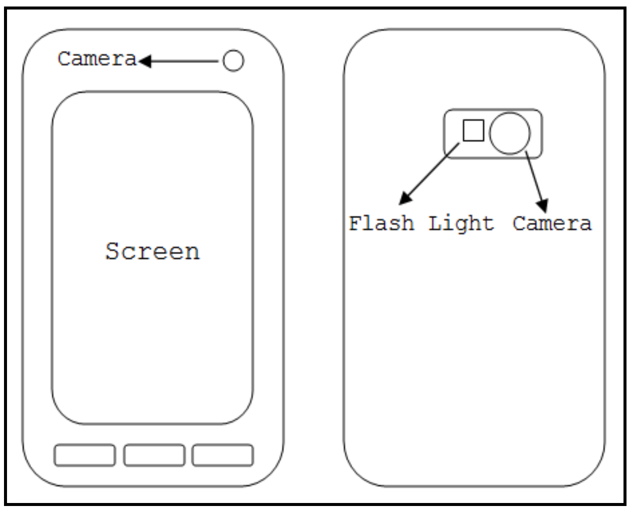


Fig 1: Block diagram of a mobile phone showing a) Front side camera above the screen facing the user and b) Back side camera facing away from the user.

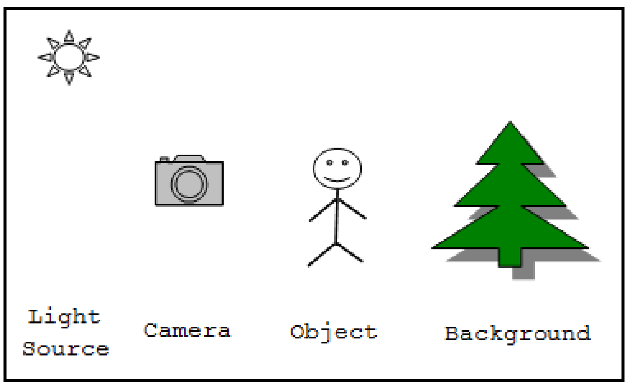
2. ANALYSIS OF EXISTING ART

This section builds the problem statement, analyses the current art in this area and tries to evaluate its merits and demerits.

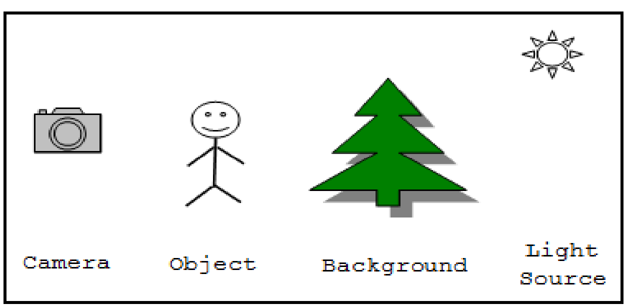
One of the important considerations for taking good photographs is to ensure that the object to be photographed is sufficiently illuminated and that the source of light does not cause any shadows, dark areas or glares on the object and the camera.

Good results are obtained by having the source of light and the object on either sides of the camera as shown in Figure 2(a). Figure 2(b) illustrates the case where the placement of source of light is such that it casts a shadow on the wrong side thereby creating dark areas on the subject and giving a glare on the camera lens.

When the picture to be taken is of a permanent structure like an edifice or of a landscape, it is not possible to avoid the unfavorable conditions. Dedicated cameras make use of polarizers, filters and such other accessories to minimize the glare [1]. However, this is not possible with cameras embedded on mobile devices. Another possible solution is to subject the captured image (with artifacts and distortions) to image processing applications to do an offline restoration. Sometimes multiple pictures are taken in a sequence using different parameters and image quality is improved [2], [9]. Image processing applications are resource intensive and require powerful processors/graphics cards and considerable amount of free memory to produce desired results. Secondly, generally the mobile screen is too small (though we now have large screen tablets) and the input device is usually the touch which makes it difficult to use manual image processing tools to correct or restore the image. Though, the processing power and system memory in latest smart phones has gone up considerably, it is still short of producing output in a short span of time using software processing.



(a)



(b)

Fig 2: The most preferred and the least preferred positions of different objects while photographing.
(a) Camera in between the source of light and the object to be photographed produces well illuminated good pictures.
(b) Source of light and object on one side of the camera does not produce good results.

Even if we are ready to compromise on the processing time, the power consumed by these applications is too high to be worthy of using them as a solution. Another disadvantage of mobile photography comes into picture while shooting in low light conditions. In a darker surrounding, the flash light needs to be enabled (auto enabling option is present in most of the camera applications). The flash being of low capacity is just sufficient to illuminate the object which is closer to it and the background remains totally dark. These kinds of pictures are difficult to process even with software and require lot of effort from the user.

State of the art camera sensors come with auto focus feature [3], [4] that automatically focuses the objects in its focal plane and blur out other background objects. This feature might not be very desirable some times when the user wants a uniform focal plane.

3. PROPOSED METHOD

This section details the proposed method and provides an analysis of the advantages of the technique.

Advanced system-on-chips (SoCs) found in latest gadgets come with camera controllers that can interface multiple cameras using various interface standards. Figure 3 shows the simplified block diagram of one such camera controller found in a latest SoC. It has four ports that can be used to control four camera sensors. The controller sends commands to the camera sensors and receives the data from them. This data could be either in raw format (RGB, YCbCr, etc) or compressed format like JPEG.

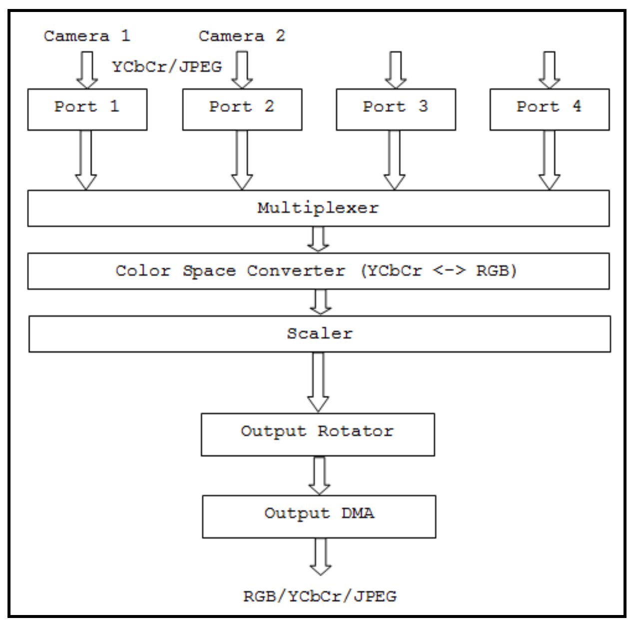


Fig 3: Simplified block diagram of a camera controller on a state of the art mobile SoC that can interface four different camera sensors.

During the normal operation, the data captured by the sensor in YCbCr or JPEG form (depending on the sensor and its configuration by the software) is sent to the multiplexer, which in turn sends the data to the color space converter. If the data needs to be converted to RGB format for doing some image processing or sending it for display on a LCD like in the case of camera preview, it is done here. The scaler resizes the image if required. On a need basis, the output rotator rotates the images by some fixed degree of rotation and sends the output to the DMA block. The camera controller also has the capability to impart certain basic special effects to the image like embossing, sepia, negative, etc. These kinds of effects are also provided by advanced camera sensors. Reference [5] claims a method to process dual images on a FPGA system but the paper does not provide sufficient implementation details.

The proposed method consists of making use of the front camera for capturing the object and the back camera for capturing the background images. This scenario is illustrated in Figure 4 below.

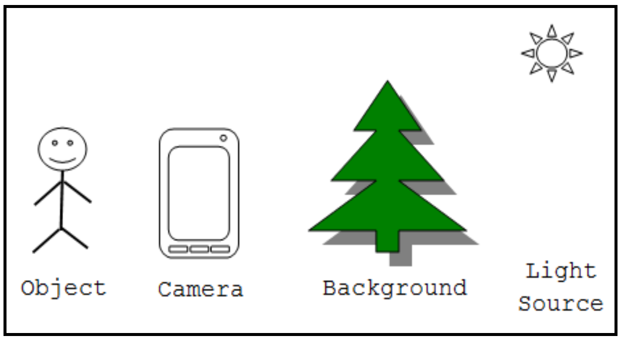


Fig 4: Representation of the proposed technique where the camera is placed in between the object and the background. The front camera captures the object and the back camera captures the background.

The front camera is a low resolution camera of 1.3 MP and the rear camera is of 5 MP resolution. Front camera is connected to port 1 and back camera is connected to port 2 of the camera controller. The images are captured at a small time offset of around 50 milliseconds (20 FPS) in YCbCr 420 format and stored in memory mapped regions. The image captured from the front camera, called foreground image, is sent to the image signal processor (ISP) module for further processing. This includes the following steps:

- Edge detection [6]
- Extraction of the image
- Blending it with the background image

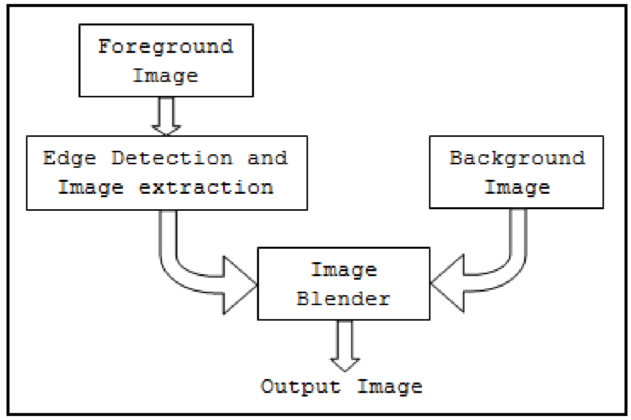


Fig 5: Block diagram of the image processing module.

Since the image transformation operations are supported in the hardware itself, they do not cause any performance overhead and can be applied selectively either to foreground or background image only.

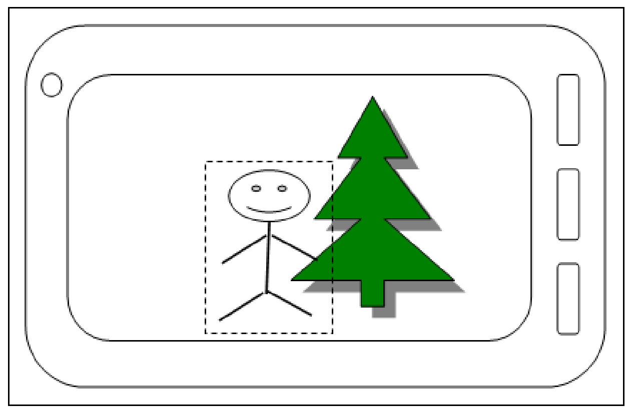


Fig 6: The extracted foreground image window (in dotted lines) is provided to the user for placing on the background image. Once done the two images are blended together and saved.

The extracted foreground window is presented to the user for placing it on the background image. This is done using alpha blending and overlaying feature provided by the display hardware. Thus there is no software processing overhead involved. This is shown in Figure 6.

Reference [7] describes a flash/no-flash based technique to extract the foreground image. Though this is an effective method to extract foreground image, this method is not suitable for mobile phones for the following reasons:

1. The flash in a mobile camera is not so powerful to produce enough differentiating brightness between foreground and background images.
2. Using flash too often consumes lot of power and drains the battery quickly.
3. The method itself is demonstrated under low light conditions and might not work under bright sunlight.

The above argument holds good for [8], [9] also.

Discrete two-dimensional convolution is an important operation in any image processing algorithm [10]. Consider an image of MxN resolution. The convolution of this image involves MxN multiply and accumulate operations for each sample of the image. Thus the number of mathematical operations is directly proportional to the number of samples in the image. Higher the resolution of the image, more are the samples and hence more number of computations involved. Optimized convolution algorithms require much lesser operations for performing convolution on a sample. But even in this case, the total number of operations increases with the size of the image.

Similarly, any image processing operation also requires temporary buffers for storing intermediate results. This is generally equal to the size of the image itself. Thus larger resolution images require more memory for processing.

4. EXPERIMENTAL RESULTS

The experimental analysis shows the limitations of current photographic techniques in mobile device cameras and brings out the advantages of the proposed technique. Figure 7 shows the case where an object is photographed at twilight hour in the evening around 6.30 PM with flash enabled. However, due to limited intensity of flash light which is short range, the object under consideration (potted plant) is fully illuminated whereas the background is completely dark.

Figure 8 shows the case where the same object is shot with a low resolution camera (1.3 MP) with flash and the background image is shot with high resolution camera (5 MP) without flash and the two images are blended into a single image. For easy comparison, the potted plant was rotated by 180º and the image flipped along y-axis. Now the background image appears in its natural light and the foreground image appears in completely illuminated form thus giving a better image.

Fig 7: Photograph taken using back camera of mobile phone with flash enabled.

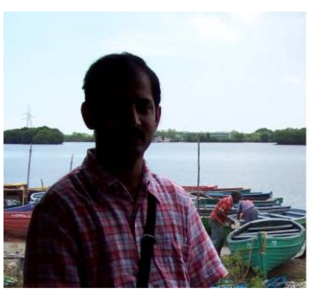
Fig 8: Image captured using two cameras separately and blended together. The foreground image taken using low resolution camera with flash to capture the plant and the background image (building) captured using high resolution camera without flash.

Once we have the two images obtained separately from two cameras, we have a lot of flexibility to manipulate them for producing any kind of special effects. Exemplary use cases are shown below in Figure 9 and Figure 10. Similar results were obtained when the source of light was behind the object. Figure 11(a) is the photograph taken using only the back camera. The face appears dark as the source of light is behind the person. Figure 11(b) is the blended picture taken using low resolution front camera for the person and high resolution back camera for the background. The picture looks uniformly illuminated. Figure 11(c) shows the complete process of creating the final blended image from the foreground and background images as described in Figure 5.

This technique also gives us the freedom to selectively zoom either the foreground image or the background image which is not possible with a single (lens) camera. Smile detection and face detection algorithms work well when the number of faces in the frame is less and the processing area is small. This also improves the processing time. By enabling these detections only on the front camera, we can reduce the overall computing, performance and power overheads on the device.

Fig 9: Foreground Image in grey scale and background image in natural mode.

Fig 10: Background image in sepia and foreground image in natural color.



(a) **(b)**

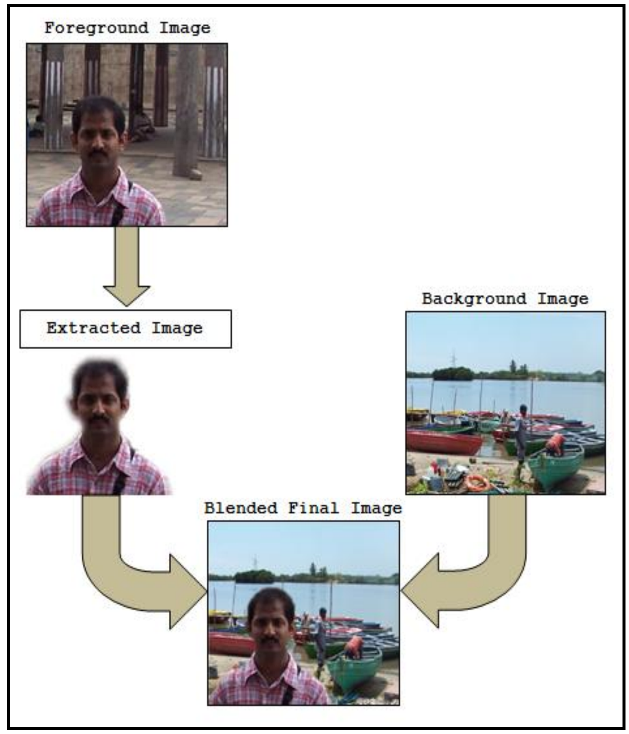


Fig 11: (a) Picture taken using only the back camera. (b) Picture taken using dual cameras (front and back) and blended together. (c) The image processing flow diagram with inputs, output and intermediate results.

Table 1 gives the details of the front and back camera of the mobile device used for the experiments. From the table it is clear that we need to operate only on 1.88 MB of data instead of the entire 7.18 MB for processing. This reduces the memory requirement and computational overhead thereby reducing the power consumed.

Table 1. Parameters of front and back camera modules

Parameter	Front Camera	Back Camera
Resolution	1.3 MP	5 MP
Width x Height	1280x1024	2592x1936
Frame Size (YCbCr420)	1.88 MB	7.18 MB

Since the front camera is of lower resolution, the quality of picture obtained is lower than that obtained from the back camera. Thus the proposed method is more suitable when regular photography using single camera makes it hard to obtain good quality pictures. The front camera needs to focus on the object properly for better results during edge detection and image extraction. This method can also be used when quick results in selectively adding certain special effects to foreground and background images are required irrespective of the external photographic conditions.

5. CONCLUSION

This paper demonstrated the technique of utilizing the hardware capability provided by modern day mobile SoCs in capturing and processing digital still images to produce good quality pictures under unfavorable conditions. The results are produced online in short time and the user does not have to depend on an offline image processing application for imparting certain basic special effects to the image. This technique can be applied to video recording as well with certain optimizations for achieving desired performance.

6. REFERENCES

[1] Pasricha, N. A Professional's Basic Photography. Black Rose Publications.

[2] Lasang, P., Ong, C. P. and Shen, S. M. 2010. CFA-based motion blur removal using long/short exposure pairs. IEEE Trans. Consumer Electron. Vol. 56, issue 2, 332-338.

[3] Rahman, M. T. and Kehtarnavaz N. 2008. Real-time face-priority auto focus for digital and cell-phone cameras. IEEE Trans. Consumer Electron. Vol. 54, issue 4, 1506-1513.

[4] Sohn, S. M., Yang, S. H., Kim, S. W., Baek, K. H., and Paik, W. H. 2006. SoC design of an auto-focus driving image signal processor for mobile camera applications. IEEE Trans. Consumer Electron. Vol. 52, issue 1, 10-16.

[5] Andorko, I., Corcoran, P. and Bigioi, P. 2011. A dual image processing pipeline camera with CE applications. In Proceedings of the IEEE International Conference on Consumer Electronics

[6] Castleman, K. R. 1998. Digital Image Processing. Prentice Hall.

[7] Sun, J., Kang, S. B., Xu, Z. B., Tang, X. and Shum, H. Y. 2007. Flash cut: foreground extraction with flash and no-flash image pairs. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition

[8] Petschnigg, G., Agrawala, M., Hoppe, H., Szeliski, R., Cohen, M. and Toyama, K. 2004. Digital photography with flash and no-flash image pairs. ACM Transactions on Graphics. Vol. 23, issue 3, 664-672.

[9] Drew, M. S., Lu, C. and Finlayson, G. D. 2006. Removing shadows using flash/noflash image edges. In Proceedings of the IEEE International Conference on Multimedia and Expo

[10] Jain, A. K. 2001. Fundamentals of Digital Image Processing. Prentice Hall of India.