表格图像解析
及其评价方法研究

梁艳慧　计算机应用技术
导师姓名：王亦洲

摘要

随着数字化文本信息在当今时代的广泛普及和应用，文本图像结构和内容的智能分析逐渐成为图像处理和计算机视觉领域的研究热点。表格，作为数字文本图像中经常出现的一种信息表示形式，其结构分析、内容提取和编辑及解析结果的合理评价等具有重要的研究和实用价值。

本文在对表格图像结构进行充分调研的基础上，基于所提出的表格物理结构模型和逻辑语意模型对表格图像进行解析，实现了一个基于友好人机交互界面的表格图像标注工具。同时，基于标注工具提供的标准表格图像解析结果，提出了一种合理的评价方法来评估任意给定的表格图像的解析结果。

表格的物理结构模型针对表格的物理特征进行建模，包括表格行和列的数目、分布以及表格各组成部分的位置信息；表格的逻辑语意模型则是根据表格的物理结构特征，从表格的行头和列头开始，以树形层次结构对表格进行建模，使用唯一的行列索引值来定位表格体中的数据单元。给定一幅表格图像，基于表格的物理结构模型，首先提取表格图像中的连通成分（字符元），并根据连通成分的尺寸特征和表格图像像素直方图在水平方向和竖直方向的投影轮廓，自动地定位表格图像中行分割线和列分割线的分布，得到一个最初的解析结果；之后通过友好的人机交互界面对自动的分割结果进行调整，最终得到一个标准的表格图像结构解析结果。为了标注表格各组成部分的逻辑语意信息，我们基于表格的逻辑

语意模型，从表格的行头和列头开始，以树形结构存储整个表格，给表格的每个数据单元赋予唯一的行列语意索引值。最后，以整个表格作为根节点，表格的每个行头作为根节点的子节点，其逻辑语意信息作为该子节点的属性值，表格列语意信息作为叶子节点，每个数据单元作为叶子节点的文本内容，形成一个XML文件作为输出结果。

对于任意给定的一个表格图像解析结果，我们的评价方法借鉴编辑距离（Edit Distance）的概念，通过计算它的行和列的分割线与标准的表格图像解析结果的行和列分割线的编辑距离，最终返回一个量化分数来评价当前的表格图像解析结果。此编辑距离定义为从给定解析结果到标准解析结果需要对行和列分割线进行插入、删除等操作代价的总和。为了计算每一个分割线的操作代价，我们以表格图像解析过程中提取出的字符元做为结点，构建出一个带有权重的字符元图。在计算字符原图中的边权值时，考虑字符元尺寸大小的相似度和它们之间的距离等影响因素。由此定义一个分割线的操作代价为此分割线切割过的所有边的权值之和，操作代价的大小则反应了分割线不同的错误程度（或者可谅解程度）。

实验结果表明，本文所提出的表格图像解析方法高效准确，能够提供令人满意的标准图像解析结果；同时，对于不同的表格图像解析结果评价合理，便于表格图像解析结果的后续应用。

关键词：表格图像，结构分析，语意标注，评价方法

Research on Table Image Analysis and Evaluation Method

Yanhui Liang Computer Applied Technology
Directed by Yizhou Wang

Abstract

With the wide spread and various applications of digital documents in the modern era, document image analysis gains more and more attention in image processing and computer vision fields. Table is such a ubiquitous form of communication in digital document image that its structure analysis, content extraction and editing as well as evaluation of its analysis result are of great significance.

In this paper, through a sufficient research on various table formats we propose a table model and implement a table image ground truther by a user-friendly graphical interface. Besides, we propose a novel method to evaluate table segmentation results based on the table image ground truther.

Table model includes table structure model and table functional model. Table structure model implies table's physical features, such as the number of rows and columns, the location distribution of the cut-off lines and the distribution of table's components. Table functional model, based on table structure segmentation, is built by constructing a hierarchical tree for the whole table from the stub, stub head and boxhead respectively. In the functional model, each data cell in the table body is assigned a unique row-column index. Given a table image, based on the table structure model our table image ground truther first extracts connected

components (atoms) then semi-automatically determines the locations and spans of row/column separators according to the size of atoms and projection profiles, under human supervision. After we obtain the structure of the table image, we assign different labels to different table components according to their positions, i.e. the stub, stub head and boxhead, in a hierarchical tree structure and assign each data cell its row-column index. Besides, for the convenience of further applications, we output the table analysis result as an XML file.

We evaluate a given table segmentation result by computing edit distance from its row and column separator assertions relative to ground truth. The edit distance is the sum of all the edit operation costs that correct wrong row and column separators. To compute the cost of each wrong separator, we connect the atoms which are extracted in the ground truthing process into an atom graph with weighed edges. Edge weight takes neighboring connected components' size similarities and distances into consideration. Then each edit operation cost is a function of the sum of the weights of the edges that the separator cuts through. Thus, separator errors incur different costs depending on the severity of the error, where severity roughly corresponds to how forgivable the error would be considered by a human observer.

Experimental results demonstrate the effectiveness of our table image analysis method which is able to provide satisfactory results. Besides, the proposed evaluation method is proved to be not only efficient, but also useful in formalizing the intuitive quality of different segmentations.

Keywords: Table Image, Structure Analysis, Functional Labeling, Evaluation Method