August 2021

LIGHTWEIGHT SYSTEM FOR TEXT READABILITY ENHANCEMENT OF IMAGES

Amin Barekatain

Anant Nawalgaria

LIGHTWEIGHT SYSTEM FOR TEXT READABILITY ENHANCEMENT OF IMAGES

ABSTRACT

In general, a computing device (e.g., a cellular phone, a smartphone, a desktop computer, a laptop computer, a tablet computer, a portable gaming device, a watch, etc.) may digitize a physical document (e.g., an application, a passport, a contract, etc.) by capturing (e.g., via a camera built into the computing device) an image of the physical document and, in some examples, processing (performing, e.g., optical character recognition (OCR), text extraction, translation, etc.), and storing the image (e.g., in a digital archive). Unfortunately, in some examples, the readability of the digital document may be poor due to noise, unclear text, suboptimal white balance, etc.

To address this issue and others, the computing device may execute an application to process the digital document to increase readability of the digital document without substantially altering the general appearance of the digital document by performing, for example, text readability enhancement, white-balance adjustment, and luminance normalization. The application may be lightweight, relatively fast, and on-device. For example, by processing the digital document fully on the computing device, the application may ensure the privacy of any sensitive information that may be present in the digital document.

DESCRIPTION

FIG. 1 below is a conceptual diagram illustrating a computing device 100 that executes a digitization application 108 (“application 108”) to increase a readability of a digital document. Computing device 100 may be any mobile or non-mobile computing device, such as a cellular phone, a smartphone, a desktop computer, a laptop computer, a tablet computer, a portable gaming device, a portable media player, an e-book reader, a watch (including a so-called smartwatch), a gaming controller, and/or the like. As shown in FIG. 1, computing device 100 may include a presence-sensitive display 102, one or more processors 104, one or more cameras 105, and one or more storage devices 106. Storage devices 106 may include application 108, which in turn may include an image correction module 110 (“IC module 110”), a color space conversion module 112 (“CSC module 112”), a text readability enhancement module 114 (“TRE module 114”), a white-balance adjustment module 116 (“WBA module 116”), and a luminance normalization module 118 (“LN module 118”).

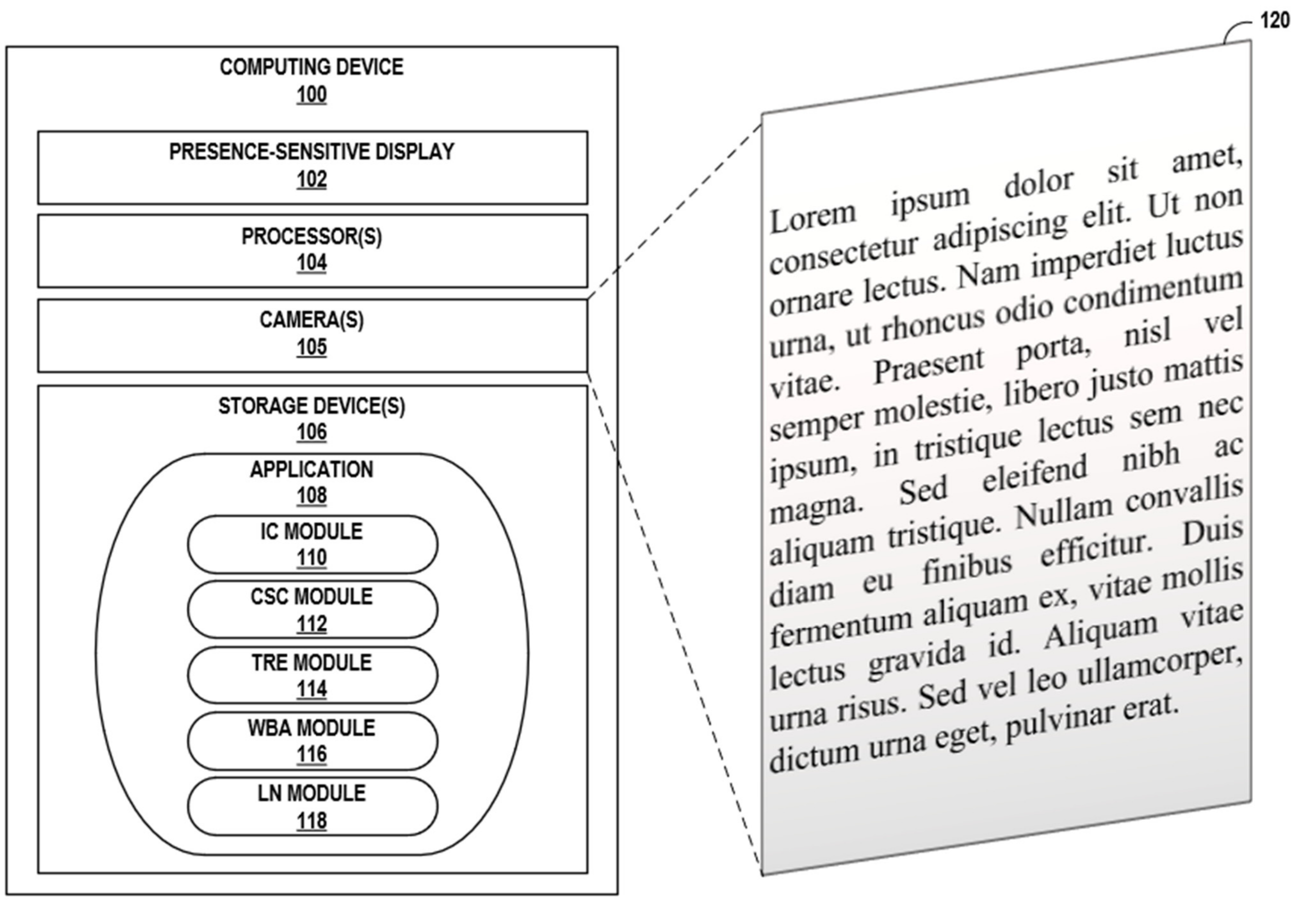

FIG. 1

Presence-sensitive display 102 of computing device 100 may be a presence-sensitive display that functions as an input device and as an output device. For example, presence-sensitive display 102 may function as an input device using a presence-sensitive input component, such as a resistive touchscreen, a surface acoustic wave touchscreen, a capacitive touchscreen, a projective capacitance touchscreen, a pressure sensitive screen, an acoustic pulse recognition touchscreen, or another presence-sensitive display technology. Additionally, presence-sensitive display 102 may function as an output (e.g., display) device using any of one or more display components, such as a liquid crystal display (LCD), dot matrix display, light emitting diode (LED) display, microLED display, organic light-emitting diode (OLED) display, e-ink, active-matrix organic light-emitting diode (AMOLED) display, or similar monochrome or color display capable of outputting visible information to a user of computing device 100.

Processors 104 may implement functionality and/or execute instructions associated with computing device 100. Examples of processors 104 may include one or more of an application specific integrated circuit (ASIC), a field programmable gate array (FPGA), an application processor, a display controller, an auxiliary processor, a central processing unit (CPU), a graphics processing unit (GPU), one or more sensor hubs, and any other hardware configure to function as a processor, a processing unit, or a processing device. Application 108 may be operable by processors 104 to perform various actions, operations, or functions of computing device 100. For example, application 108 may access one or more libraries provided by an operating system of computing device 100 to execute modules 110-118 of application 108.

Storage devices 106 may include one or more computer-readable storage media. For example, storage devices 106 may be configured for long-term, as well as short-term storage of information, such as instructions, data, or other information used by computing device 100. In some examples, storage devices 106 may include non-volatile storage elements. Examples of such non-volatile storage elements include magnetic hard discs, optical discs, solid state discs, and/or the like. In other examples, in place of, or in addition to the non-volatile storage elements, storage devices 106 may include one or more so-called “temporary” memory devices, meaning that a primary purpose of these devices may not be long-term data storage. For example, the devices may comprise volatile memory devices, meaning that the devices may not maintain stored contents when the devices are not receiving power. Examples of volatile memory devices include random-access memories (RAM), dynamic random-access memories (DRAM), static random-access memories (SRAM), etc.

In general, computing device 100 may digitize a physical document (e.g., an application, a passport, a contract, etc.) by capturing (e.g., via camera 105 built into computing device 100) an image 120 of the physical document and, in some examples, processing (performing, e.g., optical character recognition (OCR), text extraction, translation, etc.) and storing image 120 (e.g., in a digital archive, such as storage device 106 of computing device 100). Unfortunately, in some examples, the readability of the digital document may be poor due to noise, unclear text, suboptimal white balance, etc.

Computing device 100 may use conventional techniques to increase the readability of the digital document. For example, computing device 100 may separate the pixels associated with the background (“background”) of the digital document from the pixels associated with the text of the digital document (“text”) and change the color of the background of the digital document to RGB values (255, 255, 255), in this way increasing the contrast between the text and the background of the digital document. However, these conventional techniques may include various disadvantages, such as introducing noticeable artifacts in the digital documents that may make the digital document appear doctored to a user of computing device 100. Further, conventional techniques may be process-heavy, in some cases requiring a relatively long runtime (which may, e.g., harm the user experience of computing device 100) and usage of a remote server (which may, e.g., compromise the privacy of any sensitive information that may be present in the digital document).

In accordance with techniques of this disclosure, application 108 may process the digital document to increase readability of the digital document without substantially altering the general appearance of the digital document by performing, for example, text readability enhancement, white-balance adjustment, and luminance normalization. For instance, application 108 may include IC module 110, CSC module 112, TRE module 114, WBA module 116, and LN module, discussed in greater detail below. Application 108 may be lightweight, relatively fast, and on-device. For example, by processing the digital document fully on computing device 100, application 108 may ensure the privacy of any sensitive information that may be present in the digital document.

As noted above, computing device 100 may use camera 105 to capture image 120 of a physical document though computing device 100 may alternatively receive image data associated with a digital version of the physical document. In any case, to correct for apparent distortion of image 120 caused by camera 105 capturing image 120 of the document at a non-perpendicular angle, IC module 110 of application 108 may perform perspective-correction, keystone-correction, etc. For instance, IC module 110 may rotate, warp, and/or transform image 120 such that the digital document appears to be rectangular.

A color space of the digital document may initially be RGB. For reasons that will be made apparent below, CSC module 112 may convert the color space of the digital document from RGB to LAB using, for example, a predetermined algorithm known in the art. Converting the color space to LAB may help preserve the original colors of the digital document during processing. For example, during processing, application may adjust "L" (lightness) but maintain the "A" (red/green) and "B" (blue/yellow) channels.

TRE module 114 may enhance the readability of text of the digital document by operating on the "L" channel. For example, TRE module 114 may increase the "L" values of a relatively light background and decrease the "L" values of a relatively dark text to increase the contrast between the background and the text. As an example, TRE module 114 may implement global adaptive thresholding and median blur to separate pixels associated with the background of the digital document from pixels associated with the text of the digital document.

Global adaptive thresholding may be an algorithm that divides the digital document into a text portion and a background portion based on a threshold for a "L" value of a pixel and increases the contrast between them. An example algorithm may include decreasing the "L" value of a pixel by a predetermined amount if the "L" value is smaller than the threshold, and increasing "L" value by a predetermined amount if the "L" value is equal to or greater than the threshold. As image 120 captured by camera 105 may have different "L" values in different areas (e.g., due to uneven lighting conditions), the threshold for the "L" value of the pixel may be based on a small region around the pixel. In this way, TRE module 114 may use different thresholds for different regions of the same digital document, which may produce better results for the digital document with varying illumination. In some examples, a user of computing device 100 may control the parameters (e.g., kernel size, binary inversion, etc.) of the global adaptive thresholding.

Applying global adaptive thresholding to the digital document may result in noise (e.g., impulse noise, sometimes referred to as the salt-and-pepper effect) in the digital document. Accordingly, TRE module 114 may apply a median filter to remove noise from the digital document while preserving edges of the text in the digital document. For example, the median filter may replace the “L” value of each pixel with the median “L” value of the neighboring pixels. In some examples, the kernel size may be 3x3 to reduce the impulse noise for relatively small font sizes (e.g., 8 point (pt), 10 pt, 12 pt, etc.).

In another example, TRE module 114 may implement OCR and local Otsu thresholding to separate pixels associated with the background of the digital document from pixels associated with the text of the digital document. For example, TRE module 114 may execute OCR to identify the text in the digital document. TRE module 114 may then define a bounding box (e.g., a relatively small region) that includes the text. Responsive to defining all the bounding boxes, TRE module 114 may change the “L” values of the pixels of the digital document excluding the bounding boxes (e.g., increase the “L” values if the background of the digital document is relatively light). TRE module 114 may then apply local Otsu thresholding to the pixels in the bounding boxes to separate pixels associated with the background from pixels associated with the text.

In comparison to global adaptive thresholding and median blur, OCR and local Otsu thresholding may achieve better performance (e.g., less impulse noise) but be more time-consuming. In some examples, when a user operates application 108, TRE module 114 may perform global adaptive thresholding and median blur to quickly generate a preview of the digital document and perform OCR and local Otsu thresholding to subsequently generate a final version of the digital document.

Since camera 105 may capture image 120 of the physical document under non-standard lighting conditions, the color balance (i.e., the global adjustment of the intensities of the colors) of the digital document may not be neutral. For example, the digital document may have a blue, red, or yellow hue based on the light source and other factors, which may necessitate color correction. Hence, WBA module 116 may adjust the color balance such that the background appears a neutral white, in this way counteracting hues that may be introduced by artificial light, cold light (e.g., a color temperature between 3600° and 6500° Kelvin), etc.

Relatedly, application 108 may execute LN module 118 to increase readability of the digital document irrespective of the lighting conditions when camera 105 captured the digital document. Rather than change the color of the background to white (e.g., RGB values (255, 255, 255)) LN module 118 may modify the "L" values of the pixels associated with the background according to the following equation:

$$N\left(L_{x,y}\right) = L_{x,y} * k$$

where $N(L_{x,y})$ is the normalized luminance value of a pixel, $L_{x,y}$ is the luminance value of the pixel at coordinates x and y prior to luminance normalization, and k is the luminance multiplier. K may be calculated according to the following equation:

$$k = \frac{L_{max} * C}{L_{avg}}$$

where L_{max} is a maximum theoretical value of the "L" channel (e.g., 255 in Java), C is a value between 0 and 1, L_{avg} is the average "L" value of all the pixels associated with the background of the digital document. By adjusting "L" but not the "A" and "B" channels in this way, NL module

118 may increase the readability of the digital document without substantially altering the general appearance of the digital document. CSC 110 module may then convert the color space of the digital document from LAB to RGB to facilitate display of the digital document by presence-sensitive display 104.

One or more advantages of the techniques of this disclosure include reducing the amount of noise, such as impulse noise in digital documents, thereby improving readability. Additionally, the techniques provide a lightweight, fast, on-device application that may enhance the text of a digital document such that the text looks sharp and clear without appearing artificial or unnatural (e.g., due to a white background). Further, as application 108 is on-device and does not send data to a remote server, application 108 may ensure the privacy of any sensitive information that may be present in the digital document.

It is noted that the techniques of this disclosure may be combined with any other suitable technique or combination of techniques. As one example, the techniques of this disclosure may be combined with the techniques described in U.S. Patent Application Publication No. 2019/0354791A1. In another example, the techniques of this disclosure may be combined with the techniques described in U.S. Patent Application Publication No. 2007/0237394A1. In yet another example, the techniques of this disclosure may be combined with the techniques described in U.S. Patent Application Publication No. 2015/0030240A1. In yet another example, the techniques of this disclosure may be combined with the techniques described in CN Patent Application Publication No. 106388781A. In yet another example, the techniques of this disclosure may be combined with the techniques described in U.S. Patent Application Publication No. 2011/0222768A1. In yet another example, the techniques of this disclosure may be combined with the techniques described in Lufick, "Document Scanner - (Made in India) PDF Creator," Google Play, June 2, 2021.