October 17, 2019

Utilizing a Set of Processed First Frames for Expediting the Settling of Image-Processing Settings

Ruiduo Yang

YiChang Shih

Utilizing a Set of Processed First Frames for Expediting the Settling of Image-Processing Settings

Abstract:

This publication describes techniques and processes directed by an image capture manager to hasten the determination of appropriate image-processing settings, such as the "3A's" (Auto-Focus, Auto-Exposure, and Auto-White Balance), in an imaging device (*e.g.*, a digital camera, a smartphone). The image capture manager can direct an imaging processor (*e.g.*, image signal processor (ISP)) to improve the quality of the image by generating a set of processed first frames and performing object detection on the set to identify an object of interest (*e.g.*, a face). These operations are performed to the end that the object of interest may be identified earlier in the initialization process of the imaging device, and optimal image-processing settings can, consequently, be determined and applied sooner.

Keywords:

focus, Auto-Focus (AF), white balance, Auto-White Balance (AWB), exposure, Auto-Exposure (AE), 3A, imaging device, camera, face recognition, face detection, object detection, image-processing pipeline, frame, start-up, scene change

Background:

In order to capture an image using an imaging device, such as a digital camera or a smartphone with image-capturing capabilities, the imaging device is turned on and a boot-up process is initiated. After booting is complete, the imaging device can capture an image. Image capture on an imaging device will typically operate in accordance with a color image-processing pipeline (CIPP). Figure 1 illustrates the procedures of an example CIPP.

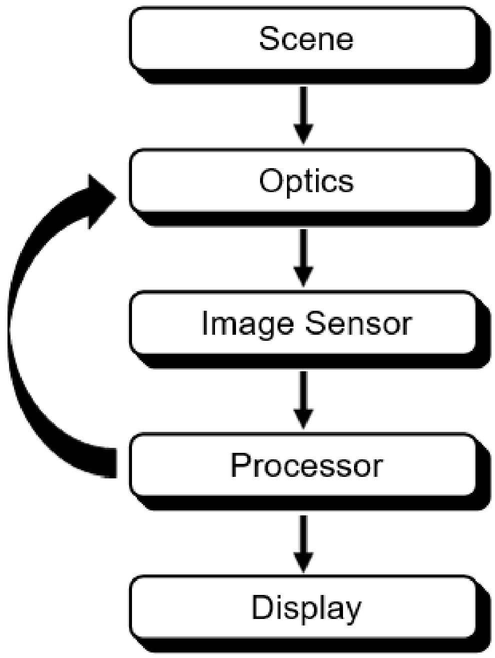

Figure 1

As Figure 1 illustrates, the CIPP operates in a loop, describing the setup required by the imaging device to generate an image. Utilizing the CIPP (the initialization process of the imaging device), the imaging device initially captures a scene under pre-set optic settings and generates RAW frame(s). A processor (*e.g.*, an image signal processor (ISP)) then applies image-processing settings to the RAW frame(s) to improve the quality of the image. The image-processing settings may include black level correction, noise reduction, 3A values (*e.g.*, Auto-White Balance, Auto-Exposure, Auto-Focus), *etc.* In addition, the imaging device may include an object detector (*e.g.*, a processor) for performing object detection.

The duration of the CIPP loop can vary depending on features of the scene, such as lighting conditions or the proximity of the object of interest. The object detector may initially have difficulty determining the presence of the object of interest due to unsuitable image-processing settings, such as low exposure. The object detector's inability to detect the object of interest causes the imaging device to receive and process additional frames. For example, if a face is poorly illuminated in an otherwise well-lit scene, then the object detector may not identify the face in the first frame. Consequently, the next frame will typically not have optic settings and/or image-processing settings adjusted to accommodate for the poor lighting on the face. In such a case, it may take the object detector several frames to finally detect the object of interest, or in this scenario, specifically, a face. Once the face is identified, then the image-processing settings will be readjusted to better expose the face. This readjustment entails more iterations through the CIPP loop that would have otherwise been unnecessary if the object detector identified the object of interest earlier in the process. In short, delays to the detection of the object of interest results in delays to the generation of a quality image for the user.

Therefore, it is desirable to have a technological solution for imaging devices where an object of interest is detected at an earlier stage within the initialization process, resulting in an expedited generation of a quality image for the user.

Description:

This publication describes techniques and processes for generating multiple inputs for face/object detection by an object detector to hasten (settle down) the determination of appropriate image-processing settings, such as the 3A's (*e.g.*, Auto-Focus (AF), Auto-Exposure (AE), Auto-White Balance (AWB)) and produce quality images on an imaging device. Figure 2 illustrates example imaging devices that can perform the operations described herein.

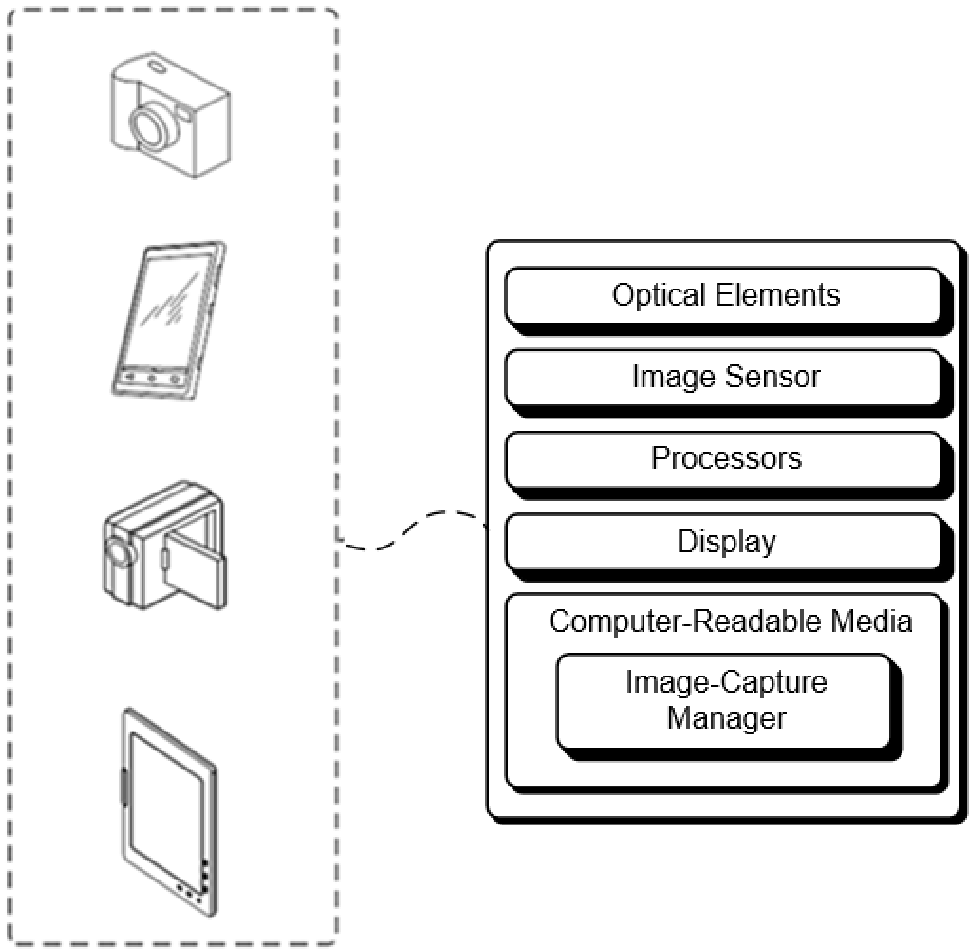

Figure 2

An imaging device can capture an image of a scene. In the example of Figure 2, the imaging device is illustrated as being a camera, a smartphone, a video recorder, or a tablet computer. The imaging device includes an image sensor (*e.g.*, a complementary metal-oxide semiconductor (CMOS) image sensor, a charged-couple device (CCD) image sensor) for detecting information used to make an image. The imaging device may include one or more optical elements (*e.g.*, a lens, a mechanical shutter, an electrical shutter, an aperture). The imaging device may also include a display for displaying a user interface. The user interface configured to receive input from a user of the imaging device. The user interface may include one or more of a touchscreen, a button, a dial, or a keypad. Inputs may include, for example, parameters that are associated with one or more imaging device settings, a selection or deletion of a captured image, or activities associated with post-processing the captured image (*e.g.*, change the rendered depth of focus of the captured image, crop the captured image, change the brightness of the captured image, change contrasting in the captured image).

The imaging device further includes at least one processor (*e.g.*, an image processor for processing images, an image detector for detecting objects). A processor may be a single or multiple core processor composed of a variety of materials, such as silicon, polysilicon, high-K dielectric, copper, and so on.

The image processor (*e.g.*, an Image Signal Processor (ISP)) of the imaging device is utilized to improve the quality of images generated by the imaging device through image-processing settings such as black level correction, noise reduction, AWB, AE, and/or AF. In some cases, an image processor may also operate as an image detector, performing object detection.

The imaging device also includes executable instructions of an image capture manager. The image capture manager may be implemented on the computer readable media (CRM) of the imaging device. The CRM may include any suitable memory or storage device such as random-access memory (RAM), read-only (ROM), or flash memory. The imaging device performs operations under the direction of the image capture manager to expedite the determination of appropriate image-processing settings (*e.g.*, Auto-Focus (AF), Auto-Exposure (AE), Auto-White Balance (AWB)). These operations include: 1) generating a set of processed first frames, and 2) performing an object detection on the set of first frames.

Upon the initialization of the image capture functionalities of the imaging device after booting is complete, a first digital image of a scene can be generated utilizing pre-set optic settings and image-processing settings. This first digital image is referred to herein as the “first frame.”

Opposed to a common color image-processing pipeline (CIPP) process where the optic settings and image-processing settings are adjusted immediately after the first frame is analyzed and additional frames are generated, the image processor processes the first frame to generate a set of first frames. Since the first frame is typically of a high dynamic range, usually a 10-bit or 16-bit image, more information is available within the first frame in contrast to the successive frames which are generally of a low dynamic range, commonly 8-bit images. This advantage makes the first frame the best candidate for digital manipulation and object detection.

In the process, the first frame is processed with a range of target image-processing settings such that each spot of the first frame will have an ideal image-processing setting applied to it at least once. For example, a range of exposure values can be set to the first frame, generating a set of first frames that range from over-exposed to under-exposed; this ensures that all areas within the first frame have been set to an ideal exposure value at least once.

Figure 3A and 3B, below, illustrate two extremes of image-processing the sharpness of a first frame.

Figure 3A

Figure 3B

In the above illustration, Figure 3A depicts the first frame that has been processed with a low sharpness, and Figure 3B depicts the first frame processed with a high sharpness. These two frames originate from the same first frame but have been processed with different target image-processing settings to produce two frames in the set of processed first frames. Figure 3A and 3B represent the two extremes, minimum and maximum respectively, in the digital manipulation of one image-processing setting, specifically sharpness. Other image-processing settings can be digitally manipulated as well, such as white balance, exposure, tone mapping, sharping, denoising, and so forth.

After generating a set of processed first frames, the imaging device performs operations under the direction of the image capture manager to detect an object of interest in the set of processed first frames. The detection of objects may be performed by an object detector. In some aspects, the image processor functions as the object detector. The object detector can receive the set of processed first frames one-by-one until the object detector identifies an object of interest. Due to the high dynamic range of the first frame and the broad range of image-processing settings in the set of processed first frames, the object detector has a greater likelihood of identifying the object of interest sooner in the initialization process.

Figure 4, below, illustrates the modification of the conventional CIPP loop to expedite the initialization process of an image devices.

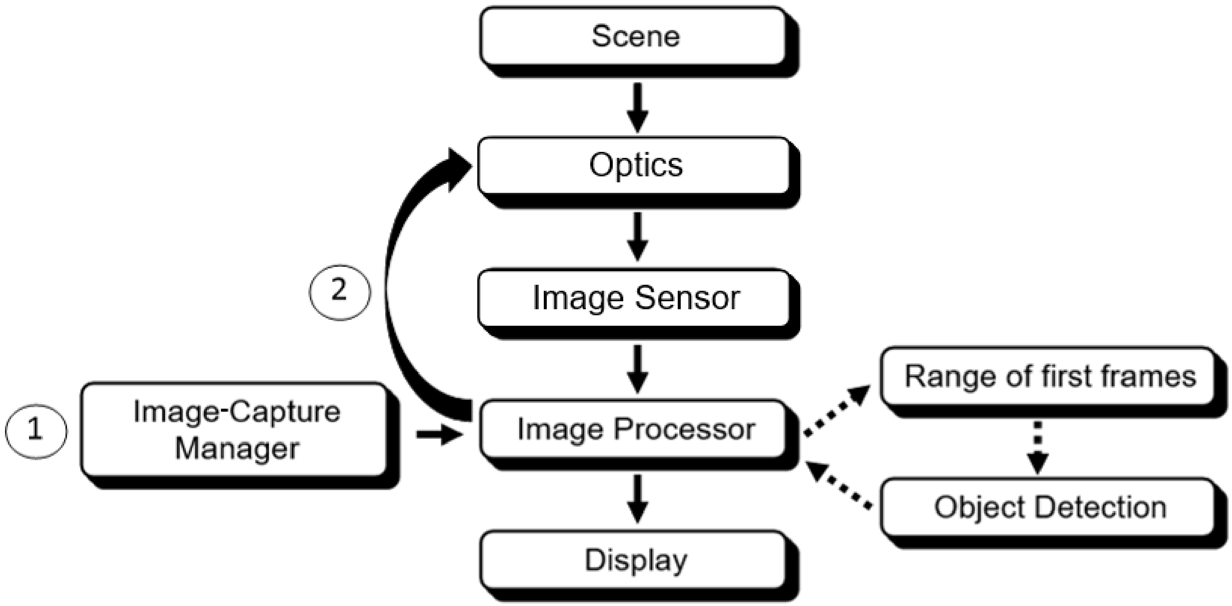

Figure 4

As illustrated in Figure 4, the imaging device will initially capture the scene under pre-set optic settings, and image sensor will produce a RAW first frame. The image processor can then receive the first frame. As illustrated by the number 1, the image capture manager can direct the image processor to generate a set of processed first frames. An object detector then performs object detection on the set of processed first frames. Once an object of interest is identified in one or more frames of the set of processed first frames, the imaging device can determine the optic and image-processing settings to utilize to generate a quality image of the scene. If necessary, the image processor, as illustrated by the number 2, can then adjust the optics, the image sensor can produce a new RAW frame, the image processor can receive the new RAW frame, and the image processor can apply the appropriate image-processing settings to the new frame to generate an image of the scene.

With the imaging device performing operations under the direction of the image capture manager, the initialization process associated with scene changes can also be expedited. For example, if the user decides to capture a different scene, then the image detector can more quickly identify an object of interest in the new scene.

In conclusion, utilizing the techniques and processes described herein, an object of interest may be identified earlier in the initialization process of an imaging device and optimal image-processing settings can be determined and applied sooner.

References:

[1] Guy Cote, Simon Wolfenden Butler, Joseph Anthony Petolino, Jr., and Joseph P. Bratt. Raw Scaler with Chromatic Aberration Correction. US Pub. 20130321675 A1, filed May 31, 2012.

[2] Mihai Ianculescu, Petronel Bigioi, Mihnea Gangea, Stefan Petrescu, Peter Corcoran, and Eran Steinberg. Real-time face tracking in a digital image acquisition device. US Pub. 20080037838 A1, filed May 24, 2007.