View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Technical Disclosure Common

Technical Disclosure Commons

Defensive Publications Series

April 2020

Improved Object Detection in an Image by Correcting Regions with Distortion

Ruiduo Yang

Yichang Shih

Chia-Kai Liang

Sam Hasinoff

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Recommended Citation

Yang, Ruiduo; Shih, Yichang; Liang, Chia-Kai; and Hasinoff, Sam, "Improved Object Detection in an Image by Correcting Regions with Distortion", Technical Disclosure Commons, (April 01, 2020)
https://www.tdcommons.org/dpubs_series/3090

This work is licensed under a Creative Commons Attribution 4.0 License.
This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

Improved Object Detection in an Image by Correcting Regions with Distortion

Abstract:

This publication describes techniques and processes for correcting distortion in an image in order to improve object detection by an object detector on an imaging device. In order to avoid missing target objects (*e.g.*, faces) in an image during object detection due to distortion, the object detector performs object detection on the image using a low-threshold value. A low-threshold value is associated with a lesser chance of missing target objects. The detection results are compared against regions of the image that are known to have distortion due to factors, such as a wide field of view (WFOV) lens of the camera. The overlapping areas of the detection results and the distorted regions are identified as candidate areas that would benefit from distortion correction. Using an algorithm, the candidate regions are corrected (*e.g.*, undistorted, cropped, down-sampled, rotated, and/or frontalized) to reduce distortion. Object detection is performed again over the corrected candidate regions, resulting in improved confidence. The object detection results can then be used by the imaging device to provide a high-quality image and a positive user experience with the imaging device.

Keywords:

Object detection, object distortion, candidate regions, facial recognition, rectification, wide field of view, WFOV, corrected regions, downsample, undistorted, cropped, frontalized, rotation, homography, object detector, threshold, confidence, perspective distortion, lens distortion

Background:

In order to improve the quality of an image taken by an imaging device, such as a digital camera or a smartphone with image-capturing capabilities, the imaging device may utilize an object detector. The object detector performs object detection to identify areas of interest within an image for image-processing settings to be optimized. In an example, the object detector may identify a face in an image and adjust image-processing settings to produce a high-quality image of the identified face within a scene.

However, the object detector may not operate as intended. Using the face example, if the face is located on a perimeter of an image, there may be perspective distortion or lens distortion that causes the object detector not to identify the face. As a specific example, a camera with a wide field of view (WFOV) lens may have perspective distortions in the corners of an image that are captured that cause the object detector to not identify the face. The object detector is usually trained on large amounts of data that may not include sufficient data from WFOV lenses. Accordingly, a warped face at the corner of the image may not be easily identified.

Therefore, it is desirable to identify and correct distorted regions in an image prior to performing object detection to optimize the detection and recognition of target objects.

Description:

This publication describes techniques to improve object detection by an object detector by identifying distorted objects in an image and performing image processing to correct the distorted objects. After correcting the one or more distorted objects, when the object detector performs object detection over the corrected objects, the object detector is more likely to correctly identify all target objects in the image. Figure 1 illustrates examples of imaging devices that may utilize object detection technology and components that may be present in such an imaging device.

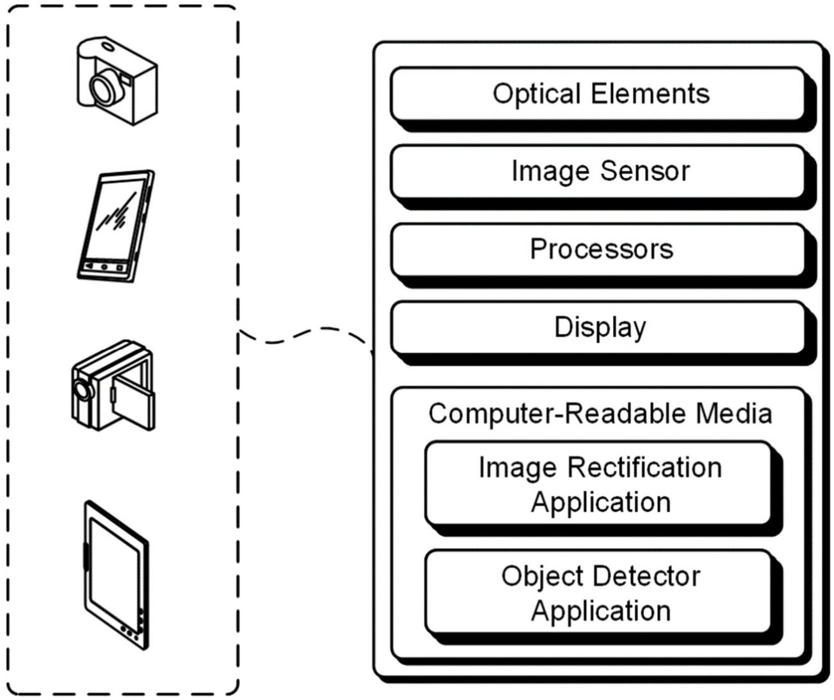

Figure 1

An imaging device can capture an image of a scene. In the example of Figure 1, 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 capture 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 may be 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, and/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, and/or activities associated with post-processing 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). An 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, Auto-White Balance, Auto-Exposure, and/or Auto Focus. In some cases, an image processor may also operate as an image detector, performing object detection (*i.e.*, an object detector).

The imaging device also includes executable instructions of an image rectification application and an object detector application. The image rectification application and the object detector application 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 rectification application to identify distorted regions in an image and perform rectification (*i.e.*, distortion correction) over the regions. The imaging device performs operations under the direction of the object detector application to 1) identify areas of interest in an image and 2) use the detection results so that the appropriate image-processing settings can be optimized for those areas. The image rectification application works with the object detection application, as detailed below.

The process for optimizing the detection and recognition of target objects first utilizes information associated with the lens that captured the image to identify regions of the image where distortion is likely. For example, an image taken with a 35 millimeter (mm) focal length lens has a field of view that is approximately 74 degrees, and there is little distortion. However, an image taken with an 11 mm lens has a field of view close to approximately 180 degrees, and the lens is considered to be a WFOV lens. Distortion is exaggerated for WFOV lenses, with objects close to the camera appearing rotated and much larger than objects farther away. For example, if taking a selfie that includes a background scene (*i.e.*, a face in the corner of the shot and the environment in the background), the WFOV lens may cause the person's face to be distorted by looking elongated and rotated in the corner. This distortion can cause the object detector to not identify objects (*e.g.*, the face). The image rectification application stores information associated with the lens that took an image to identify regions of the image where distortion is likely.

The object detector generates detection results for the image based on a threshold value that is set to be low. With regards to the threshold value, if the object detector performs detection with a low-threshold value, objects may be incorrectly identified as being target objects by the object detector. However, in this case, it is less desirable to miss an object during detection than to over-identify incorrect objects. The detection results identify regions of the image that include a detected object (*e.g.*, a face), as illustrated in Figure 2.

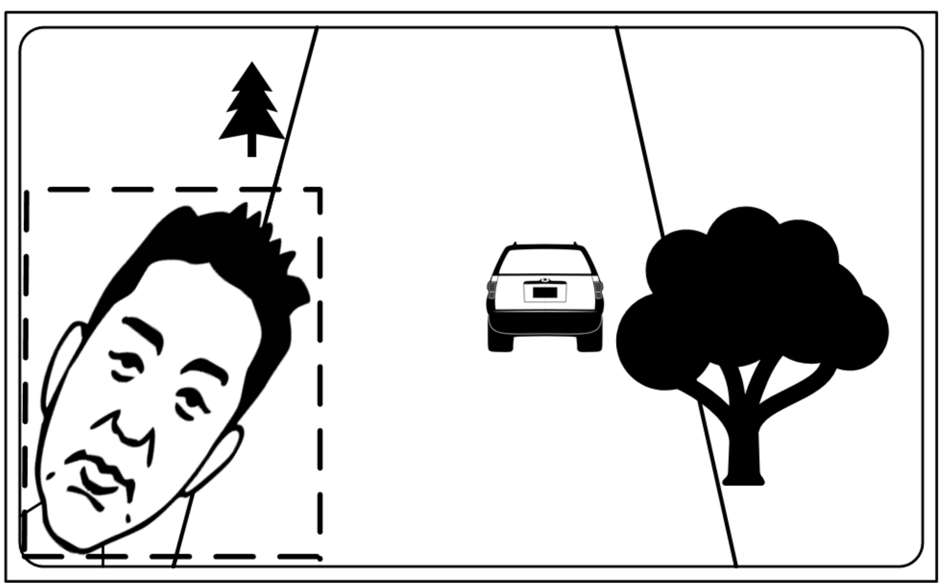

Figure 2

The image rectification application identifies overlaps between the regions where distortion is likely and regions of the image that include a detected object. These overlaps are candidate regions that may benefit from distortion correction. In an alternative embodiment, regions may default to be candidate regions based on the lens used without having detection results.

For example, the image rectification application performs rectification over the candidate region to improve perspective. The image rectification application filters the candidate regions based on factors (*e.g.*, detection confidence, location, size, face orientation, landmark location, etc.). For example, the image rectification application may downsample the candidate region to correct size then un-crop the candidate region. In an embodiment, the image rectification application may use homography to rotate and frontalize the candidate region. Homography relates the pixel coordinates in two images using rotation calculations to correct a rotation warp on an image. Homography allows the candidate region to be viewed from a different angle, which in this case, brings the candidate region to a more frontal plane (*e.g.*, 0-degree angle with face and camera). Rectification to improve perspective over the candidate region from Figure 2 is illustrated in Figure 3.

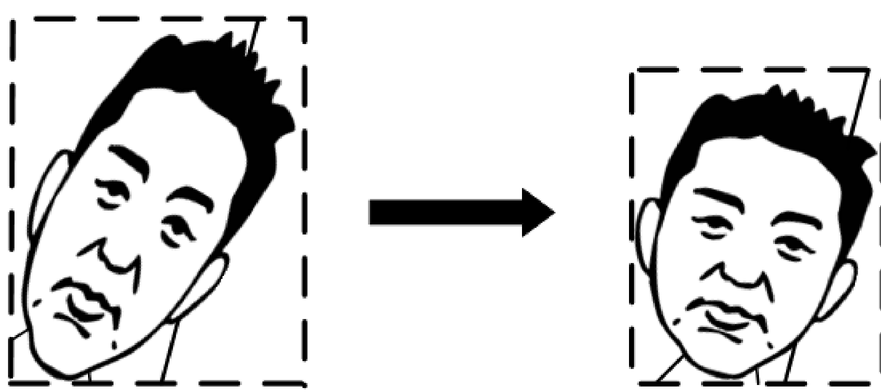

Figure 3

After the candidate regions are corrected, the object detector performs object detection using a higher threshold on the corrected candidate regions. The higher threshold makes it less likely that incorrect objects will be identified as target objects within the corrected candidate region. Alternatively, the object detection may be performed with another object detector that is more accurate.

There is a high level of confidence that the object detection results from the object detector have all target objects identified. If an object is detected over the corrected candidate region, the detection result (*e.g.*, bounding box, landmarks, key points, etc.) will be warped back to the original image space. This allows the imaging device to optimize the image-processing settings (*e.g.*, auto-exposure, auto-white balance, auto focus, etc.) to provide a high-quality image.

References:

[1] Patent Publication: US20110090303A1. Facial Pose Improvement with Perspective Distortion Correction. Priority Date: October 16, 2009.

[2] Patent Publication: US20090207266A1. Image processing device, camera device, image processing method, and program. Priority Date: February 15, 2008.