

# Extracting written lines from cheques

S. DJEZIRI<sup>1</sup>, F. NOUBOUD<sup>1</sup>, R. PLAMONDON<sup>2</sup>

<sup>1</sup>Département de mathématiques et d'informatique, Université du Québec à Trois-Rivières  
C.P. 500 Trois-Rivières (QC) Canada G9A 5H7

<sup>2</sup>Département de génie électrique et de génie informatique, Ecole Polytechnique de Montréal,  
C.P. 6079, Station Centre-ville, Montréal (QC) Canada H3C 3A7  
e-mail: djeziri@uqtr.quebec.ca, nouboud@uqtr.quebec.ca, rejean@vlsi.polymtl.ca

## Abstract

We present a method to extract written lines from complex backgrounds of cheques. The method is based on a local measure which gets positive values for ideal dark written lines, a null value for ideal edges and negative values for white ideal written lines. The distribution of the gray level according to this local measure is contained in a diamond ODUW that can be divided in three important regions : ideal written lines, uniform regions in the image, ideal dark written lines. A synthetic image is used to show that any discriminating function must cross point O of the diamond. A global process searches for a linear discriminating function to separate the region of dark written lines from the uniform region. A first evaluation on 120 cheques with 16 different complex backgrounds gives 92% of correct results. 8% contains remaining background. This method is an improvement of the previous one [15].

## 1. Introduction

Extracting handwritten information correctly from background images is the first step of a more global process including the recognition and interpretation of all items on the cheque image such as the amount, the date, the owner, etc. [1][2]. This first step may have various degrees of complexity depending on the nature of the background image. When the background is uniform, the histogram contains two modes, which can be separated by identifying the minimum between them. Global thresholding is then suitable and it is still used in many character recognition systems. For non uniform backgrounds, as where the document has a graduated illumination, researchers have investigated many other types of thresholding:

adaptive, dynamic and local thresholding [3] [4] [6]. This problem is addressed in many fields of document analysis : as address extraction from postal packages, OCR, and cheques processing. With more complex backgrounds, as in the presence of a picture on a cheque (or on another document), researchers use other techniques such as morphological approaches [5], or making a subtraction directly between a virgin model of a cheque and a real specimen [7][8]. In some special cases, when a certain hypothesis can be made on the background, as the presence of a gaussian noise or a periodic structure, the extraction is generally successful [9][10].

In general cases, when no hypothesis can be made on the background image, the problem can be expressed as follows : what operator can respond only to lines present in the image? Standard edge detectors, will respond to any line present, including handwritten lines but also contours. The problem therefore, is still complicated because we have to eliminate contour lines from the resulting image. This cannot be done without some hypotheses on the background structure.

On the opposite, it's very easy for the human eye to perform this task. The experimental study on psychovisual tests conducted by Burr and al. [11] confirms the idea suggested previously [12] [13], that the detectors of the human visual system take advantage of the even symmetry that characterizes a written line and of the odd symmetry that characterizes a contour line. By written lines, we mean lines which can be handwritten, printed or graphic lines, and by contour lines, we mean the frontier of regions. In a previous work [14] we have presented a simple measure which gets positive values for dark written lines. In another work [15], we used this difference between of written lines and contour

