

种模式；介绍、总结了多个经典二值化算法；利用直线检测结果等多个定量标准来评价二值化算法的好坏；研究了图像扫描的分辨率与二值化处理之间的关系。本文的主要创新点在于：(1)提出了一种用于整票二值化的算法。该算法基于直方图的最大梯度值，具有简单、高效、可靠的特点，能够较好的满足支票自动化识别系统中整票二值化的需要。(2)将直线检测结果用于对二值化算法效果的评价。此种定量的评价方式能够更好的评价、反映二值化结果对后继处理（直线检测）的影响。

关键字：二值化，图像分割，直方图，最大梯度值，支票图像处理

Abstract

Binarization is a common technique used for segmenting image. The essence of binarization is using a threshold of gray level to classify the pixels in an image. As an important image segmenting technique, binarization has great impact on the final recognition result of image recognition system. A image recognition system usually has two subsystems: preprocessing subsystem and pattern recognition subsystem. The preprocessing subsystem has to segment the target and extract characteristics from the target. Binarization is the main method to segment target. No matter how well the recognition algorithm is, image recognition system can not finish its recognition task correctly, if the target is poorly segmented.

Check recognition system uses DIA (Document Image Analysis) and OCR (Optical Character Recognition) techniques to process the check images. Check recognition system has great academical values and commercial values and has being studied by many scholars recently. In the check recognition system, a check firstly scanned by specially designed scanner in order to getting a digitalized image. Then, the critical step of the check image preprocessing is binarization. As a foundation for the further process, binarization extracting the interesting objects from the whole image has huge impact on the check recognition system.

This paper studies some of the binarization algorithms based on the references [9][10][11][12][13][14]. After analyzing the histograms of gray images of 2000 cheques, we found a clue, which can be used for segmenting

image. With the clue, we developed a new binarization algorithm based on the maximum gradient of the histogram, making the target object stand out in the binary image. Comparing several other commonly used binarization algorithms, the algorithm proposed by the paper has been proven in the simulating tests to be more feasible and advanced. The contributions of this paper are: Concluded three different kinds of check image processing patterns. Introduced and studied several binarization algorithms. Introduced the results of line-detecting and some other measurements to evaluate the quality of binarization. Studied the relationship between the scanning resolution and image. The innovative works of this paper are: (1) Proposed a new and effective and reliable algorithm based on the maximum gradient of the histogram of the image for check image binarization. (2) Introduced the results of line-detecting to evaluate the quality of binarization.

**Keywords: binarization, cheque image preprocessing, histogram,
maximum gradient, segmenting image**

西南财经大学

学位论文原创性及知识产权声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。因本学位论文引起的法律结果完全由本人承担。

本学位论文成果归西南财经大学所有。

特此声明

学位申请人：罗彬杰

2007年12月28日

1. 绪论

1.1 研究背景

当前，图像模式识别系统已经越来越多的走进了人们的生产、生活领域，例如：邮政编码自动识别系统、指纹识别系统、人脸识别系统、车牌识别系统等。而在我国当前的金融业务中，由于多种原因，票据（支票、汇票、表格、内部交换票据等）处理长期以来不得不依赖人工手段（人工分散处理、纸质库房保存、人工查询）。随着中国国民经济持续快速增长，金融行业的业务量也随之成倍增加。人工处理票据的效率有限，单纯依靠增加人手的办法，无疑会加大银行的人力成本。因此银行票据的自动化处理需求，应运而生。由于支票识别系统自身具有较高的学术价值以及商业价值，已经越来越多地受到人们的重视。

国内有多个学术团体在进行支票识别系统的研究，这些学术团体主要包括：南京理工大学、北京邮电大学模式识别与智能系统实验室等。南京理工大学在中文期刊上发表了多篇有关支票识别系统方面的文章^{[1][2][3][4][5]}。北京邮电大学模式识别与智能系统实验室则完成了国家十五期间的一个名为“银行票据 OCR 系统中的成套关键技术”的 863 项目。从公开文献来看，以上两个团体对支票识别的研究，都是建立在打码后的支票上的。在国内，进入商业实用阶段的支票自动识别系统，目前尚无公开文献报道。

从文献[6]可以看出，在国外，一个名为：A2iA Bank Check Recognition System 的支票自动化处理系统已经进入实用阶段。该系统由多个子系统组成，能够处理银行内部的大部分票据。这些票据包括：支票，存款凭条，借记卡、

信用卡申请表单, 银行内部票据。该系统的支票识别子系统从 1996 年开始, 在法国的多家银行和金融机构内开始运行, 处理结果平均可接受率达到了 60-70%, 错误率为 1% (接近人工处理的错误率)。由文献[7]可知, 研究人员在 A2iA 系统的基础上, 进行了适应中国汉字的本地化研究, 并且已经取得了一些成果。关于 A2iA 的更多技术细节介绍可以参看文献[8]。

一个完整的支票识别系统一般由图像预处理子系统和模式识别子系统所构成。而二值化技术作为图像预处理子系统中的核心技术之一, 它将直接影响是否能获得正确的待识别目标以及能否正确的识别待识别目标。本文以关于二值化技术的综述性文献[9][10][11][12][13][14]为基础, 对二值化技术及其相关问题进行讨论, 同时将其在支票识别预处理中的应用作一个论述。针对支票自识别系统的具体特点, 经过研究和总结, 本文提出了一种基于直方图最大梯度值的、用于整票二值化的算法。

1.2 二值化技术概述

二值化技术作为图像分割技术中的一种, 本质上是通过阈值(Threshold)的选择将数字图像中的像素点分成两类, 一类用 1 表示, 视觉上通常为白色, 称为前景(Foreground); 另一类用 0 表示, 视觉上通常为黑色, 称为背景(Background); 或者表示的方法正好与以上描述相反。二值化问题必须针对具体问题具体分析, 目前还不存在一种通用的二值化算法, 也正因如此, 二值化问题一直都是图像处理领域的一个经典问题、经典难题。由于其在图像处理领域的重要性和困难性, 从上世纪七十年代起, 就吸引了大批科研人员对其进行持续不断的研究。早期关于二值化的公开文献请参看文献[15][16][17][18][19][20][21][22]。

人们长期的研究也使得二值化技术在很多领域得到了实际应用。例如: 车牌识别系统^[23]、OCR 系统^{[24][25][26]}、汽车导航系统^{[27][28]}、道路跟踪系统^[29]、染色体分析^[30]、细胞图像分析^[31]、银行票据自动化处理系统^[32]。

1.3 支票识别系统概述

现金支票作为一种重要的纸制文档，包含了大量的关键信息：金额、印章、支票号码、银行代码、支票用途等。现金支票经过专用、高速扫描仪扫描，形成一幅支票数字图像。利用该数字图像，辅以现代计算机技术，可以方便地实现支票的存储、检索、传输，极大地简化了过去传统的支票管理、使用方式。支票的自动化处理，结合银行的支票处理业务流程，从技术实现考虑，可以有以下几种实现模式。

模式一：支票的识别、真伪鉴别工作由人来完成，人工录入支票的相关信息（包括扫描后的数字图像）到数据库，便于日后的查询、管理，此种模式的技术实现最简单，应用价值低。

模式二：支票的识别、真伪鉴别工作由人来完成，人工处理支票的同时，使用打码机在支票的下边缘打码（打码支票如图 1 所示）。打码过程由人输入指定信息到打码机，机器完成打码。需要人输入的信息包括：支票号码、行号、帐号、交易码、金额。打码后再扫描。自动处理系统对扫描形成的数字图像进行识别，将识别结果和支票图像自动存入数据库。此种模式，技术实现难度一般，因为目前印刷体数字识别的算法成熟、商用范围广泛、识别率已经逼近 100%。此种模式的应用价值一般，因为人并未从识别、鉴别等工作中解脱出来。

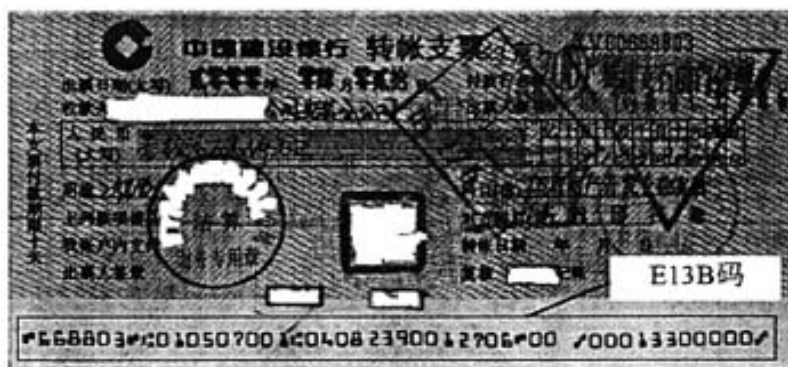


图 1 真实的打码支票

模式三：支票由人工集中后，直接送进高速扫描仪。扫描后形成支票的数字图像。由支票自动处理系统来完成支票相关信息的识别工作，识别结果自动写入指定数据库。同时系统应作出拒识判断。被拒识支票由人工完成识别任务。此种模式的技术实现难度极大，但应用价值较之前两种更大。

1.4 二值化技术在支票识别预处理中的应用

在很多有关图像识别的系统中，二值化通常占据着重要的地位^{[32][9]}。在支票自动化处理系统中，支票经过扫描后，支票图像预处理的第一步是获得一个好的二值化基础，该基础对后继处理（金额栏提取、字符识别等）有着至关重要的影响^[10]。文献[7]提出的支票自动化处理流程如图2所示。

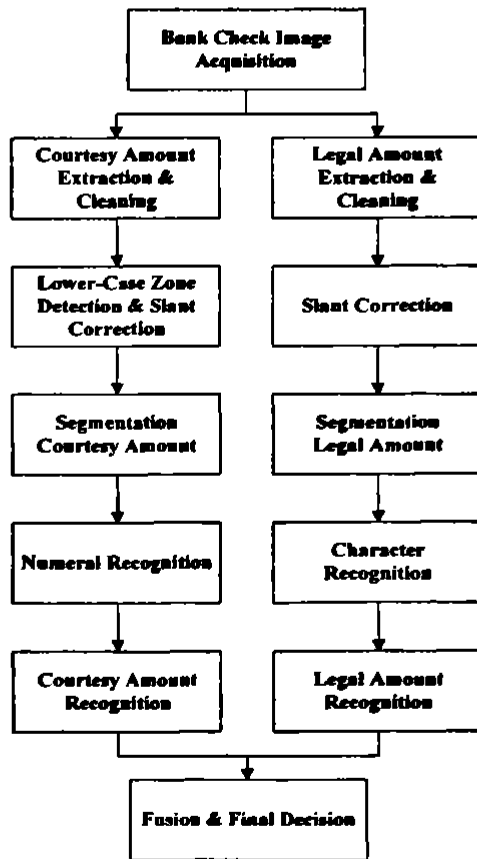


图2 文献[7]提出的支票自动化处理流程图

图 2 所示流程比较概要的介绍了中文支票的处理流程。流程图中虽然没有明确标示出二值化技术在 A2iA 系统中的具体应用。但是文献[6]指出：在流程图的第一步：金额栏(amount field)的获取中不可避免地使用了二值化技术。由文献[7]可以看出，此文实验所用支票为中国香港地区的支票；A2iA 系统所针对的也是欧美等国的现行支票。考虑中外支票的诸多差异，针对我国当前所使用的现金支票，按照第三种支票自动化处理模式，本文提出了一个更加详细，更加适用的支票处理流程(如图 3 所示)。其中的某些模块的存在性、重要性仍然处于探索阶段。该流程图如图 3 所示。其中的实线所示模块为主要模块；虚线所示模块为可选择模块，斜体字所示模块属于图像预处理子系统。

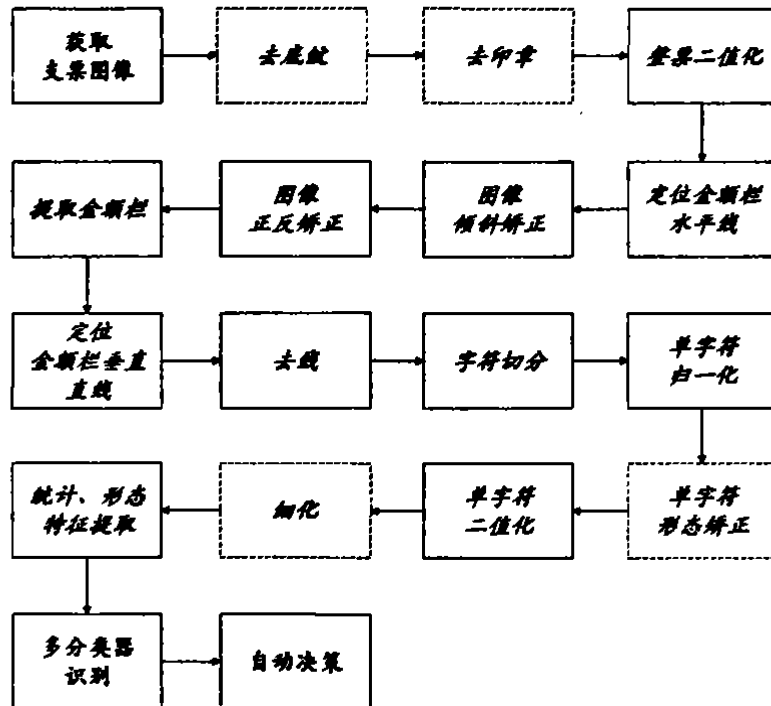


图 3 本文提出的支票自动化处理流程图

在图 3 所示流程中，虚线框所示模块的存在性，目前尚无定论，在公开文献中有一些报道，说明了这些模块的作用。但是，针对我们的具体情况，其存在的意义还需要进一步的研究。从研究的角度来看，某些模块的存在性，只有在系统整体基本完成后，利用后继模块的处理效果或者最终处理效果来证明。本文则主要就个人对“整票二值化”模块的一些研究工作进行论述。

1.5 本章小结

本章主要从三个方面：二值化技术概述、支票识别系统概述以及二值化技术在支票识别预处理中的应用来介绍本文的研究背景。在二值化技术概述中介绍了二值化的基本定义以及二值化算法的发展过程。在支票识别系统概述中总结了基于本国支票的、三种支票识别系统的实现模式。在 1.4 节中提出了更加详细的支票处理流程（图 3 所示），并利用此流程图来说明二值化技术在支票识别预处理中的应用。

2 . 二值化算法

二值化技术作为一种图像分割技术，由于其在工程应用中发挥着重要的作用，长期以来吸引了大量工程技术人员对其进行研究。本章将根据文献 [9][10][11][12][13][14] 为基础，对多种二值化算法及其相关问题进行讨论。本章 2.3 节介绍全局二值化算法，2.4 节介绍局部二值化算法，2.5 节介绍专门用于文本图像的二值化算法，2.6 节介绍两种定量评价二值化算法的准则。

2.1 二值化算法的分类及各类特点

对二值化技术的分类，通常的分类方法将其分为两类：全局二值化算法和局部二值化算法。全局二值化算法，选择单一阈值，该阈值对全局适用。局部二值化算法，则是根据局部信息选择一个阈值，该阈值仅对局部适用。

全局二值化算法，基于对从图像中所得到的某种信息（例如：灰度直方图、笔画宽度统计图）的分析，来获得一个单一阈值 T ，然后根据公式 1，进行二值化。

$$B_{i,j} = \begin{cases} 0, & \text{if } I_{i,j} \leq T \\ 1, & \text{if } I_{i,j} > T \end{cases} \quad \text{公式 1}$$

公式 1 中的 $B_{i,j}$ 表示二值图像 B 中坐标 (i,j) 处的像素值。 $I_{i,j}$ 表示灰度图像 I 中坐标 (i,j) 处的像素值。如果 $I_{i,j}$ 小于等于阈值 T ，那么将此像素点在二

值化结果中应当被赋值为 0, 如果 $I_{i,j}$ 大于阈值 T , 那么将此像素在二值化结果中应当被赋值为 1; 或者正好相反 (将小于等于 T 的赋值为 1, 大于 T 的赋值为 0)。全局二值化算法具有处理速度快的特点, 能够满足普通图像的二值化需要, 但是, 由于使用的是单一阈值, 因此, 不能克服光照不均给图像二值化带来的影响。

局部二值化算法则根据局部信息 (局部像素点灰度平均值、梯度值等), 通过某种准则, 计算获得局部阈值 T , 然后, 对每一个局部按照公式 1 进行二值化。由于对同一张图片, 采用了多阈值分割, 因此, 局部二值化算法能够克服光照不均给图像二值化带来的影响。但是, 相对全局二值化算法而言, 局部二值化算法的处理速度较慢。

而对此常用分类方法, 也有些研究人员持不同意见, 文献[11]就将二值化算法分为了六大类。

- | | |
|------------------|----------------------------------|
| 第一类: 基于直方图形态的。 | 例如: 利用“峰-谷”特性的算法 ^[34] |
| 第二类: 基于聚类方法的。 | 例如: Otsu 算法 ^[20] |
| 第三类: 基于熵的。 | 例如: 文献[35]的算法 |
| 第四类: 基于分割目标物属性的。 | 例如: 文献[36]的算法 |
| 第五类: 基于空间的方法。 | 例如: 文献[22]的算法 |
| 第六类: 局部适应的方法。 | 例如: Bernsen 算法 ^[37] |

因为对于同一个算法从不同的角度来看, 可以划归到不同的类别中, 例如: Otsu 算法^[20]是非参数的算法, 也是基于分类准则的算法。所以, 本文不在算法分类问题上作过多追究, 采用了较为通用的分类方法 (全局、局部两类) 来对二值化算法进行介绍。

2.2 二值化技术发展简介

二值化技术的发展过程可简地要由图 4 所表示。早期出现的二值化算法主要是全局二值化算法, 例如文献[20][34][35]中提出的算法, 随后, 人们注意到局部二值化算法相对全局算法而言, 具有更好的分割能力, 尤其是在光

照不均匀的情况下，局部二值化算法能够完成全局算法所不能完成的任务^{[38][39]}。再后来，人们对二值化算法的研究逐步进入了更加细分的专业领域，利用某一类图像所具有的共性来完成图像分割，文献[40]指出“大多数算法依赖于统计分析，而并未注意文档图像的自身特点”。例如，在文档图像中，可以利用笔画粗细的统计特征完成对文本图像的分割^[39]。

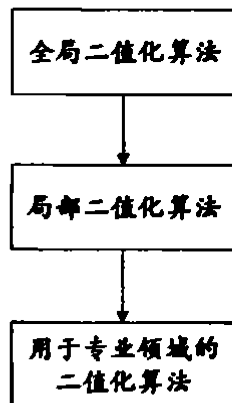


图4 二值化技术发展过程

2.3 全局二值化算法

大多数的全局二值化算法都基于灰度图像的某种统计信息。最为常见的就是利用灰度直方图来选择最佳阈值。灰度直方图是图像的一种统计信息，其中有多种可以利用的信息，例如：直方图的形状、直方图的“峰谷”特性、熵等。关于灰度直方图的图例，可以参考图5。

2.3.1 按比例选择的方法

从本文所搜集的公开文献资料来看，P-tile 算法^[41]应该可以算作最早的一种二值化算法。此算法假设图像由较暗（灰度值较低）的物体和白色背景（灰度值较高）组成；并且还知道目标物体所含像素点的数目所占比例。按照此比例将一定数量的像素点分割出来。此种方法只能用于一些比较特殊的场合，在事先不知道目标物体比例的情况下，无法取得满意效果的。

2.3.2 利用直方图的“谷”的算法

文献[12]将这一类的方法称为“Mode Method”，文献[11]将此类方法归属于基于直方图形态的方法。文献[42]最早提出了利用直方图的“峰-谷”(peak-and-valley)特性来选择图像分割的阈值。该类算法都基于这样一种理想的假设：灰度直方图具有强烈的“双峰”(bimodal)特性，并且两峰的大小差不多，且被一个深“谷”所分开。而两个“峰”之间的“谷”就是将两类物体分割开的最佳阈值。典型的“双峰”特性的灰度直方图如图5所示：

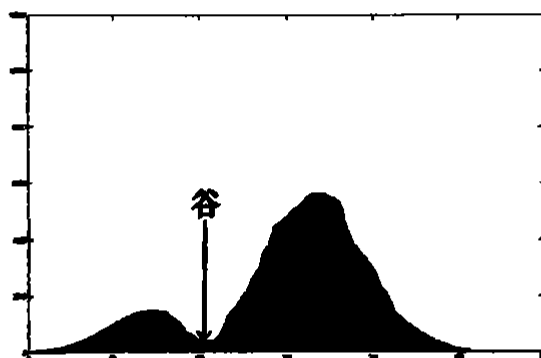


图5 典型的“双峰”灰度直方图

但是，在实际情况下，直方图的形态变化较多。例如，“双峰”之间的“谷”可能会很平坦，两个“峰”的大小差异可能很大，或者呈现出“单峰”(unimodal)、“多峰”(multimodal)形态。针对这些特例，有一些文献提出了解决这些问题的方法。例如，针对“平谷”和两峰差异过大问题，文献[34]提出，在建立直方图的时候，只考虑具有较大（这里需要人工设定一个阈值或者比例来度量“较大”这个标准）“Laplacian”值的像素点，这样建立的直方图将具有更加“深”的“谷”，从而便于利用“谷”来确定最佳阈值。针对直方图中可能出现一些“假”的“峰-谷”，文献[43]利用高斯核函数对直方图进行平滑，然后进行“峰-谷”的定位。文献[43]认为图像中的两类物体具有较低的对比度时，直方图容易呈现出强烈的“单峰”性质。文献[44]还提出，在建立直方图的时候只对具有较低梯度值的像素点进行统计，可以获得更好的具有“峰-谷”特性的直方图。本文认为，如果将此类方法假设为一种通用的二值化算法，那么就应该考虑到，任意一幅图像的直方图中“峰”的个数根本无法预

测，也就是说，我们事先不会有任何先验知识来对“峰”的个数进行判断。此问题也是这类基于“峰-谷”特性的算法所无法克服的（可以应用在某些特殊领域）。

2.3.3 基于凸包的算法

图像中如果含有具有一定区分度的两类物体时，其直方图会体现出一定的“双峰”性质。如果直方图表现出较弱的“双峰”性质时，通过对其直方图进行某种转换、平滑，从而使其具有较强的“双峰”特性是可能的，例如文献[43]中提出的算法。然而转换后的直方图也有可能还是表现为强烈的“单峰”性质，此时，就不应再利用“谷”点来寻找最佳分割阈值。文献[43]就对平滑后仍然呈现强烈“单峰”性质的直方图，提出使用最大曲率来寻找最佳分割阈值。

最大凸残差算法由文献[45]提出。其思想是：当直方图表现为强烈“单峰”性质的时候，可以在直方图的“峰”的“肩”部，找到一个合适的用于分割的阈值。文献[45]认为：如果将直方图视为一个平面区域，不论是利用“谷”来定位最佳阈值，还是利用“肩”来定位最佳阈值；这两种形态在数学上都可以反映为最大凸残差。因此，可以通过计算区域的凸包以及最大凸残差来确定最佳分割阈值。文献[11]仍然将此算法归属于基于直方图形态的方法中。最大凸残差算法的示意图，如图6所示。

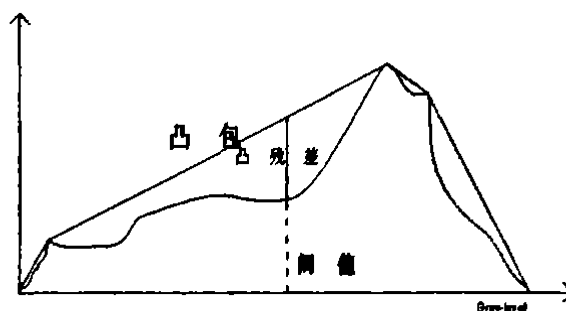


图6 最大凸残差算法的图示

文献[46]对基于凸包的算法进行了一些改进，通过计算直方图的指数凸包(exponential hull)来完成最佳阈值的选择。

2.3.4 Otsu 算法

Otsu 算法^[20]是图像分割算法中的经典算法。在文献[12][13]的测试中，都成为了全局二值化算法中得分最高的算法。但文献[47]也同时指出了 Otsu 算法在前景、背景的像素点数目分布极不对称的情况下会失效的情况。该算法假设图像的灰度总共为 L 级，将灰度直方图在某一阈值处分为两类，第一类 (C_0) 包括灰度直方图中灰度级为: $0, 1, 2, 3 \dots t$ 的像素数目；第二类 (C_1) 包括灰度直方图中灰度级为: $t+1, t+2, t+3, \dots, l-1$ 的像素数目；假设 σ_w^2 , σ_B^2 , σ_T^2 和 n_i 分别表示：类内方差，类间方差，总方差和灰度级 i 所具有的像素数量。令 $\eta = \frac{\sigma_B^2}{\sigma_T^2}$ ，那么最佳阈值 $T = \arg \min(\eta)$ 。其中：

$$n = \sum_{i=0}^{l-1} n_i, \quad p_i = \frac{n_i}{n}, \quad \sigma_T^2 = \sum_{i=0}^{l-1} (i - \mu_T)^2 p_i, \quad \mu_T = \sum_{i=0}^{l-1} i p_i$$

$$\sigma_B^2 = w_0 w_1 (\mu_1 - \mu_0)^2, \quad w_0 = \sum_{i=0}^t p_i, \quad w_1 = 1 - w_0,$$

$$\mu_1 = \frac{\mu_T - \mu_0}{1 - w_0}, \quad \mu_0 = \frac{\mu_0}{w_0}, \quad \mu_i = \sum_{i=0}^l i p_i$$

除了采用 $\eta = \frac{\sigma_B^2}{\sigma_T^2}$ 作为准则来判断最佳阈值。还可以有 $\lambda = \frac{\sigma_B^2}{\sigma_w^2}$ 和 $\kappa = \frac{\sigma_T^2}{\sigma_w^2}$ 两种

准则来选择最佳阈值。

文献[48]对 Otsu 算法作出了评价。文献[49]递归使用 Otsu 算法，从而完成最佳阈值的选择。文献[50]提出了 2-D 的 Otsu 算法。Otsu 算法提出以后，人们又根据其思想，建立了一些新的分类准则，从而完成对像素点的分类任务，例如文献[51][52]中所提出的算法。

这类算法得以实现的一个重要原因，是由于灰度图像本身仅有 256 个灰度级别 (0-255)，利用当代计算机的高速处理能力，可以利用穷举法来找到最佳阈值。文献[53]认为，阈值选择问题是一个组合优化问题，同时提出利用

ACO(Ant Colony Optimization)算法来改进 Otsu 算法, 避免穷举, 提高 Otsu 算法的效率。

2.3.5 基于最小误差的算法

Minimum Error Thresholding 算法^[54]也是一种比较经典的全局算法。该算法在文献[11]的对比测试中取得了多个第一。该算法将灰度直方图视为前景物体与背景物体的混合概率密度函数的估计。假设混合概率密度函数为

$p(g)$, 并 $p(g) = \sum_{i=1}^2 p_i p(g|i)$, 其中 p_i 是先验概率密度,

$p(g|i) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{g-\mu_i}{2\sigma_i^2})$, 那么分割阈值就转换为解下面这个二次方程:

$\frac{g-\mu_1}{\sigma_1^2} + \log_e \sigma_1^2 - 2\log_e p_1 = \frac{g-\mu_2}{\sigma_2^2} + \log_e \sigma_2^2 - 2\log_e p_2$ 。但是, 对于一张数字图

像的灰度直方图而言, 混合概率密度函数 $p(g)$ 的参数 μ_i, σ_i^2 和 p_i ($i=1, 2$) 都是未知的。为了不去计算这些参数的估计值, Kittler 和 Illingworth 提出了一个准则 $J(t)$, 通过计算 $J(t)$ 取得最小值时所对应的 t , 将其作为最佳阈值。

$J(t) = 1 + 2\{p_1(t)\log_e \sigma_1(t) + p_2(t)\log_e \sigma_2(t)\} - 2\{p_1(t)\log_e p_1(t) + p_2(t)\log_e p_2(t)\}$

其中,

$$p_1(t) = \sum_{g=0}^t h(g), \quad p_2(t) = \sum_{g=t+1}^{l-1} h(g),$$

$$\mu_1(t) = \frac{\left\{ \sum_{g=0}^t h(g)g \right\}}{p_1(t)}, \quad \mu_2(t) = \frac{\left\{ \sum_{g=t+1}^{l-1} h(g)g \right\}}{p_2(t)},$$

$$\sigma_1^2(t) = \frac{\left\{ \sum_{g=0}^t [g - \mu_1(t)]^2 h(g) \right\}}{p_1(t)}, \quad \sigma_2^2(t) = \frac{\left\{ \sum_{g=t+1}^{l-1} [g - \mu_2(t)]^2 h(g) \right\}}{p_2(t)}$$

文献[55]证明了: Otsu 算法使得原始灰度图像与二值图像之间的平方误

差达到最小；如果将直方图视为两个具有共同方差的正态分布所形成的混合概率分布，那么 Otsu 算法也就等价于在混合概率模型中找到了条件分布的最大似然值；Minimum Error Thresholding 算法是基于平均最小错分类率的，如果假设直方图为两个具有不同方差的正态分布所形成的混合概率分布，那么 Minimum Error Thresholding 算法也就等价于在混合概率模型中找到了联合分布的最大似然值。如果可以对图像的直方图具有一定的先验知识，那么我们就可以选择合适的算法。当图像由两类物体构成，并且知道这两类物体具有不同的均值、相同的方差，那么，选择 Otsu 算法是比较合适的；如果两类物体具有不同的均值、不同的方差，那么，选择 Minimum Error Thresholding 是比较合适的。本文认为，关键的问题在于我们事先不能对直方图的分布作出正确的判断，这也应当是设计一个二值化算法自动选择系统的关键所在。对于一个“多峰”直方图虽然可以使用类似于文献[49]的方法来完成阈值选择，但是，我们更加需要的是一种能够确定直方图中到底含有几类物体的能力。将有限混合模型中某些算法应用于图像分析，或许可以使具有这种能力。

2.3.6 Moment Preserving 算法

Moment Preserving 算法^[66]希望在二值图像中保留原图像中的矩。假设灰度图像中共有 l ($g=0, 1, 2, \dots, l-1$) 级灰度，对灰度图像中的 i 阶矩 m_i 有如下定义：

$$m_i = \frac{1}{n} \sum_{g=0}^{l-1} g^i h(g), \quad i=1, 2, 3$$

其中 n 是图像中所有像素点的数目。将图像分为两类 j , $j=0$ 表示第一类, $j=1$ 表示第二类，那么对灰度图像中的矩 b_i 定义如下：

$$b_i = \sum_{j=0}^1 p_j z_j^i$$

其中

$$p_0 = \frac{\sum_{g=0}^l h(g)}{n}, \quad p_1 = \frac{\sum_{g=t+1}^{l-1} h(g)}{n}, \quad z_0 = \sum_{g=0}^l g, \quad z_1 = \sum_{g=t+1}^{l-1} g$$

根据二值化图像和原始灰度图像中的一阶矩相等、二阶矩相等、三阶矩相等，可以得到以下方程：

$$b_1 = m_1 \quad b_2 = m_2 \quad b_3 = m_3$$

三个方程，三个未知数，方程有唯一解。通过解以上三个方程可以得到 P_0 ，

先直接将 P_0 的解析表达给出如下：

$$P_0 = \frac{z - m_1}{(c_1^2 - 4c_0)^{1/2}}$$

其中，

$$c_0 = \frac{m_1 m_3 - m_2^2}{m_2 - m_1^2}, \quad c_1 = \frac{m_1 m_2 - m_3}{m_2 - m_1^2}, \quad z = \frac{1}{2} \{ (c_1^2 - 4c_0)^{1/2} - c_1 \}$$

最后再根据下式：

$$P_0 = \frac{1}{n} \sum_{g=0}^l h(g)$$

来确定阈值 t 。

文献[57][58]对此算法作出了一些改进。此算法需要一定的先验知识(将直方图分成几类)，所以，也不具有自动决定分类数目的能力。

2.3.7 基于熵的算法

基于熵的算法，也是全局算法中的一类重要算法。文献[59]对此类方法进行了专题总结和比较。此类算法将直方图视为一个信息源，利用不确定性来描述其中所包含的信息。最早由 Pun^[60]将熵的理论用于图像分割。现对 Pun^[60]的算法描述如下：

如果将 t 作为分割阈值， p_0, p_1, \dots, p_{l-1} 是图像各级灰度的概率。

那么可以获得后验熵 H_0 和 H_1 :

$$H_0 = -\sum_{g=0}^l p_g \log_e p_g \quad , \quad H_1 = -\sum_{g=t+1}^{l-1} p_g \log_e p_g$$

Pun 提出可以通过确定 $H = H_0 + H_1$ 的上限的最大值来找出最佳分割阈值。Pun

证明了, 可以通过确定下试的最大值来确定 H 的上限的最大值:

$$\frac{H_t \log_e p_t}{H_t \log_e \max\{p_0 \dots p_t\}} + [1 - \frac{H_t}{H_T}] \frac{\log_e (1 - p_t)}{H_T \log_e \max\{p_{t+1} \dots p_{l-1}\}}$$

其中

$$H_t = -\sum_{g=0}^t p_g \log_e p_g \quad , \quad H_T = -\sum_{g=0}^{l-1} p_g \log_e p_g \quad , \quad p_t = \sum_{g=0}^t p_g$$

随后 Kapur^[61]对 Pun 的算法作出了改进, Kapur^[61]的算法如下:

令

$$A: \quad \frac{p_1}{p_d}, \frac{p_2}{p_d}, \frac{p_3}{p_d}, \frac{p_4}{p_d}, \frac{p_5}{p_d} \dots \dots \frac{p_6}{p_d}$$

$$B: \quad \frac{p_{s+1}}{1-p_d}, \frac{p_{s+2}}{1-p_d}, \frac{p_{s+3}}{1-p_d} \dots \dots \frac{p_L}{1-p_d}$$

其中,

$$p_d = \sum_{i=1}^s p_i$$

定义:

$$H(t) = -\sum_{i=1}^s \frac{p_i}{p_d} \log\left(\frac{p_i}{p_d}\right) - \sum_{i=s+1}^L \frac{p_i}{1-p_d} \log\left(\frac{p_i}{1-p_d}\right)$$

当 $H(t)$ 取得最大值时, 所对应的 t 作为最佳分割阈值。

2.3.8 基于直方图形态拟合的算法

文献[62]提出了一种基于直方图拟合的算法。文献[62]指出：如果直方图呈现出“双峰”，分割阈值的选择是很容易的。但是，现实中的很多图像的直方图表现各种形态，我们需要一种能够较好处理各种形态直方图的算法。

文献[62]提出使用 *bilevel* 函数来拟合直方图，并且提出了两种用于拟合的误差准则 $E(t)$ 和 $S(t)$ ，当 $E(t)$ 或者 $S(t)$ 取得最小值时，所对应的 t 为最佳分割阈值。

$$E(t) = \sum_{i=0}^t [i - m_1(t)]^2 + \sum_{i=t+1}^L [i - m_2(t)]^2$$

$$S(t) = \frac{1}{N_1 - 1} \sum_{i=0}^t [i - m_1(t)]^2 + \frac{1}{N_2 - 1} \sum_{i=t+1}^L [i - m_2(t)]^2$$

其中：

$$m_1(t) = \frac{\sum_{i=0}^t ih_i}{\sum_{i=0}^t h_i}, \quad m_2(t) = \frac{\sum_{i=t+1}^L ih_i}{\sum_{i=t+1}^L h_i}, \quad N_1 = \sum_{i=0}^t h_i, \quad N_2 = \sum_{i=t+1}^L h_i$$

由于分割阈值 t 的取值范围有限，[62]使用穷举法找出满足条件的 t 。同时，还提出了一种利用先验知识的、定量的评价二值化结果的方法。

该算法并不依赖于某种事先的假设，对直方图形态的变化并不敏感，在处理直方图呈现“单峰”形态时，应该优先考虑此算法。

2.4 局部二值化算法

局部二值化算法利用了局部信息，为每一个局部指定一个适合的阈值，利用该阈值对该局部进行二值化。局部二值化算法相对全局二值化算法而言，最明显的特点并不是利用了局部信息（某些全局二值化算法也会考虑局部信息，例如[51]中的算法），而是多阈值的特点。文献[38]指出：无论如何来选择，单一阈值，都不可能满足光照不均图片的分割需要。从哲学角度来看，

局部二值化算法符合具体情况具体对待的原则。但是，全局二值化算法也可以被应用于局部阈值的选择，例如文献[16]，在对图像分块以后，每块都利用全局二值化算法来选择阈值。

文献[10][12][13]对局部二值化算法进行了总结。根据文献[10][13]的测试结果来看，相对于全局二值化算法，局部二值化算法在处理效果上体现出了明显的优势，但效率相对也有所降低。在文献[12]的对比测试中，由于选择的评判标准都是基于统计量的，因此比较有利于全局二值化算法。从[12]的结果来看，局部二值化算法并未表现出突出优势。

2.4.1 Chow&Kaneko 算法

Chow&Kaneko 算法^[16]是最早的一种局部二值化算法，它同时给出了对直方图“双峰”性质的定量描述，以此描述来判定直方图是否是“双峰”。该算法大致描述如下：

- (1) 将待处理图像分块
- (2) 对每一块进行直方图检测。如果满足某种条件，则选择直方图的“谷”作为此块的分割阈值；如果不满足事先设定的条件，则暂时不对此块设定分割阈值
- (3) 分块图像现在有两种情况：具有阈值和不具有阈值。通过利用已有的阈值对不具有阈值的块进行插值，从而使每块都具有一个阈值。
- (4) 利用每块的阈值对每块进行二值化，最后在拼接成为一张完整的二值化图像。

Chow&Kaneko 算法^[16]利用下式对某块图像计算直方图的平均值 μ 和标准方差 σ ：

$$\mu = \frac{1}{N} \sum F(i)i \quad , \quad \sigma = \sqrt{\frac{1}{N} \sum F(i)(i-\mu)^2}$$

其中 $F(i)$ 表示此块图像中灰度值为 i 的像素点的数目， N 是此块图像中所有像素点的数目总和。当 $\sigma \leq 3$ 时，暂时不对此块图像进行阈值选择。当 $\sigma > 3$ 时，对此块图像是否需要设定阈值，还需要判定其直方图是否具有“双峰”性质，

其判断过程如下:

(1) 对此块图像的直方图进行基于最小方差的高斯拟合。

(2) 拟合结果满足如果某种事先设定的阈值, 则此块图像的直方图具有“双峰”性质, 按照给定的公式来计算出阈值。否则不设定其阈值。

Chow&Kaneko 算法中的基于最小方差的高斯拟合过程如下: 当

$\sum [f(i)-F(i)]^2$ 取得最小值时所对应的参数(P_1 、 σ_1 、 μ_1 、 P_2 、 σ_2 、 μ_2)为

最佳参数, 其中 $f(i) = \frac{P_1}{\sigma_1} e^{-\frac{(i-\mu_1)^2}{2\sigma_1^2}} + \frac{P_2}{\sigma_2} e^{-\frac{(i-\mu_2)^2}{2\sigma_2^2}}$ 。Chow&kaneko 算法使用

hill-climbing 算法来求 $\sum [f(i)-F(i)]^2$ 的最小值, 该算法的三个参数设置如下:

EST(Estimate of the Minimum Function Value)=1.0, EPS(Expected Absolute

Error)=0.1, LIMIT(Maximum Number of Iterations)=30。 $f(i)$ 的参数初始化按照

如下方法完成: 利用 $F'(i) = \frac{F(i-2)+2F(i-1)+3F(i)+2F(i+1)+F(i+2)}{9}$ 对

直方图 $F(i)$ 进行平滑处理。假设图像具有 l 级灰度 ($0 \dots l-1$), 在平滑后的直

方图上选择一个“谷”点, 该“谷”点所对应的灰度值为 ν , 它将直方图 $F(i)$

分成两部分 N_1 和 N_2 :

$$N_1 = \sum_{i=0}^{\nu} F(i), \quad N_2 = \sum_{i=\nu+1}^{l-1} F(i),$$

$$\mu_1 = \frac{\sum_{i=0}^{\nu} F(i)i}{N_1}, \quad \mu_2 = \frac{\sum_{i=\nu+1}^{l-1} F(i)i}{N_2},$$

$$\sigma_1 = \sqrt{\frac{1}{N_1} \sum_{i=0}^{\nu} F(i)(i-\mu_1)^2}, \quad \sigma_2 = \sqrt{\frac{1}{N_2} \sum_{i=\nu+1}^{l-1} F(i)(i-\mu_2)^2}$$

$$P_1 = \frac{N_1 \sigma_1}{\sum_{i=0}^{\nu} e^{-\frac{(i-\mu_1)^2}{2\sigma_1^2}}}, \quad P_2 = \frac{N_2 \sigma_2}{\sum_{i=\nu+1}^{l-1} e^{-\frac{(i-\mu_2)^2}{2\sigma_2^2}}}$$

在完成直方图拟合以后, 可以获得一组最佳参数, 这些参数如果同时满足

$\mu_2 - \mu_1 > 4, 0.1 < \frac{\sigma_1}{\sigma_2} < 1.00, \sigma_{12} < 0.8$, 则认为直方图具有“双峰”。其中

$$\sigma_{12} = \frac{\min_{i \in \{\mu_1, \mu_2\}} f(i)}{\min[f(\mu_1), f(\mu_2)]}$$

如果认定直方图具有“双峰”，那么其阈值 t 可以通过解下面的二次方程来确定：

$$\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right)t^2 + 2\left(\frac{\mu_2}{\sigma_2^2} - \frac{\mu_1}{\sigma_1^2}\right)t + \frac{\mu_1^2}{\sigma_1^2} - \frac{\sigma_2^2}{\sigma_1^2} + 2 \ln \frac{p_2 \sigma_1}{p_1 \sigma_2} = 0$$

假设某分块图像阈值还没有被指定。该块图像的阈值将由下式来确定：

$$T(x, y) = T(x+1, y) + T(x, y+1) + T(x, y-1) + \frac{1}{\sqrt{2}} \{T(x+1, y+1) + T(x+1, y-1) + T(x-1, y+1) + T(x-1, y-1)\}$$

其中 $T(x-1, y-1)$ 为 $T(x, y)$ 所在块的左上方的那块图像的阈值，其它均如图 7 所示，同时还要求 $T(x, y)$ 所在块的、非对角领域块，至少有一个块拥有阈值。

| | | |
|---------------|-------------|---------------|
| $T(x-1, y-1)$ | $T(x, y-1)$ | $T(x+1, y-1)$ |
| $T(x-1, y)$ | $T(x, y)$ | $T(x+1, y)$ |
| $T(x-1, y+1)$ | $T(x, y+1)$ | $T(x+1, y+1)$ |

图 7 Chow&Kaneko 算法示意图

在完成对分块图像的按块插值以后，每一个分块图像都有了一个阈值。然后再对整个图像每块的阈值按块为单位，进行平滑。平滑所用模板如图 8 所示。

| | | |
|----------------------|---|----------------------|
| $\frac{1}{\sqrt{2}}$ | 1 | $\frac{1}{\sqrt{2}}$ |
| 1 | 2 | 1 |
| $\frac{1}{\sqrt{2}}$ | 1 | $\frac{1}{\sqrt{2}}$ |

图 8 Chow&Kaneko 算法平滑模板

最后再利用双线性插值(bilinear interpolation)完成对每个像素点的阈值的确定。假设像素点 P 与四个分块图像 A, B, C, D (其块阈值分别为 T_A, T_B, T_C, T_D)相邻,如图9所示。

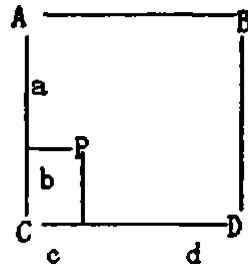


图9 Chow&Kaneko 算法插值示意图

P 点的阈值 t_p 可以由下式计算:

$$t_p = \frac{1}{(a+b)(c+d)}(bdT_A + bcT_B + daT_C + caT_D)$$

然后利用这些阈值对整个图像进行分割。

Chow&Kaneko 算法^[16], 作为最早的局部二值化算法, 对于二值化问题的研究开辟了新的途径。随后 Nakagawa^[15]对 Chow&Kaneko 算法进行了一些改进, 考虑了直方图呈现“三峰”时的阈值选择问题, 并且证明了在某些测试图片上 Chow&Kaneko 算法的分割结果还不如单一阈值的分割结果。同时, 文献[38]也指出 Chow&Kaneko 算法^[16]对图像进行分块的时候并没有考虑图像的内容; 并且在求取局部阈值的时候也只是考虑了局部的统计信息, 并没有充分利用局部均值、梯度等其它信息。由于只是利用了局部的统计信息, 分块后的每块图像如果过小, 就会导致统计信息不足等问题。

2.4.2 Bernsen 算法

Bernsen 算法^[37]假设以 p 点为中心的大小为 $r \times r$ 窗口 Z , 窗口内部的最大值为 g_{\max} , 最小值为 g_{\min} 。利用 $\frac{g_{\min} + g_{\max}}{2}$ 可以计算出一个阈值 T 。令局部对比度 $c = g_{\max} - g_{\min}$, 当 $c < l$ (l 为一事先设定的值)时, Z 为一个整体, 由阈

值 T 来决定整个窗口 Z 是属于背景还是前景; 当 $c \geq l$ 时, 阈值 T 只用来决定 p 点是属于前景还是背景。对图像中的所有像素点进行如上操作, 从而完成图像的二值化。文献[13]将 r 和 l 都设定为 15, 取得了满意的效果。

该算法作为早期的局部二值化算法, 相对于 Chow&Kaneko 算法^[14], 具有思路简单的特点, 因此也经常被其它文献所引用。Bernsen 算法不再像 Chow&Kaneko 算法那样局限于局部的统计信息, 而是着眼于局部图像的对比度和灰度均值信息。

2.4.3 Niblack 算法

Niblack 算法^[43]假设以 p 点为中心的大小为 $r \times r$ 的窗口 Z , 其窗口内部像素点的灰度平均值为 m 、标准差为 S , 那么可以通过 $m + ks$ (k 为一事先设定的值) 计算得到一个阈值用于判断 p 点是属于前景点还是前背景点。文献[13]推荐将 r 设定为 15, k 设定为 -0.2 。

该算法不仅考虑了局部的平均值, 同时将标准差也纳入了阈值计算的范畴。在文献[13]的测试中, Niblack 算法取得了不错的排名。文献[14][40]都对此算法进行了改进。

2.4.4 Yanowitz&Brukstein 算法

Yanowitz&Brukstein 算法^[38]作为最早提出“阈值曲面”(“threshold surface”)的算法, 为局部二值化算法的发展又提供了一个新的发展方向。现在对该算法描述如下:

- (1) 对图像进行平滑。使用模板, 将模板中心点处的像素值替代为整个模板的平均值。
- (2) 在平滑后的图像上, 求取灰度值的梯度值, 形成梯度图像。
- (3) 对梯度图像进行细化, 形成单像素的边界图像, 从而确定局部区域梯度最大值。
- (4) 在平滑后的图像上, 确定局部梯度最大值所对应的像素点, 将这些点作为插值点, 进行一个递归的插值。

(5) 递归插值过程描述如下:

- a. 对阈值曲面进行初始化。边界点的灰度值保持不变, 其它点的灰度值全部赋值为 0
- b. 定义第 n 次插值的留数为 $R_n(x, y) = p_n(x+1, y) + p_n(x-1, y) + p_n(x, y+1) + p_n(x, y-1) - 4p_n(x, y)$ 其中 $p_n(x, y)$ 表示第 n 次插值时, 位置 (x, y) 处像素点的灰度值。
- c. 递归插值公式如下所示: $p_{n+1} = p_n(x, y) + \frac{\beta R_n(x, y)}{4}$, 为了达到快速收敛 Southwell^[66] 建议 β 取值区间为 $(1, 2)$ 。Yanowitz 将 β 设定为 1。递归插值的结束条件是: 某次插值过程所得到的最大留数小于某个事先设定的值。

(6) 插值结束后, 利用插值形成的“阈值曲面”对图像进行二值化。

(7) 对二值化结果进行一个后处理:

- a. 对所有前景物体(像素点的值为 1)按照连通性进行编号。确定每个编号物体的边界像素点, 求这些点所对应的梯度值(第(2)步获得的)的平均值, 如果该平均值小于事先设定的一个值, 那么就将这个物体变为背景(像素点的值为 0)。
- b. 将前景、背景反色后。再对所有的前景物体进行步骤 a 和 b 的处理。

此算法的最大特点是利用边界像素点的灰度值来建立了阈值曲面, 这样就使得分割后的图像与原始图像在内容上保持了一定的一致性。该算法的关键点在于边界像素点的确定。Yanowitz&Brukstein 算法^[38]并未对步骤(1)(2)(3)的具体操作进行详细的描述, 算法实现者应当对这几步有自己的考虑。而且, 步骤(1)(2)(3)的本质其实是确定图像中物体的边界, 算法实现者也可以采用其它的边缘检测算子来替代[38]中的方法。该算法的步骤(7)作为一种后处理过程可以与其它二值化算法搭配使用, 文献[13]证明了此后处理过程与某些算法搭配使用, 对二值化结果有明显改善。文献[69][70]提出了新的建立“阈值曲面”的方法, 同时对 Yanowitz&Brukstein 算法的某些步骤作出了改进。

2.5 专门用于文本图像二值化算法

经过长期的发展，虽然研究人员针对图像二值化技术提出了大量的算法，但是，人们也逐渐发现，没有一种通用的二值化算法或者技术可以应对所有类型图像的二值化处理。因此，人们对二值化算法的研究转向了更加具体的应用方向，例如：二值化技术在处理文本图像中的应用。

2.5.1 Integrated Function 算法

White 和 Rohrer 在文献[64]中最早提出了针对文本图像二值化的专用算法——“Integrated Function”算法。[64]指出：某些局部二值化算法使用了“对比度”这个指标来参与分割阈值的决策，例如 Bemsen 算法^[37]；但是对于某些背景比较复杂的文本图像而言（例如：支票图像）：前景与背景的对比如度低于背景内部某些物体之间的对比度。

针对文本图像的这个特点，文献[64]提出采用：文本宽度、文本图像边缘锐度、对比度，三个指标来共同完成阈值选择。阈值选择过程如下：

- (1) 使用“活动算子”来定位边缘及靠近边缘的像素点。“活动算子” $A(i, j)$

定义如下：

$$d_x(i, j) = u(i-1, j) - u(i+1, j) \quad , \quad d_y(i, j) = u(i, j-1) - u(i, j+1)$$

$$a(i, j) = |d_x(i, j)| + |d_y(i, j)| \quad , \quad A(i, j) = \sum_{n=-1,0,1} \sum_{m=-1,0,1} a(i+n, j+m)$$

其中 $u(i, j)$ 表示坐标 (i, j) 处像素点的灰度值

- (2) 定义一个近似于 Laplacian 算子的算子 d_{xy} 如下：

$$d_{xy}(i, j) = u(i+2, j) + u(i-2, j) + u(i, j+2) + u(i, j-2) - 4u(i, j) \quad \text{根据}$$

$A(i, j)$ 和 $d_{xy}(i, j)$ 可以定义 $s(i, j)$ 如下：

$$s(i, j) = \begin{cases} 0 & \text{当 } A(i, j) < T \\ + & \text{当 } A(i, j) \geq T \text{ 并且 } d_{xy}(i, j) < 0 \\ - & \text{当 } A(i, j) \geq T \text{ 并且 } d_{xy}(i, j) \geq 0 \end{cases}$$

- (3) 通过 $s(i, j)$ 的处理, 图像中的较暗部分中的边缘被标记为“+”较亮部分中的边缘被标记为“-”, 其它的像素点被标记为 0。那么属于笔画内部的像素点应该被标记为 0 或者“+”。
- (4) 在经过 S 函数处理以后的图像上, 使用“Elementary Correlation”技术来完成二值化: 任何属于笔画内部的像素点, 或者被标记为“0”或者被标记为“+”, 并且具有以下特征序列: $\dots, -, +, \dots, [S(i, j) = 0 \text{ 或者 } +], \dots, +, -, \dots$; 其中, 最后一个和第一个省略号表示任何 +、-、0 的排列组合, 中间两个省略号表示 +、0 的任何排列组合。同时还可以利用笔画宽度来进一步改善二值化结果。

该算法作为最早针对文本图像处理而提出的二值化算法, 将笔画宽度条件引入了二值化过程, 启发了后继研究者的思路。文献[71]对“Integrated Function”算法的处理效果进行了测试, 并指出其缺陷, 进而对其进行了一些改进。

2.5.2 Integral Ratio 算法

文献[65]针对字符目标二值化提出了一个专门的算法。此算法是一个分两步定位最佳阈值全局算法。算法的第一步是利用 Quadratic Integral Ration 技术来确定图 10 中的 A 点和 C 点, 从而将直方图分成三类: 前景 (Foreground)、背景 (Background) 和混合 (Fuzzy) 部分。

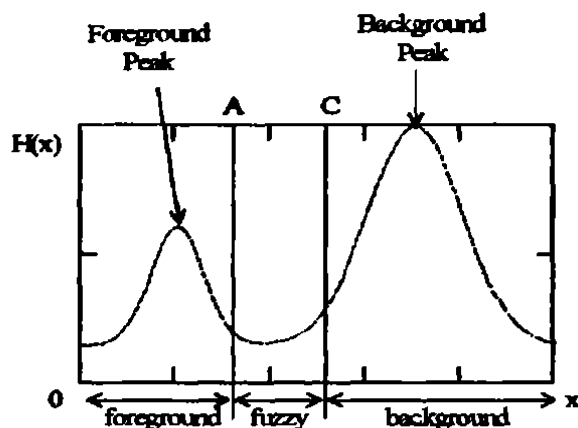


图 10 Integral Ratio 算法示意图

像素点的灰度值如果属于 fuzzy 部分,那么是很难决定其究竟该属于前景还是背景。对于灰度值介于 A、C 部分之间的像素点,需要利用其它的信息来决定其归属。如果我们将分割阈值选择为 C 点所对应的灰度值,虽然可以保留全部的前景点,但是,有些本应属于背景的物体也会被保留下来,这是一种保守的阈值分割策略;如果将分割阈值选择为 A 点所对应的灰度值,虽然所有的背景点都可以被去掉,但是,也会损失部分前景点,这是一种乐观的分割策略。为了追求平衡,最佳分割阈值 t 应属于 $[A,C]$ 。而如何确定 t ,文献[65]根据经验,提出了按以下四种情况来确定:

- (1) 当字符由毡尖笔书写时,最佳分割阈值为 $C - \frac{C-A}{2}$
- (2) 当字符由圆珠笔书写时,最佳分割阈值为 $C - \frac{C-A}{10}$
- (3) 当字符由铅笔书写时,最佳分割阈值为 C
- (4) 当字符由未知类型的笔所书写完成时,最佳分割阈值为 $C - \frac{C-A}{10}$

文献[65]对二值化技术的研究具有一定积极的意义。首先文献[65]提出了,针对文本图像的二值化,需要具体问题具体对待。而不是企图设计一个大而全的算法。其次,文献[65]定性的研究了书写工具,对阈值分割的影响。但是,此算法也对直方图的形态作出了预先的假设(直方图呈现明显的“双峰”),然而在现实情况中,就算是文本图像,也会因为各种原因而不具有明显的“双峰”特性。

2.5.3 Modified Niblack 算法

Sauvola 在[40]中提出了针对文本、图像混排的数字图像，应当分别使用不同的二值化算法。其中，对于文本图像的二值化，Sauvola 提出将 Niblack 的算法^[63]进行改进，从而适用于文本图像的二值化处理。Sauvola 改进后的算法如下：

$$\text{阈值 } T = m \times \left[1 + k \times \left(\frac{s}{R} - 1 \right) \right], \text{ 其中 } R \text{ 是一个动态变化范围、表示窗口内部像}$$

素点灰度值的标准差， k 是一个正数。在 Sauvola 的测试中，他将 R 取值为 128， k 取值为 0.5。通过增加参数来降低算法对噪声的敏感度。

2.6 两种定量评价二值化结果的方法

数字图像处理，长期以来依靠人工对处理结果进行评价，例如文献[13]中所使用的主观评价准则。但是，由于自动化系统的发展需要以及大样本的处理需要，人们针对如何自动、定量评价二值化算法的处理效果，提出了一些评价算法，例如文献[12]将文献[66]的评价方式引入了二值化算法评价。本节所介绍的两种定量的评价准则都是基于一种理想化的假设的：前景点与背景点所属物体具有一致性，一致性从数学上来描述可以反映为方差值小于某个阈值。但是，现实中被划分到前景（背景）中的所有物体还是具有极大的差异的，因此，也并非处处都可以使用量化的准则去评价二值化的效果。就目前为止，本文认为数字图像处理的最主要评价方式还是应当以人眼观察为主，其它评价方式为辅。

2.6.1 Uniformity Measure 评价方法

Uniformity Measure 最早是由文献[66]提出的，文献[12]最早将其用于对二值化结果的定量评价。定义：

$$U(t) = 1 - \frac{\sigma_1^2 + \sigma_2^2}{C}$$

其中 $\sigma_i^2 = \sum_{(x,y) \in R_i} [f(x,y) - \mu_i]^2$, R_i = 分割区域 i , $f(x,y)$ 表示坐标 (x,y) 处像

素点的灰度值, $\mu_i = \frac{\sum_{(x,y) \in R_i} f(x,y)}{A_i}$, A_i 表示分割区域 R_i 中像素点的数目。在

二值化问题中, 分割后的区域一般只有两个, 所以 i 的取值为 1 或者 2。 C 是一个用于归一化的参数。

2.6.2 Shape Measure 评价方法

Shape Measure 评价方式最早被文献[12]用于二值化图像的评价, 其计算步骤如下:

用某种梯度算子计算图像中每一个点的梯度值 $\Delta(x,y)$ 。其中 $\Delta(x,y)$ 由下式计算:

$$\Delta(x,y) = \left[\sum_{k=1}^4 D_k^2 + \sqrt{2} D_1 (D_3 + D_4) - \sqrt{2} D_2 (D_3 - D_4) \right]^{\frac{1}{2}}$$

其中, $D_1 = f(x+1,y) - f(x-1,y)$, $D_2 = f(x,y-1) - f(x,y+1)$

$$D_3 = f(x+1,y+1) - f(x-1,y-1), \quad D_4 = f(x+1,y-1) - f(x-1,y+1)$$

如果某个点的灰度值大于其领域的平均值, 那么将其梯度值赋为正值“+”, 否则将其梯度值赋为负值“-”, 对图像中的所有点的梯度值进行正、负

标记。计算 $S = \frac{\sum_{(x,y)} \text{Sgn}(f(x,y) - \overline{f_{N(x,y)}}) \Delta(x,y) \text{Sgn}(f(x,y) - t)}{C}$, 其中, $\overline{f_{N(x,y)}}$

表示以 (x,y) 为中心的邻域 N 的灰度平均值, t 是最佳分割阈值, C 是一个用

于归一化的参数, 当 $x \geq 0$ 时, $\text{Sgn}(x) = 1$, 否则 $\text{Sgn}(x) = -1$ 。

2.7 本章小结

本章按照图 4 的顺序对二值化算法进行了介绍。着重介绍了几种经典的二值化算法，并且在介绍的同时加入了自己对算法的认识和评论。本章在介绍二值化算法的同时，也介绍了二值化结果的定量评价方式。

在文献[10][12][13][14]等文献中可以看到对多种二值化算法的评价结果，由于采用的标准、以及二值化处理的目标、目的的不同，评测结果也不尽相同。全局二值化算法和局部二值化算法具有各自的优势，全局二值化算法具有速度快特点，但是无法满足光照不均图像的二值化需求；局部二值化算法的处理速度较慢，但是能够较好的克服图像光照不均带来的影响。工程实践者，在面对具体的需求时（处理速度的需求，图像中是否存在光照不均的情况等需求），应当视具体情况的不同、结合前人的经验以及具体的实验结果来选择合适的二值化算法。

3 . 整票二值化

3.1 问题描述

在支票自动化识别应用中，支票上的数字是待识别的对象。作为获取单个字符的必要步骤，我们首先需要做到的是金额栏的提取。根据本文提出的支票自动化处理流程图(图 3)，流程中的“整票二值化”是为金额栏的提取提供一个二值化基础。而金额栏的提取由定位和提取两步构成，其中金额栏的定位则主要依赖于平行直线(图 11 所示)的检测。平行直线检测的效率和准确性对后继的其它处理具有决定性的作用。因此我们就必须在二值化结果中使得我们的目标——平行直线，尽可能的突出出来。

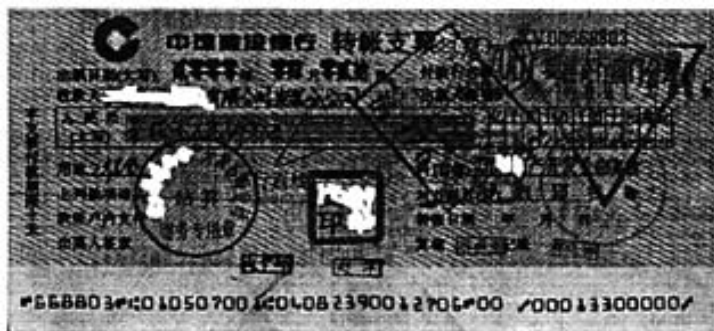


图 11 支票中金额栏的平行直线

在二值化结果中，被检测目标——平行直线应当被视为前景物体，标记为 1；而其它的物体(印章、印刷物等)应当被视为背景物体，标记为 0。在二值化过程中，如果将过多的物体标记为前景物体，那么需要检测的像素点数量就会很多、目标物体不能够突出显示出来，平行直线检测的效率会降低

甚至发生错误：如果将过多的物体标记为背景物体，那么被检测的目标也有可能被部分或者全部标记为 0，那么，平行直线的检测结果可能是错误的或者完全无法检测到目标。因此，有必要寻找一种算法，它能够为每一张支票图像选择一个最佳分割阈值，使得二值化图像的前景物体中含有较多的目标物体的像素同时含有较少的非目标物体的像素。同时，用于整票二值化算法还要考虑效率和效果的平衡性。在文献[13]的测试环境中，根据主观评判准则，局部值化算法的处理效果好于全局二值化算法，但是局部二值化算法所耗时间也要远多于全局二值化算法。综合考虑，本文认为整票二值化所选算法应当是某种全局二值化算法。

3.2 整票二值化效果的评价准则

对于二值化结果的评价，在文献[13]中，作者提出了以下几条主观评价准则：

- (1) 直线断裂次数及缺口大小
- (2) 残缺的符号及文本等物体
- (3) 模糊的直线、符号及文本
- (4) 完全消失的物体
- (5) 单一区域内的噪声情况

而我们认为此五条准则在自身的实际情况中不适用。因为，我们所采用的平行直线检测算法是基于霍夫变换的，该算法具有较好的抗断线和抗干扰的能力；同时，为了能够足够突出目标物体，非目标物体应当尽可能多的被标记为背景，这也与[13]中的评价准则所矛盾。

3.2.1 基准直线的定义

通常来讲，在数字图像中，由人的肉眼所看到的视觉上的一条直线，并非是一条单像素的直线，而是由多条单像素直线所构成的直线(如图 12 所示)，由于霍夫变换最后所定位的直线为单像素直线，因此，我们也有必要在灰度

图像中定义一条单像素直线，作为基准直线，用于处理效果的评价。

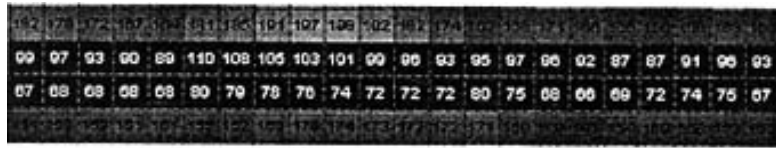


图 12 多条单像素直线构成的直线

我们定义：

$$I = ML + DL \quad \text{公式 2}$$

其中， ML 表示一条单像素直线中，所有像素点的灰度值的平均值； DL 表示一条单像素直线中，所有像素点的灰度值的标准差。在多条备选直线中，具有最小 I 值的直线被确定为基准线。一张支票图像中，金额栏的一组平行直线由多条单像素直线构成，根据公式 2，我们可以为一组平行直线中的每一条线都选择一条基准直线。

二值化结果中，视觉上的一条直线同样是由多条单像素直线所构成的，通过霍夫变换可以定位视觉直线中的一条单像素直线，然后再在这条直线的相邻区域内根据公式 2 来确定基准直线。

3.2.2 利用基准直线评价二值化效果

在确定基准直线以后，我们可以定义

$$R = \frac{PL}{PO} \times 100\% \quad \text{公式 3}$$

其中 PL 为基准直线像素点的数量， PO 为二值化结果中所有前景点的数量。那么由公式 3 可以对目标物在二值化结果中的突出程度作一个定量评价， R 值越大表示，目标物越突出，二值化效果也就越好。同时二值化效果的好坏还可以用提取金额栏的效率和准确率来进一步评价。

3.3 基于直方图最大梯度值的二值化算法

3.3.1 图像统计分析

我们从北京邮电大学获取的支票数据库中选择了 2000 张支票进行了如下统计分析。

(1) 统计分析 1 (直方图“双峰”性质检测)

使用文献[16]提出的检测方法,对 2000 张支票进行了检测,检测数据如表 1 所示。

表 1 “双峰”性质检测结果

| 测试 支票图像数量 | 直方图 具有“双峰”的图像 数量 | 直方图 不具有“双峰”的图 像数量 |
|--------------|------------------------|-------------------------|
| 2000 | 1 | 1999 |

从统计分析 1 的结果 (表 1) 来看,灰度支票图像的直方图绝大部分并不符合“双峰”假设,因此,我们作出推断:使用基于“双峰”假设的二值化算法是不恰当的。

(2) 统计分析 2 (自定义统计分析)

定义 $mh(i)$ 和 $sh(i)$:

$$mh(i) = \frac{\sum_{n=1}^{2000} h_n(i)}{2000} \quad \text{公式 4}$$

$$sh(i) = \sqrt{\frac{\sum_{n=1}^{2000} [h_n(i) - mh(i)]^2}{2000 - 1}} \quad \text{公式 5}$$

其中 $h_n(i)$ 表示第 n 张支票的灰度值为 i 的像素点的数目。

根据公式 4 和公式 5 可以建立如图 13 所示的直方图:

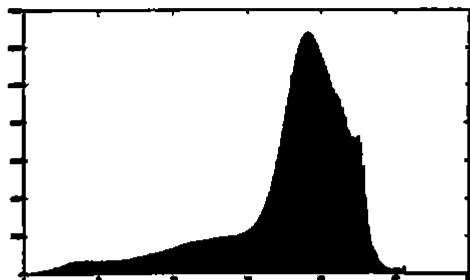


图 13 根据公式 4 建立的直方图

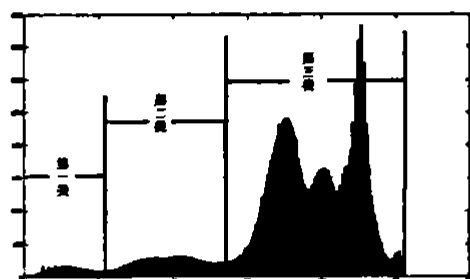


图 14 根据公式 5 建立的直方图

从图 13，很难发现任何可用于图像分割的信息。但是，从图 14，可以看出，该直方图中包含有三类物体（已经标记在图 14 中）。通过对支票中平行直线的像素点的灰度值的抽样中（人工抽样），我们发现，平行直线所属的大部分像素点的灰度值属于第二类，少量像素点的灰度值属于第一类。通过统计分析，发现了隐藏在 2000 张支票中的一种模式。将该模式用于图像分割，提出了基于直方图最大梯度值的二值化算法。

3.3.2 基于直方图最大梯度值的二值化算法

通过以上分析，本文认为灰度支票图像的理想直方图应当如图 15 所示，其最大梯度值已标记于图 15 中。

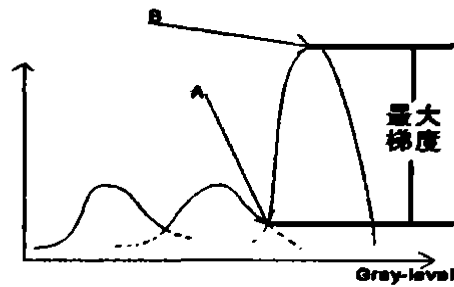


图 15 灰度支票的理想直方图

我们用于确定最佳分割阈值的算法，数学表达如下：

$$rbs(i) = \max_{j=1,2,\dots,ms} [h(i+j) - h(i)] \quad \text{公式 6}$$

$$T^* = \arg \max_{i=0,1,\dots,255-ms} rbs(i) \quad \text{公式 7}$$

其中， T^* 就是我们所选择的最佳阈值， $h(i)$ 表示灰度级 i 所具有的像素数目， $ms = [256/3]$ 是由我们的经验选择的求梯度时的最大步长。

3.4 算法对比测试

3.4.1 客观评价测试

利用北京邮电大学提供的支票图像数据库，我们在其中选择了 2000 张 (200dpi, 256gray-level) 不需要旋转矫正的图像，进行了一系列的面向直线检测的算法评价测试，其基准直线的位置由公式 2 确定。在不同的二值化基础上，进行直线检测所得的精确度，如表 2 所示。

表 2 直线检测精确度

| 直线间距 | Otsu[20] | MSE[62] | HCA[46] | ME[61] | 最大梯度值算法 |
|------|----------|---------|---------|--------|---------|
| 0 | 638 | 584 | 703 | 643 | 643 |
| 1 | 923 | 832 | 891 | 859 | 988 |
| 2 | 325 | 507 | 283 | 396 | 369 |
| 3 | 105 | 70 | 85 | 92 | 0 |
| >=4 | 9 | 7 | 38 | 10 | 0 |

其中“直线间距”表示，在二值化基础上进行直线定位所得位置与主线真实位置间的距离

同时我们还进行了以下一些测试，获得的数据如表 3 所示。

表 3 直线检测综合性能指标

| 测试指标 | Otsu[20] | MSE[62] | HCA[46] | ME[61] | 最大梯度值算法 |
|------|------------|------------|------------|------------|------------|
| AVR | 1.4751% | 2.2542% | 1.3415% | 1.7720% | 2.5920% |
| AVT | 146 | 118 | 154 | 137 | 111 |
| ATB | 0.0776(s) | 0.0746(s) | 0.0756(s) | 0.8320(s) | 0.0866(s) |
| ATL | 21.2098(s) | 15.4320(s) | 23.2027(s) | 17.5550(s) | 12.6138(s) |
| ALG | 0.9620 | 1.0425 | 3.2735 | 0.9835 | 0.8630 |

其中，AVR 是公式 3 中的 R 值的平均值，AVT 是算法所选择阈值的平均值，ATB 是二值化所耗时间的平均值，ATL 是直线检测所耗时间的平均值，ALG 是直线间距的平均值

根据表 2 和表 3 所反映的测试结果来看，本文所提出的最大梯度值算法，为直线检测提供了更好的二值化基础。由表 2 可知，在最大梯度值算法所提供的二值化基础上进行直线检测，所得的精度更高；由表 3 可知，在最大梯度值算法所提供的二值化基础上进行直线检测，所耗时间最少；同时，由表 3 可知，AVR 值能够客观的评价二值化算法的处理结果。

3.4.2 主观评价测试

利用从银行获得的样票，在不同分辨率下扫描后，进行二值化，通过人工来观察、评价二值化效果的好坏。100dpi 情况下的二值化结果，请见附录 1；200dpi 情况下的二值化结果，请见附录 2。通过对二值化处理效果的主观观察，可以得出以下一些结论。

结论 1: 在 100dpi 条件下, 所测试算法的处理结果都不够理想。这是由于所测试的全局二值化算都是基于灰度直方图的。而灰度直方图作为一种统计信息, 由于过低的扫描分辨率, 致使统计量不足, 从而不能选择合适的二值化阈值。同时, 过低的分辨率本身也就会降低图像的质量, 导致视觉效果不佳。

结论 2: 在 200dpi 情况下, 本文提出的二值化算法(详情请见 3.3.2 节)取得了最佳主观效果(图 25): 较好地保留了我们金额栏的平行直线, 同时也较好地减少了干扰像素的存在。MSE 算法^[62]效果最差, 我们所需要的信息几乎被完全去除, 这也直接导致了提取金额栏的失败。

结论 3: 文献[67]指出, 分辨率在 600dpi 时, 能够较好的反映物体的细节。同时考虑效果与性能, 我们认为用于识别处理的支票, 其扫描分辨率最少应该为 200dpi。

3.5 本章小结

在本章中, 通过统计分析(图 14 所示), 我们发现了一些可以用于二值化的信息, 并且利用此信息设计出了一种新的、针对支票图像处理的二值化算法。该算法用于图 3 中的“整票二值化”步骤。通过 3.4 节(算法对比测试), 可以证明, 本文所提出的“基于直方图最大梯度值的二值化算法”能够为后继处理(直线检测)提供一个较好的二值化基础。

4. 总 结

本文以关于二值化技术的综述性文献[9][10][11][12][13][14]为基础,对二值化技术及其相关问题进行讨论,同时,针对支票自识别系统的具体特点,经过研究和总结,提出了一种基于直方图最大梯度值的、用于整票二值化的算法。该算法较好的满足了支票识别系统预处理中整票二值化的需要(表3所反映数据)。同时本文还将直线检测的结果(表2所反映数据)首次运用到了对二值化结果的评价中。最后,本文还总结了支票自动化处理系统实现的三种模式(详见1.3节);利用实验结果展现了图像扫描的分辨率与二值化处理之间的关系(详见3.4.2节)。

根据表2和表3所反映的测试结果来看,本文所提出的最大梯度值算法,为直线检测提供了更好的二值化基础。由表2可知,在最大梯度值算法所提供的二值化基础上进行直线检测,所得的精度更高;由表3可知,在最大梯度值算法所提供的二值化基础上进行直线检测,所耗时间最少;同时,由表3可知,AVR(详见3.2.2节公式3)值能够客观的评价二值化算法的处理结果。

由于时间、精力有限,本文对最新的二值化相关文献跟踪不足,其它研究者可以以本文为基础继续相关问题的研究和跟踪;同时,对于1.4节图3所示流程中的“单字符二值化”也是一个没有很好解决的二值化问题,有待于进一步的研究。

参考文献

- [1] 张重阳, 杨静宇, 张艳. “支票大写金额图像分割策略” 计算机工程, 2006年, 第32卷, 第24期
- [2] 娄震, 陆建峰. “支票图像中打印体字符的分割和识别” 计算机工程, 2003年, 第29卷, 第9期
- [3] 张重阳, 娄震, 杨静宇. “银行支票中小写金额的提取” 中文信息学报, 2003年, 第17卷, 第2期, pp.42-47
- [4] 娄震, 胡钟山, 杨静宇. “支票自动处理系统中的图像处理及识别” 南京理工大学学报, 1999年, 第23卷, 第3期, pp.273-277
- [5] 张重阳, 陈强, 徐勇, 杨静宇. “支票中手写体数字串的提取与分割” 计算机工程, 2005年, 第31卷, 第3期, pp.163-164
- [6] N.Groski, V.Anisimov, E.Augustin, O.Baret, D.Price and J.-C.Simon, “A2iA Check Reader: A Family of Bank Check Recognition Systems,” A2iA, 222 av. du Maine, 75014 Paris, France; <http://www.a2ia.com>
- [7] Hanshen Tang, Emmanuel Augustin, Ching Y.Suen, Olivier Baret, Mohamed Cheriet. “Recognition of Unconstrained Legal Amounts Handwritten on Chinese Bank Checks”. Proceedings of the 17th International Conference on Pattern Recognition
- [8] N. Gorski, V. Anisimov, E. Augustin, O. Baret, and S.Maximov, “Industrial bank check processing: the A2iA CheckReader™,” *Int. Journal on Document Analysis and Recognition (2001) 3*, Springer, Berlin, Germany, 2001, pp. 196-206
- [9] C.-I Chang, Y. Du, J. Wang, S.-M. Guo and P.D. Thouin. “Survey and comparative analysis of entropy and relative entropy thresholding techniques”. *IEE Proc.-Vis. Image Signal Process.*, 2006. Vol. 153, No. 6, pp: 837-850
- [10] ø.D. Trier and A.K. Jain. “Goal-Directed Evaluation of Binarization Methods”. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1995. vol. 17, NO.12, pp: 1191~1201
- [11] M.Sezgin, B.Sankur. “Survey over image thresholding techniques and quantitative performance evaluation”, *J. Electron. Imag.* 2004. 13 (1) 146 -165
- [12] P. K. Sahoo, S. Soltani, A. K. C. Wong, Y. C. Chen. “A survey of thresholding techniques,” *Comput. Vis., Graph., and Image Processing*, 1988. vol. 41, pp: 233-260
- [13] J.Kittler and J.Illigworth, “Minimum error thresholding”, *Pattern Recognition*. 19, 1986, pp:41-47
- [14] Graham Leedham, Chen Yan, Kalyan Takru, Joie Hadi Nata Tan and Li Mian. “Comparison of Some Thresholding Algorithms for Text/Background Segmentation in Difficult Document Images” Proceedings of the Seventh International Conference on

Document Analysis and Recognition (ICDAR 2003)

- [15] Yasuo Nakagawa and Azriel Rosenfeld. "Some experiments on variable thresholding" *Pattern Recognition*, 1979. vol.11, pp:191-204
- [16] C.K. Chow and T.Kaneko. "Automatic detection of the left ventricle from cineangiograms" *Computers and Biomedical Research*, 1972. Vol.5, pp:388-410
- [17] J. S. Weszka and A. Rosenfeld. "Threshold evaluation techniques," *IEEE Trans. Syst. Man Cybern.* 1978. SMC-8, 627-629
- [18] J. Weszka and A. Rosenfeld, "Histogram modification for threshold selection" *IEEE Trans. Syst. Man Cybern.* 1979. SMC-9, 38-52
- [19] T. W. Ridler and S. Calvard, "Picture thresholding using an iterative selection method," *IEEE Trans. Syst. Man Cybern.* 1978. SMC-8, 630-632
- [20] N. Otsu, "A threshold selection method from gray level histogram," *IEEE Trans. Syst. Man Cybern.* 1979. SMC-9, 62-66
- [21] R. L. Kirby and A. Rosenfeld. "A note on the use of (gray level, local average gray level) space as an aid in threshold selection," *IEEE Trans. Syst. Man Cybern.* 1979. SMC-9, 860-864
- [22] N. Ahuja and A.Rosenfeld, "A note on the use of second-order gray-level statistics for threshold selection," *IEEE Trans. Syst. Man Cybern.* 1975. SMC-5, 383-388
- [23] H. A. Hegt, R. J. Haye, and N. A. Khan, "A high performance licenseplate recognition system," *Proc.IEEE Int. Conf. Systems, Man, Cybernetics*, pp:4357-4362,1998
- [24] T. Abak, U. Baris, B. Sankur, "The performance of thresholding algorithms for optical character recognition," *ICDAR'97*, 1997, pp:697-700
- [25] B.Yu, .K.Jain, and M.Mohiuddin, "Address Block Location on Complex Mail Pieces," *Int'lConf. Document Analysis and Recognition*, pp.897-901, Ulm, Germany, 18-20 Aug.1997
- [26] H.Bunke, M.Roth, and E.G.Schukat-Talamazzini, "Off-Line Cursive Handwriting Recognition Using Hidden Markov Model," *Pattern Recognition*, vol.28, no.9, pp:1399-1413,1995
- [27] J.F. Reid and S.W. Searcy. "Vision-based guidance of an agricultural tractor," *IEEE Contr. Syst. Mag.*, 1987. vol.1, pp:39-43
- [28] A.M.Waxman et al. "A visual navigation system for autonomous land vehicles," *IEEE J.Robotics Automat.* 1987. Vol.RA-3. Pp: 124-141
- [29] M.A.Turk, D.G.Morgenthaler, K.D.Greban, and M.Marra. "Video road-following for the autonomous land vehicle," *IEEE 1987 Int. Conf. Robotics and Automation*, vol.1, pp:273-280
- [30] D.Mason, I.J.Lauder, D.Rutoritz and G.Spowart "Measurement of C-bands in human chromosomes," *Comput. Biol. Med*, 1975. vol.5, pp:179-201
- [31] J.M.S.Prewitt and M.L.Mendelsohn, "The analysis of cell images," *Ann. N.Y.Acad.Sci.* 1966. vol.128, pp:39-43
- [32] Amer Dawoud and Mohamed S.kamel. "Iterative Multimodel Subimage Binarization for

- Handwritten Character Segmentation" IEEE TRANS ON IMAGE PROCESSING, VOL.13, NO.9, 2004
- [33] B. Gatos, I. Pratikakis, and S.J. Perantonis, "Adaptive degraded document image binarization," Pattern Recognition, vol. 39, pp. 317-327, 2006
- [34] J.S.WESZKA, R.N.NAGEL, A.ROSENFELD. "A Threshold Selection Technique" IEEE TRANSACTIONS ON COMPUTERS., 1974, DECEMBER, pp: 1322-1326
- [35] ANDREW K.C. WONG and P.K.SAHOO. "A Gray-Level Threshold Selection Method Based on Maximum Entropy Principle" IEEE TRANSACTION ON SYSTEM, MAN, AND CYBERNETICS. VOL.19, NO.4, 1989, pp:866-871
- [36] Ying Liu and Sargur N. Srihari. "Document Image Binarization Based on Texture Features". IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 19, NO. 5, MAY 1997.,pp: 540-544
- [37] J. Bernsen. 1986. "Dynamic thresholding of gray-level images," Proc Eighth Int'l Conf.on Pattern Recognition, Paris, France, pp:1251-1255
- [38] S.D. Yanowitz, A.M. Bruckstein, A new method for image segmentation, CVGIP 46 (1989) 82-95
- [39] Xiangyun Ye, Mohamed Cheriet and Ching Y. Suen. " Stroke-Model-Based Character Extraction from Gray-Level Document Images" IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 10, NO. 8, 2001, pp:1152-1161
- [40] J. Sauvola*, M. PietikaKinen. "Adaptive document image binarization". Pattern Recognition 33 (2000) 225-236
- [41] W. Doyle, Operation useful for similarity-invariant pattern recognition, J. Assoc. Comput. Mach. 9,1962, 259-267.11. G. Fekete, J. O. Eklundh
- [42] J. M. S. Prewitt and M. L. Mendelsohn, The analysis of cell images, in Ann. New York Acad. Sci. Vol128, pp 1035-1053, New York Acad. Sci., New York, 1966
- [43] Du-Ming Tsai. "A fast thresholding selection procedure for multimodal and unimodal histograms". Pattern Recognition Letters 16 (1995) 653-666
- [44] D.P.Panda and A. Rosenfeld, "Image segmentation by pixel classification in(gray level, edge value) space", IEEE Trans. On Computers, Vol. C-27, pp:875-879,1978
- [45] A.Rosenfeld and P. De La Terre, Histogram concavity analysis as an aid in threshold selection, IEEE Trans. Systems Man Cybernet. SMC-13, 1983, 231-235
- [46] R. J. Whatmough, "Automatic threshold selection from a histogram using the exponential hull," Graph. Models Image Process. 53, 592-600, 1991
- [47] J.Kittler and J.Illigworth. "On Threshold Selection Using Clustering Criteria." IEEE Transactions on Systems, Man, Cybernetics, 15:652-655, 1985
- [48] H. Lee and R. H. Park, "Comments on an optimal threshold scheme for image segmentation," IEEE Trans. Syst. Man Cybern. SMC-20,741-742, 1990
- [49] M. Cheriet, J. N. Said, and C. Y. Suen. "A Recursive Thresholding Technique for Image Segmentation". IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 7, NO. 6, 1998, pp:918-921

- [50] J. Z. Liu and W. Q. Li, "The automatic thresholding of gray-level pictures via two-dimensional Otsu method," *Acta Automatica Sin.* 19, 101-105, 1993
- [51] Songcan Chen, Daohong Li. "Image binarization focusing on objects" *Neurocomputing* 69 (2006) 2411-2415
- [52] S.H. Kwon, Threshold selection based on cluster analysis, *Pattern Recognition Lett.* 25 (9) (September 2004) 1045-1050
- [53] Zhiwei Ye, Zhaobao Zheng, Xin Yu, Xiaogang Ning "Automatic threshold selection based on ant colony optimization algorithm" *Neural Networks and Brain, 2005. ICNN&B '05*
- [54] J.Kittler and J.Illigworth, "Minimum error thresholding", *Pattern Recognition.* 19, 1986, pp:41-47
- [55] T. Kurita, N. Otsu and N. Abdelmalek. "Maximum likelihood thresholding based on population mixture models" *Pattern Recognition.* Volume 25, Issue 10, October 1992, Pp: 1231-1240
- [56] W. Tsai. "Moment-preserving thresholding: A new approach", *Comput. Vision Graphics Image Process.* 29, 1985, 377-393
- [57] S. C. Cheng and W. H. Tsai, "A neural network approach of the moment-preserving technique and its application to thresholding," *IEEE Trans. Comput.* C-42, 501-507, 1993
- [58] E. J. Delp and O. R. Mitchell, "Moment-preserving quantization," *IEEE Trans. Commun.* 39, 1549-1558, 1991
- [59] C.-I Chang, Y. Du, J. Wang, S.-M. Guo and P.D. Thouin. "Survey and comparative analysis of entropy and relative entropy thresholding techniques." *IEE Proc.-Vis. Image Signal Process.*, Vol. 153, No. 6, December 2006
- [60] Pun T. 1980. "A new method for gray-level picture thresholding using the entropy of the histogram" [J]. *Signal Processing*, 2: 223- 237
- [61] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, A new method for gray-level picture thresholding using the entropy of the histogram, *Comput. Vision Graphics Image Process.* 29, 1985, 273-285
- [62] N. Ramesh, J.-H. YOO, I.K. Sethi. 1995. "Thresholding based on histogram approximation" *IEE Proc.-Vis. Image Signal Process.*, Vol. 142, No. 5 pp: 271-279
- [63] W. Niblack, *An Introduction to Image Processing*, pp. 115-116, Prentice-Hall, Englewood Cliffs, NJ, 1986
- [64] J.M. White and G.D. Rohrer, "Image Thresholding for Character Image Extraction and Other Applications Requiring Character Image Extraction," *IBM J. Research and Development*, vol. 27, pp. 400-411, 1983
- [65] Y. Solihin and C. G. Leedham, "Integral ratio: A new class of global thresholding techniques for handwriting images," *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-21, 761-768, 1999
- [66] M. D. Levine and A. M. Nazif, Dynamic measurement of computer generated image

-
- segmentations, IEEE Trans. Pattern Anal. Mach. Zntell. P&H-7,1985, 155-164
- [67] Y. Solihin, "A Toolset of Image Processing Algorithms for Forensic Document Examination," Master's thesis, School of Applied Science, Nanyang Technological Univ., Singapore, 1997
- [68] V.R.Southwell, "Relaxation Methods in Theoretical Physics", Oxford University Press, Oxford, 1946
- [69] Ilya Blayvas, Alfred Bruckstein and Ron Kimmel. "Efficient Computation of Adaptive Threshold Surfaces for Image Binarization" Pattern Recognition, Volume 39, Issue 1, January 2006, Pages 89-101
- [70] Dinggang Shen, and Horace H. S. Ip "A Hopfield neural network for adaptive image segmentation: An active surface paradigm" Pattern Recognition Letters Volume 18, Issue 1, January 1997, Pages 37-48
- [71] ø.D. Trier and T.Taxt, "Improvement of "Integrated Function Algorithm"for binarization of document images"

附录

附录 1 (100dpi 时处理结果)



图 16 Otsu 算法^[20]的处理结果

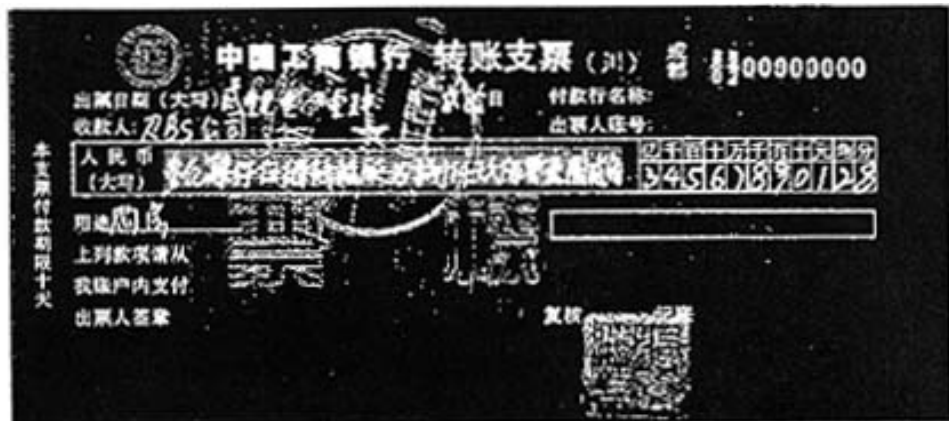


图 17 最大凸残差算法^[45]的处理结果

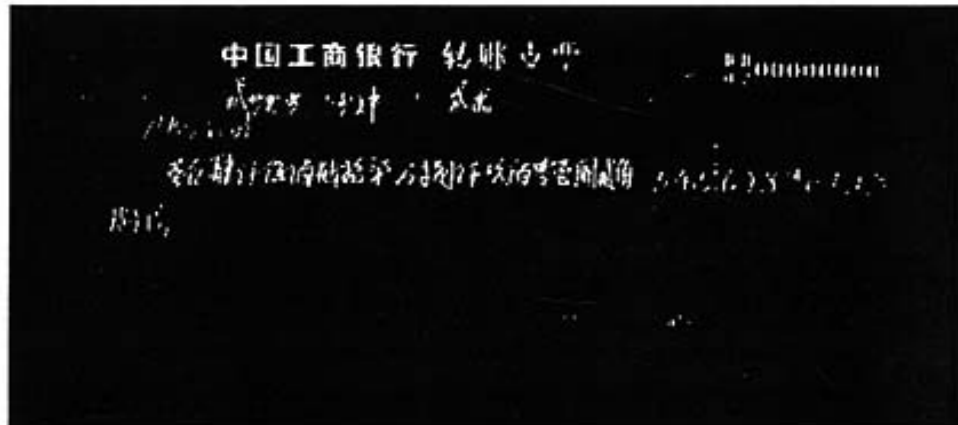


图 18 MSE 算法^[62]的处理结果

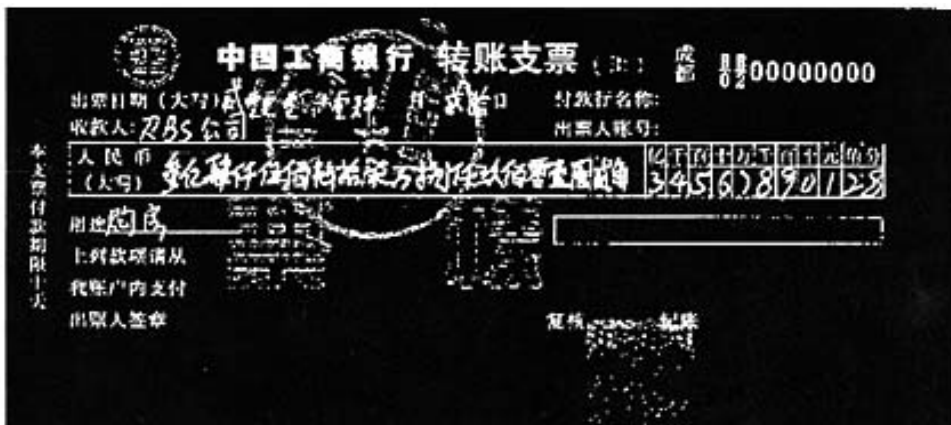


图 19 最大熵算法^[61]的处理结果

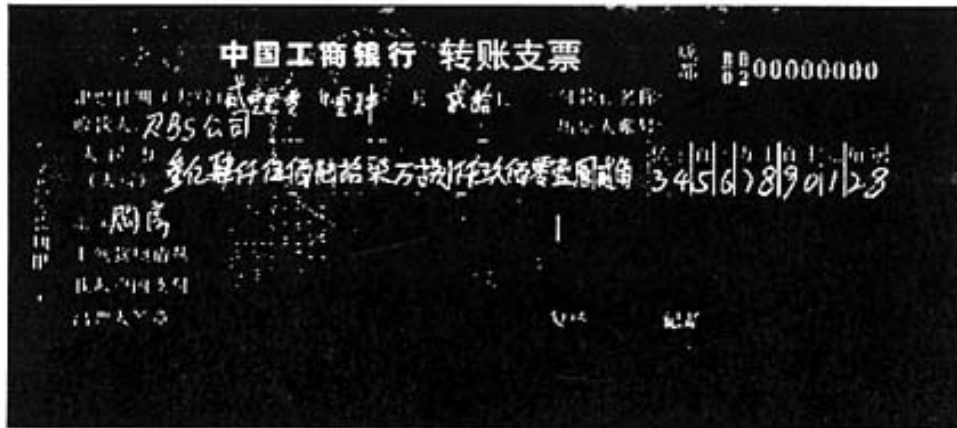


图 20 最大梯度值算法的处理结果

附录 2 (200dpi 时处理结果)



图 21 Otsu 算法^[20]的处理结果



图 22 最大凸残差算法^[45]的处理结果

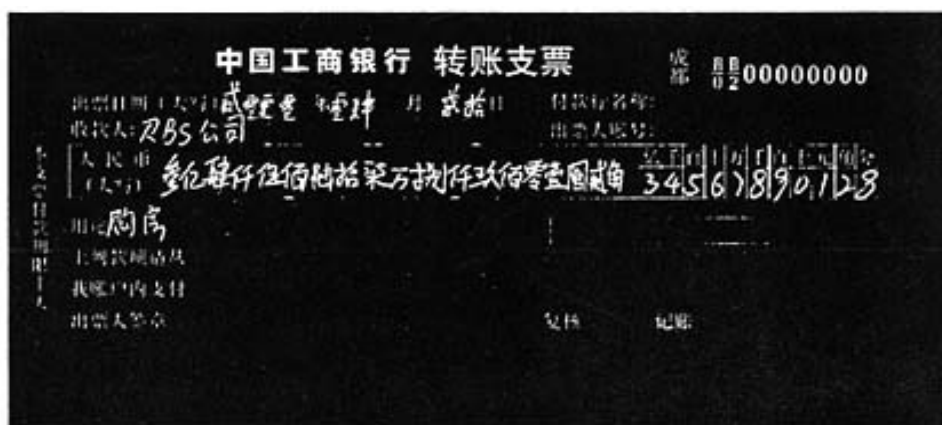


图 23 MSE 算法^[62]的处理结果

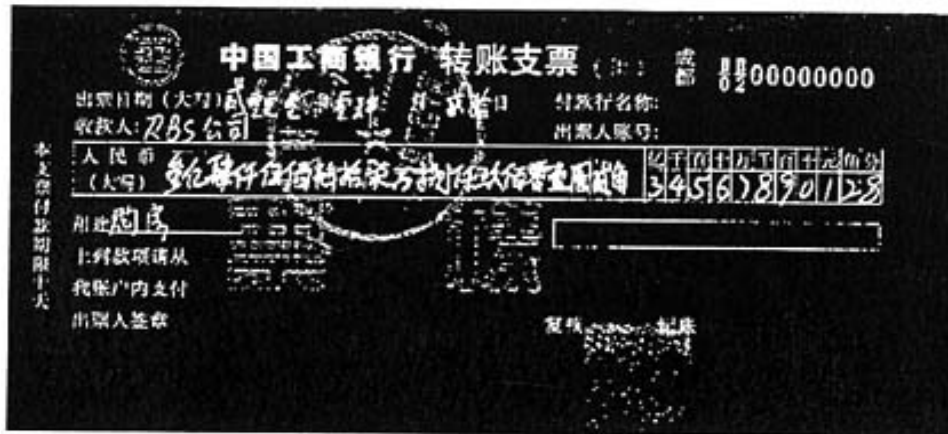


图 24 最大熵算法^[6]的处理结果

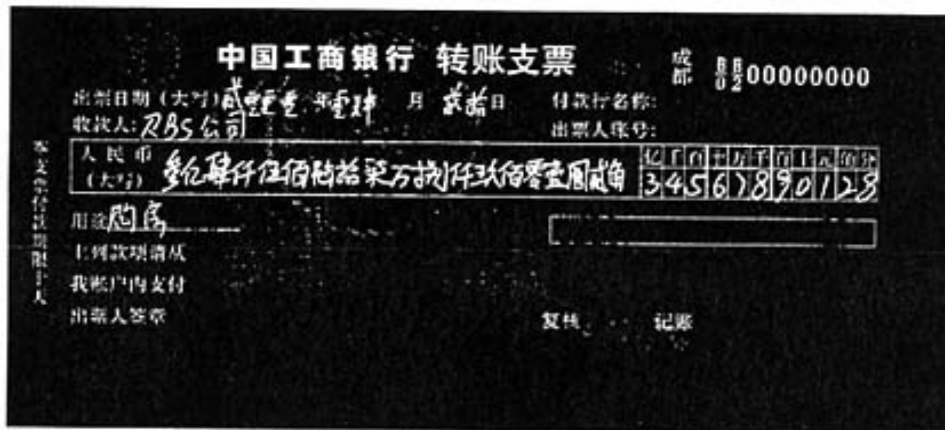


图 25 最大梯度值算法的处理结果

后 记

本文较为全面的分析了二值化技术以及在支票识别中的应用。同时针对具体的应用提出了更加适用的二值化算法，通过较为全面的测试，证明了本文所提出算法的可行性及优越性。二值化技术作为一种重要的图像分割技术，在发达国家已经对其进行了相当多的研究和技术积累，例如：美国 NASA 在 1993 年对二值化技术做了一个总结性的报告。希望本文的总结工作能够为相关技术人员提供一定的帮助。

由于本人能力、时间有限，本文对最新二值化技术文献的跟踪做的不好，但愿后来的研究人员可以在本文的基础上，再进一步的发展。本文对多种二值化算法都进行了准确的描述，完全能够满足研究人员算法实现的需要。本文引用了大量的文献，希望本文可以作为一本二值化技术及其应用的查询目录，帮助研究人员向更加深入的方向发展。

数字图像作为一个离散空间，维数有限；且离散整数能够进行的数学运算也是有限的，二值化算法的研究者通过引入各种数学量、数学模型（均值、方差、熵、滴水模型等）来完成对像素点的分类任务。但是，本文完成的，对大批量的同类图像作统计分析，尚数首次。通过利用统计分析所表现出来的分类信息，从而完成对支票图像的首次二值化（整票二值化）。

致 谢

本人希望在此通过文字向我的导师——贾华丁教授，表示诚挚的谢意，在人生成长的关键时期，是您给予了我极大的帮助与引导，使我踏上了学术研究的道路。经济信息工程学院在研究生阶段也给予了我很多关照与帮助，很多老师都在学术方面给予了我指引，谢谢学院、谢谢老师。同时还要感谢：陈伯承、代星科、何志、何焯、李嘉禧、唐博、唐万丽、邢阳，在文献查找和实验设计方面给予的帮助。最后要感谢我的父母及家人，感谢父母的养育之恩，感谢家人对我的关心和无私奉献。谢谢大家！

在读期间科研成果目录

| 在读期间科研成果目录 | | | | |
|------------|---|-----------------------|------|---|
| 序号 | 题 目 | 刊物或出版社 | 排名情况 | 备注 |
| 1 | A new binarization algorithm based on maximum gradient of histogram | IEEE Computer Society | 第二作者 | The fourth international conference on image and graphics, Chengdu, China, 2007 |
| 2 | 基于磁盘下层过滤驱动的 USB 存储设备安全控制技术 | 计算机应用与研究 | 第三作者 | 2008 年 4 月 出版 |
| 3 | 基于 HSC 的进程隐藏检测技术 | 计算机应用与研究 | 第三作者 | 2008 年 9 月 出版 |