1. Abstract

This study proposes an approach to improve the results of an Optical Character Recognition (OCR) engine by utilizing depth map information from a time of flight sensor in order to correct the warping in a book spread image. This approach does not assume a uniform distribution of the pages of the book spread, but rather that there are varying levels of distortion due to changes in the curvature within each page. The corrections are achieved by modifying the lens equation to take into account different height points on the book spread. In addition, the current resolution limitation from the time of flight device is overcome by scaling and matching the pixel depth data to an image taken by a camera with higher resolution, in order to be able to read the text once the curvature correction is performed. The implementation results support the assertion of improved reading accuracy, which in turn highlights the merits of using this approach based on depth maps in order to correct for the book curvature. Applications of such a system design are in its use as a book reader for persons with visual impairment or as means to digitize books and other bound documents.

2. Introduction

• The purpose of this suggested approach is the correction of a book spread
• The approach relies on being able to obtain an accurate depth map of the book spread
• In this case, the time of flight device was chosen due to its higher processing speed and light weight while maintaining a good level of depth extraction accuracy

3. Book spread correction

• Conceptual Design and Geometry of the Book Reader
  • V is defined as the distance of image sensor from the lens
  • U is the distance of the base of the platform from the lens
  • Point (X,Y,U) is a point along the book spread, with a height of H(X,Y)
  • The corrected location (x, y) on the captured image plane can be expressed as
    \[ x = x' \times \left( \frac{U - h(x',y')}{U} \right) \]
    \[ y = y' \times \left( \frac{U - h(x',y')}{U} \right) \]
  • Further more the length extension (L_i) can be derived from the change in height (dh) and the distance of the pixels along each row (dx)
    \[ L_i = (dh_i - dx_i)^{1/2} \]
  • Then using this length (L_i), the new extended locations can be determined

6. Conclusion

• The results showed, that with the proposed mathematical derivations, the image can be corrected and the warping effect is attenuated.
• The apparent curvature of the book spread is shown to have been flattened and extended correctly and the character recognition rate has been increased.
• Future work will focus on finding alternative approaches to further improve OCR accuracy.

5. Results

• Flattening and linear extension of the page and use of ABBYY FineReader 12

8. Acknowledgments

This research is supported through NSF grants CNS-0959985, CNS-1042341, CNS-1429345, HRD-0833093, IIP 1383922 and IIP-1230661. The support of CAHSI and the Ware Foundation are greatly appreciated.

References