Alignment, Scale and Skew Correction for Optical Mark Recognition Documents Based

Erik Miguel de Elias
Computer Science Department
Technological Institute of Aeronautics
São José dos Campos, Brazil
edeelias@acm.org

Paulo Marcelo Tasinaffo
Computer Science Department
Technological Institute of Aeronautics
São José dos Campos, Brazil
tasinaffo@ita.br

Roberto Hirata Junior
Institute of Mathematics and Statistics
University of São Paulo (USP)
São Paulo, Brazil
rhiratajr@usp.br

***Abstract*—Acquiring an OMR (Optical Mark Recognition) reading equipment can be impracticable to some companies or individuals due to its costs. Computational software solutions can be more attractive, but they require specific page format or page layout, such as specific marks to be used when recognizing any OMR document. In this paper, we propose a way to treat skew, translation, scale and alignment using a base document as reference due to the intrinsic characteristics of the problem. Key points are found by a pattern matching algorithm and used for the document image transformation. The method does not require specific layout, needing less formatting, allowing non-experts to create the form using ordinary software and scanners. Two experiments were executed: one with 40 images distorted randomly from a document clipping, and the second one with 1034 images of real student tests. Both experiments reached high overall accuracy.**

***Index Terms*—OMR, MCQ, assessment, skew, scale, scoring systems, answer sheets.**

I. Introduction

There are tools and computational approaches to working with OMR, instead of specific optical reading equipment. Gaikwad [1] developed an application of OMR using an ordinary scanner. A robust and low-cost automated data entry system to deal with survey research, using an ordinary scanner to input an image of the questionnaire to the system without special OMR machine is proposed by Sanguansat [2]. Zampirolli et al. in [3] presents a solution using simple Morphological operations for segmentation. Perez-Benedito et al. compare AMC and GEXCAT for the development of GEECA [4], being those some of the proposal available in the literature.

A digitized image (acquired by a scanner or equipment with cameras) may present artifacts as distortion, scale, rotation, and translation. These artifacts can prevent or hinder a good result in the next steps of marks recognition and data obtaining process. For a document batch or group with the same layout, the process can be more assertive and precise if the document is equally positioned, that is, with less interference from the problems aforementioned.

Tilt or rotation can be corrected using techniques and algorithms known as ”skew correction.” This correction is frequently based on approaches using written baselines to detect skew angle [5], [6], and on Hough transform, or Projection profile, or Fourier transform, or Radon transform, or Principal Component Analysis (PCA), or Principal Component Analysis with Wavelet transform, or Moments with Wavelet transform. An extensive review on these techniques can be found in [7].

In broader terms, due to its characteristics, skew can present an excellent correction in terms of inclination deformation or in finding an approximate angle. For instance, the Projection Profile method (already mentioned before) presented a 100% success rate with ±0.3º confidence in the studies by Mahajan, Jaidka, and Apoorva [8].

Phase Correlation is one approach that enables to address the four deformations aforementioned, not to mention the vast amount of research and literature available about Fourier Transform. Luke A. D. Hutchison and William A. Barrett propose a Fourier-Mellin based on Fourier Transform approach to register documents using tables [9]. The algorithm presented good results when solving alignment deformation and a reasonable result for the other deformations commonly found in photographs. Still, it is essential to choose a sound reference image, and the results are good when the algorithms are adapted for the specific goal of documents with tables. Besides that, the authors use a specific method to separate foreground from background, to treat the shear, the translational, and other problems discussed by them in the paper.

There is the possibility to treat the distortions by template matching approach. When finding the template, it could be easily repositioned, but it may not be possible in case there are scale related problems. It should also be considered that template image might have been altered by students scribbles and therefore would not be a match.

Considering how the OMR resource was designed, it can be considerably less affected by those deformations. However, for this to occur, specific points must be defined and found to ensure the comprehension and dispositions of the marks and layout as in [10]. Moreover, there are cases where the software itself executes document printing in a way to recognize layout, facilitating data reading as in [11]. This strategy limits the user when using the same software for confectioning and reading, aside from demanding specific training for using this software. Besides that, solutions that use specific marks and design software, lack versatility, because they are heavily tied up to