DOCUMENT IMAGE PROCESSOR WITH DEFECT DETECTION

BACKGROUND OF THE INVENTION

The present invention relates generally to the field of digital image processing, and more specifically to the problem of identifying pixel locations with questionable pixel color values which are questionable due to defects in a scanner.

A document is digitized by dividing an image of the document into pixel locations and assigning a pixel color to each pixel location, where the assigned pixel color corresponds to a color of an area around the pixel location on the document being digitized. Typically, the digitizer divides all documents into the same pixel locations, for example, a 300 DPI (dots per inch) digitizer will divide 8½×11" documents into the same 2550 by 3300 pixel locations every time. In some applications, digitizing also includes the process of recognizing patterns of the pixels, such as done in character recognition applications. However, the digitization process of interest herein is just the process of obtaining a pixel array which forms a digital image representative of the document being scanned/digitized.

Defects in the scanning mechanism of a digital copier or computer scanner can introduce errors in the digital images produced by such mechanisms. For example, if the glass platen of a copier is scratched deeply enough that dirt accumulates in the scratch or the scratch affects the refraction of light through the glass, the scratch might appear as a dark line added to the digital image. If the pixel locations associated with these defects can be determined, image restoration techniques can be used to correct the errors by adjusting the pixel color values of pixels at those pixel locations which are flagged as defect locations. Note that a pixel at a defect location need not have an erroneous color value, just a questionable value due to the defect at that location.

A particularly good system for restoration of digital images containing defects at known locations is disclosed in U.S. patent application Ser. No. 08/045,954, filed Apr. 12, 1993, which is assigned to the assignees of the present application and is entitled "Restoration of Images with Undefined Pixel Values."

U.S. Pat. No. 5,214,470, issued to Denber on May 25, 1993, describes a method for isolating defects in which a blank sheet of paper is placed on the scanning platen and the locations of pixels with dark color values in the resulting digital image of the blank sheet are flagged as defect locations. This method is particularly inconvenient and requires the manual intervention of the user. In addition to this method being inconvenient, it may also result in the overinclusion and underinclusion of defect locations in a defect listing.

Overinclusion of defects occurs when there is dirt or dark smudge on the blank sheet, which is then interpreted to be a defect in the scanner itself. Underinclusion occurs where a white defect is present on the scanning platen, such as would be caused by spilled correction fluid adhering to the platen, since white defects cannot be detected from a scan of a white sheet of paper.

Other defect correction mechanisms include manual entry of defect locations, which is even more cumbersome and often must be completed before documents are scanned. Another cumbersome defect correcting process uses image retouching software to allow a user to "brush" away defects in a digital image.

From the above it is seen that an improved system for automatically detecting defect locations in a digital image is needed.

SUMMARY OF THE INVENTION

The present invention solves the aforementioned difficulties with defect detection in scanned digital images. In one embodiment, the invention provides for automatic detection of defect pixel locations from the digital images of an accumulation of documents scanned using the scanner for which defects are to be detected. More specifically, a list of tentative defect locations is kept, and as each document is scanned, entries are added to the list. If only darkening defects are of concern, only the locations of dark pixels are noted in the list. If a subsequent document contains a dark pixel in a tentative defect location present in the list, a count for that tentative defect location is incremented, but if the location contains a light pixel, the count is reduced or the entry for that location is eliminated altogether. If a count for a location is incremented above a threshold, the tentative defect is flagged as a defect and a defect detector records or flags that pixel location as being an actual defect location.

In some variations of the invention, lightening defects are tracked as well as darkening defects. In the specific case of black and white bi-level digitization, a darkening defect always (or in the presence of quantization and alignment noise, almost always) produces a black pixel at the defect location regardless of the color of the document at that location. Similarly, a lightening defect produces a white pixel regardless of the document. The invention also works in the more general case, where pixel colors are any two colors, or where pixel colors are selected from many colors and/or shades of gray. For explanation purposes, pixel colors are often generally referred to as dark and light.

In some embodiments, a memory-saving technique is used to reduce the size of the defect list. Instead of having an entry in the defect list for each tentative defect location of a digital image, the number of entries in the defect list is capped at a specified maximum size and a tentative defect location is added to the list only if an entry is available. To avoid the bias of the defect list being filled with pixel locations from only the first portion of documents scanned, not all defects are put on the list right away. Cursors stored in cursor registers define a rolling window covering less than all the possible defect locations. If a defect is detected, but is not within the rolling window, it is not put on the defect list. However, since the rolling window moves each time a document is scanned, eventually covering the entire document space, a defect will eventually be a candidate for addition to the defect list.

One additional advantage over the prior art of manually scanning a white piece of paper is that intermittent errors can also be detected, with proper adjustment of the thresholds. For example, if a CCD array contains an element which reports, over a series of scans, either a black pixel or a white pixel at random with equal probability, the present invention will detect it, while the prior art only has a fifty percent chance of detecting it, i.e., the black pixel must appear when the white paper is being scanned.

In an alternate method which detects both white and black errors, each page image, or frame, is accumulated until an error map evolves from the accumulation of many images. Because most images for which a copier or scanner are used are mostly white pixels, the page frames are very compressible. For detecting black error pixels, the frames can be made even more compressible by passing sets of frames through