

На правах рукописи

09 16 апреля 99

ЛТБ ОД

6 ИЮН 1999

ЗУЕВ Константин Алексеевич

5
**Система идентификации структуры
печатных документов**

05.13.14 – системы обработки информации и управления

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата технических наук

Москва - 1999

Работа выполнена на кафедре физики Московского государственного университета леса и в компании ABBYY Software

Научный руководитель: заслуженный деятель науки РФ,
доктор технических наук,
профессор Харченко В. Н.

Официальные оппоненты: доктор технических наук,
доцент Ретинская И.В.
кандидат технических наук,
с.п.с. Мазо Б.Л.

Ведущая организация: Рос НИИ информационных технологий и
систем автоматизации проектирования.

Защита состоится 14 мая 1999 г. в 13⁰⁰ часов
на заседании диссертационного совета Д 053.31.06 при Московском государственном
университете леса по адресу: Мытищи-1, ул. Институтская, д. 1, МГУЛ, ауд. 313.

С диссертацией можно ознакомиться в библиотеке МГУЛ.

Автореферат разослан «__» _____ 1999 г.

Ученый секретарь диссертационного совета, кандидат технических наук,
доцент Дашков А.А.

У 236.23, 0

Общая характеристика работы

Актуальность темы

Документы, напечатанные на бумаге, во многих случаях остаются наиболее удобным средством передачи информации. Однако обработка этой информации требует ее перевода в электронную форму, что и осуществляется системами оптического распознавания текстов – системами OCR (Optical Character Recognition).

В настоящее время эти системы достаточно устойчиво распознают тексты среднего качества. Это позволяет использовать их для преобразования больших массивов текстовой информации. Однако зачастую документ обладает определенной *логической структурой*, например может содержать заголовки, аннотацию, ключевые слова и прочие *реквизиты*. Эта логическая структура должна быть в явном виде отражена в электронном представлении документа. Более того, существует определенный класс документов, так называемые *формы* (платежные поручения, декларации, анкеты, счета), распознавание которых бесполезно без выделения их реквизитов. Таким образом, дальнейшее развитие систем OCR лежит на пути анализа логической структуры печатных документов.

Основной проблемой является то, что расположение реквизитов строго не фиксировано ни относительно листа, ни относительно друг друга. Наиболее перспективный подход к решению данной проблемы основывается на методах структурного распознавания образов. Однако существующие методы были изначально разработаны для распознавания дискретных образов. Изображение же печатного документа является настолько мелкодискретным, что необходимо параметрическое представление.

Цель диссертационной работы

Целью диссертационной работы является разработка системы идентификации логической структуры печатных документов для распознавания печатных форм с не фиксированным расположением полей.

Исходя из цели работы, с учетом анализа проблемы и существующих методов были поставлены следующие задачи:

1. Разработать метод структурного распознавания образов, позволяющий учесть все необходимые геометрические свойства документа, в том числе:
 - разработать алгоритм распознавания образа;
 - создать средства, позволяющие описывать структуру образа.
2. Адаптировать разработанный метод структурного распознавания образов к задаче идентификации логической структуры печатных документов.
3. Создать программную систему, осуществляющую компиляцию описания образа, реализующую алгоритм структурного распознавания и процедуры выделения структурных элементов на изображении печатных документов.

Методы исследования

В диссертации использовались методы оптимизации, элементы теории формальных языков, теории распознавания образов, теории нечетких множеств и нечеткой логики.

Научная новизна

1. Разработана новая система идентификации изображений печатных документов, позволяющая учесть все необходимые геометрические свойства документа.
2. Разработан метод структурного распознавания образов, не связанный с формой представления распознаваемых данных, свойствами структурных элементов.
3. Создан язык структурных описаний, позволяющий описывать структуру образа и отношения между его элементами, основываясь на их параметрическом представлении.
4. Разработана методика инверсии отношений при компиляции структурного описания печатного документа, позволяющая оптимизировать процесс его создания, повысить скорость и вероятность правильного распознавания.

Практическая значимость работы

1. Разработанная система идентификации логической структуры изображения печатных документов позволяет распознавать формы, не имеющие строго фиксированного расположения полей, в частности:
 - формы платежного поручения;

- справки о доходах физического лица.
2. Созданная программа распознавания платежных поручений может быть использована при автоматизации работы с данными финансовыми документами. В настоящий момент программа внедрена:
- более чем в 30 государственных и коммерческих банках, в том числе в 9 отделениях Сбербанка;
 - в налоговых службах РФ.
3. Разработанный метод структурного распознавания образов позволяет распознавать образы, состоящие из достаточно сложных структурных элементов, связанных произвольным набором отношений. Наряду с рассмотренным в диссертационной работе применением, он также используется для распознавания рукописных символов.

Достоверность

Достоверность полученных результатов обеспечивается тем, что используются современные методы и алгоритмы, проверенные на практике в других областях компьютерной математики. Достоверность созданной системы распознавания форм подтверждается хорошим совпадением результатов автоматической идентификации с эталонными, выявленным в ходе тестирования системы, проведенных, в том числе, Сбербанком России.

Апробация работы

Основные положения диссертационной работы докладывались и обсуждались на 4-ой Международной конференции по анализу и распознаванию документов ICDAR (Forth International Conference on Document Analysis and Recognition, Ulm, Germany, 1997 г.), Международной научной конференции "Перспективные технологии автоматизации" (г. Вологда, 1998 г.), Отраслевой конференции по документообороту DOCFLOW (г. Москва 1998 г.).

На защиту выносятся

1. Система идентификации логической структуры печатных документов на основе метода структурного распознавания образов.

2. Модель описания структуры образов и метод структурного распознавания, не имеющие ограничений на типы используемых структурных элементов и вид отношений, связывающих структурные элементы образа.
3. Язык структурных описаний как средство описания образа для его последующего распознавания предложенным методом структурного распознавания.
4. Методика инверсии отношений, позволяющая априорно вычислить дополнительные ограничения на выделяемые структурные элементы.

Публикации

По теме диссертации опубликовано 8 статей и докладов на конференциях.

Структура и объем диссертации

Диссертационная работа состоит из введения, 5 глав, заключения, списка литературы и 4 приложений. Общий объем 153 стр., в том числе 137 стр. основного текста, 20 рисунков, 5 таблиц, 93 наименования списка литературы.

Основное содержание работы

Во введении обоснована актуальность темы исследования, сформулированы цель и задачи исследования, представлены структура, объем и общее содержание диссертации, практическая значимость и основные положения, выносимые на защиту.

В первой главе проведен анализ задачи идентификации логической структуры печатных документов и рассмотрены существующие методы ее решения. Определены доступные для автоматического анализа характеристики документа, в которых отражается его логическая структура: содержимое текстовых элементов, их шрифтовое оформление и геометрическая структура документа. Показана необходимость интегрального учета данных характеристик для создания наиболее универсального метода идентификации логической структуры печатных документов.

Проведен сравнительный анализ информативности основных характеристик печатных форм – основного класса документов, в целях распознавания которого проводились данные исследования. Показана неактуальность использования информации о шрифтовом оформлении и достаточность применения простых

существующих методов для анализа текстового содержимого. Сделан вывод о необходимости сосредоточить усилия на более полном использовании информации о геометрической структуре распознаваемых документов.

Исследованы общие свойства геометрической структуры печатных документов. Выделены основные *структурные элементы*, т.е. элементы, составляющие геометрическую структуру изображения документа: разделительные линии и текстовые области. Определены общие свойства структурных элементов – существует несколько типов структурных элементов, каждый из которых задается определенным множеством атрибутов, в частности прямоугольником, определяющим положение структурного элемента на странице. Сформулированы основные положения о геометрической структуре документа в терминах структурных элементов:

- существуют ограничения на абсолютное и относительное расположение структурных элементов;
- существуют структурные вариации, когда один и тот же элемент может находиться в абсолютно разных местах на различных изображениях;
- структурные элементы могут быть необязательными или пропадать в процессе сканирования.

Проведен анализ существующих методов идентификации геометрической структуры печатных документов. Выделены четыре основных подхода к решению данной задачи. Это, во-первых, методы, использующие фиксированное расположение полей. Они надежны, эффективны и просты, поэтому применяются практически в каждой коммерческой системе распознавания форм, однако имеют существенный недостаток – ограничение области применения специально спроектированными машинночитаемыми формами.

Другой подход ориентируется на документы, обладающие табличной структурой. Он позволяет анализировать широкий класс обычных документов, в частности выделять реквизиты научных статей. Системы такого типа разрабатываются, например, в институте INRIA (Франция), в исследовательском центре CEDAR (Center of Excellence in Document Analysis and Recognition) университета Нью-Йорка (США).

Третий, наиболее перспективный, подход основывается на методах структурного распознавания образов, в частности применяются алгоритмы сопоставления графов соседства. Граф соседства получается, если для каждого элемента документа определены его отношения (выше, ниже, левее, правее) с соседними элементами. Идентификация сводится к сравнению и сопоставлению граф соседства документа и граф соседства его описания. Вместе с тем, существующие методы не позволяют при задании отношений указать допустимый диапазон расстояний, а также не позволяют описывать структурные варианты. Поэтому, несмотря на использование дополнительной информации, такой как положение относительно страницы и распознанный текст, многие типы документов все еще требуют создания специализированных систем. Такие системы и составляют четвертый подход к задаче идентификации логической структуры печатных документов.

Структурные методы идентификации документа разрабатываются, например, в компаниях Деймлер-Бенц и Сименс. Их ограниченность вызвана тем, что лежащие в их основе алгоритмы структурного распознавания образов изначально были ориентированы на анализ дискретных объектов с дискретными отношениями. Изображение же документа является очень мелкодискретным, поэтому необходимо работать с параметрически заданными элементами и отношениями. Подобные проблемы решаются также и в других системах распознавания, например в системе распознавания рукописных символов «Графит», разработанной в НИЦЭВТ. Однако применение этих методов к задаче идентификации структуры документа требует их значительной переработки. Актуальным, следовательно, является разработка метода структурного распознавания образов, допускающего произвольное представление распознаваемых данных и структурных элементов.

С учетом анализа существующих систем и методов идентификации логической структуры печатных документов принята трехуровневая архитектура, разделяющая:

- декларативное описание анализируемого вида документов;
- модель структуры документа и предметные алгоритмы распознавания структурных элементов;

- метод структурного распознавания образов, сопоставляющий описание и изображение.

Отделение декларативного описания документа и алгоритмов идентификации является необходимым условием для создания общего метода идентификации печатных документов. Разделение предметной модели и частных алгоритмов от общего метода сопоставления, во-первых, обеспечивает большую гибкость и расширяемость, а во-вторых, позволяет адаптировать метод сопоставления для распознавания других типов образов.

Последующие главы посвящены разработке метода структурного распознавания образов и его адаптации для решения задачи идентификации логической структуры печатных документов.

Во второй главе описан предлагаемый обобщенный метод структурного распознавания образов. На основании сформулированных свойств геометрической структуры документа выдвинуты следующие дополнительные требования к методу структурного распознавания:

1. Абстрагированное представление структурных элементов, в частности параметрическое.
2. Иерархичность и вариантность описания образа.
3. Возможность задавать произвольные *отношения* – взаимные ограничения на свойства структурных элементов.
4. Возможность задавать необходимые параметры распознавания структурных элементов, в том числе вычислять их по контекстной информации.
5. Необходимость оценки достоверности вариантов распознавания для принятия решения в случае неоднозначности.

С учетом данных требований разработана общая структура модели, в которой выделены три основные сущности – *описание, гипотеза и генератор*: $\langle D, H, g \rangle$

Описание D является декларативной компонентой, описывающей структуру распознаваемого образа. Описание включает как *терминальные*, т.е. примитивные структурные элементы, для которых указываются необходимые параметры распознавания: $D = P_1 \times \dots \times P_M$, так и *нетерминальные*, например *составные элементы*, для которых указаны их составляющие подэлементы и отношения между

ними: $D_C = D_I \times \dots \times D_M \times R$. В частности, весь образ является, как правило, составным элементом верхнего уровня.

Гипотеза H является вариантом распознавания структурного элемента. Гипотеза обладает множеством атрибутов, достаточным для идентификации варианта распознавания. Гипотезе соответствует также число, определяющее степень достоверности данного варианта распознавания: $H = A_I \times \dots \times A_M \times Q$.

Генератор g является абстракцией алгоритма распознавания соответствующего типа структурных элементов. Пользуясь описанием, т.е. параметрами распознавания, и исходными данными, т.е. изображением в том или ином представлении, генератор выдвигает гипотезы о расположении структурного элемента (рис. 1): $g: D \times I \rightarrow H$

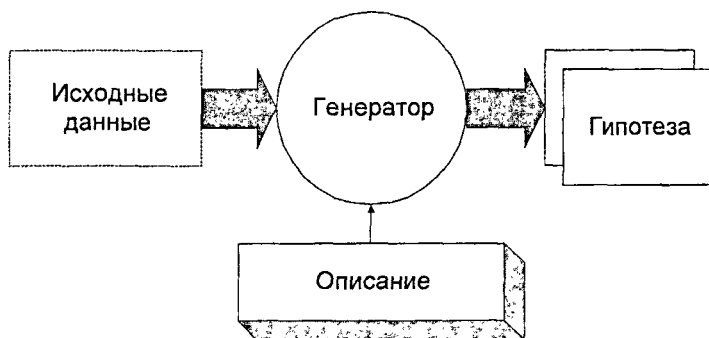


Рисунок 1: Общая структура модели

Для взаимосвязи данных компонент разработаны соответствующие программные интерфейсы и протоколы.

Описанная модель удовлетворяет выдвинутому требованию абстрагированного представления структурных элементов – при ее специализации уточняется множество атрибутов, которыми обладают гипотезы, и определяются параметры реализуемых в виде генераторов предметных алгоритмов распознавания структурных элементов.

Для того чтобы общая модель была способна решать задачу сопоставления иерархического структурного описания и распознаваемых данных, в ней конкретизированы описание, гипотеза и генератор для наиболее важного, вместе с

тем универсального типа элементов – составного элемента. Структура модели, конкретизированная для составного элемента, представлена на рисунке 2.

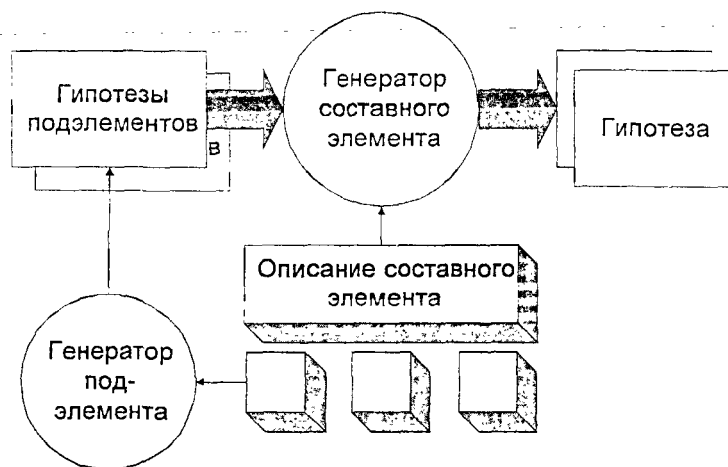


Рисунок 2: Структура модели для составного элемента

Гипотеза, соответствующая составному элементу, в качестве своих атрибутов имеет гипотезы своих подэлементов: $H_C = H_1 \times \dots \times H_M \times Q$. Генератор составного элемента на основании выдвинутых гипотез подэлементов формирует гипотезу составного элемента: $g_c: D_C \times (H_1 \times \dots \times H_M) \rightarrow H_C$. При этом образуется так называемое дерево перебора, узлы которого соответствуют гипотезам подэлементов, а каждый путь от корня к листьям – одной из гипотез распознаваемого составного элемента. Задача состоит в том, чтобы построить пути с лучшей оценкой качества, что осуществляется с помощью стандартного алгоритма A^* , позволяющего управлять перебором.

Тот факт, что гипотезы являются как исходными данными, так и результатом работы генератора, позволяет иметь иерархическое и рекурсивное описание. Рекурсивное описание позволяет описывать образы с регулярно, в том числе и неограниченно, повторяющейся структурой. Однако при этом возникает проблема бесконечной рекурсии, изучение которой в приложении к разработанному методу является темой для дальнейших исследований.

В третьей главе представлены результаты разработки языка структурных описаний. Его предназначение – описание структуры и свойств конкретного образа в соответствии с моделью, поддерживаемой описанным выше методом структурного распознавания образов. Следовательно, язык должен позволять задавать тип и параметры распознавания структурных элементов, описывать устройство составных элементов, в том числе отношения между их подэлементами. В качестве дополнительного было выдвинуто требование декларативности, упрощающей создание и сопровождение описания.

Представлена грамматика основных конструкций языка структурных описаний в нотации Бэкуса-Науэра, которые были разработаны с учетом возможности синтаксического анализа на основе LL(1) разбора:

```

<описание образа> ::= <описание составного элемента>
    | <описание ИЛИ элемента>
<описание составного элемента> ::=
    составной_элемент <имя элемента>
        {<инициализация параметра распознавания>;}*
элементы:
        {<описание подэлемента>}*
    [
отношения:
        {<логическое выражение>;}*
    ]
конец_составного_элемента
<описание ИЛИ элемента> ::=
или_элемент <имя элемента>
        {<инициализация параметра распознавания>;}*
варианты:
        {<описание подэлемента>}*
    [
отношения:
        {<логическое выражение>;}*

```

]

конец_или_элемента

<описание подэлемента> ::=

<описание терминального элемента>

| <описание составного элемента>

: <описание ИЛИ элемента>

<описание терминального элемента> ::=

<тип элемента> <имя элемента>

{<инициализация параметра распознавания>}*

[

отношения :

{<логическое выражение>}*

]

конец_элемента

При разработке части языка структурных описаний, относящейся к заданию параметров и отношений, необходимо совместить требования:

- достаточной выразительности языка, который должен позволять вычислить произвольные выражения при задании параметров распознавания структурных элементов и задать условия на атрибуты порожденных гипотез;
- декларативности и простоты описания.

На основании данных требований была выбрана парадигма функционального программирования. К сожалению, синтаксис существующих функциональных языков, вследствие универсальности, не отвечает в полной мере требованиям простоты и декларативности данного специализированного применения. Кроме того, стандартные средства компиляции и интерпретации с трудом поддаются интеграции, что и послужило основанием для разработки более простого, специализированного языка в функциональном стиле.

Язык включает стандартные математические и логические операторы, операторы сравнения и условный оператор. Поддерживается обращение к атрибутам структурных элементов, в частности к подэлементам составных элементов, вызов

методов и функции, в том числе и с ключевым заданием параметров. Допускается введение локальных констант.

Для разработанного языка созданы компилятор описания и интерпретатор для процедур вычисления параметров и проверки отношений. Использование интерпретатора обеспечивает быстрый цикл настройки-тестирования описания образа.

В четвертой главе рассмотрен способ вычисления оценки достоверности распознавания, основанный на ее вероятностной интерпретации.

Предполагается, что оценка достоверности вариантов распознавания, т.е. гипотез, осуществляется специализированными алгоритмами их распознавания, т.е. генераторами. Однако для того, чтобы сравнивать и комбинировать оценки различных структурных элементов, необходимо иметь единую шкалу и смысл этой оценки. Принято интерпретировать оценку достоверности как оценку условной вероятности:

$Q_i \sim p(t_i|I)$, где t_i – структурный элемент, а I – изображение.

Так как распознавание составного элемента реализуется непосредственно в обобщенном методе, необходимо определить правило вычисления оценки достоверности для составного элемента.

На основании вероятностной интерпретации оценки достоверности имеем:

$$p(N|I) = p(N|t_1, \dots, t_n)p(t_1|I) \dots p(t_n|I),$$

где $p(t_i|I)$ – вероятность i -го подэлемента составного элемента N ,

$p(N|t_1, \dots, t_n)$ – условная вероятность того, что данный набор подэлементов образует составной элемент N .

Следовательно, оценка достоверности распознавания составного элемента может быть вычислена по следующему правилу:

$$Q = Q_R Q_1 \dots Q_n$$

где $Q_i \sim p(t_i|I)$ – оценка достоверности i -го подэлемента составного элемента N

$Q_R \sim p(N|t_1, \dots, t_n)$ – оценка отношений

Величина $p(N|t_1, \dots, t_n)$ соответствует отношениям Q_R , которые ограничивают подэлементы в составном элементе. Таким образом, отношения должны моделировать функцию условной плотности вероятности $p(N|t_1, \dots, t_n)$. Для решения

данной задачи предложено использовать нечеткую логику в вероятностной интерпретации. В этом случае логические операции вычисляют оценку по следующим формулам:

операция ' \wedge ' по формуле: $A \wedge B \rightarrow a * b$;

операция ' \vee ' по формуле: $A \vee B \rightarrow a + b - a * b$;

операция ' \neg ' по формуле: $\neg A \rightarrow 1 - a$.

Примитивные предикаты, из которых с помощью логических операций формируются отношения, также должны вычислять оценку, т.е. формировать функцию условной плотности вероятности. Для этого предлагается использовать кусочно-линейную аппроксимацию (рис. 3).

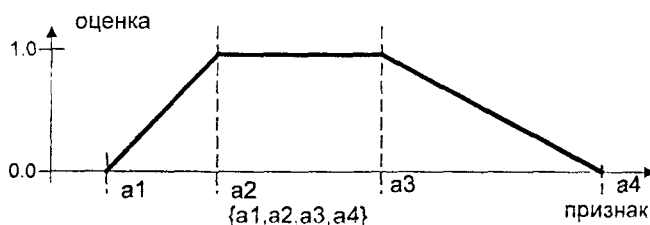


Рисунок 3: Моделирование функции условной плотности вероятности примитивным предикатом

В пятой главе проведена адаптация разработанного обобщенного метода структурного распознавания образов к задаче идентификации логической структуры печатных документов.

Методика адаптации включает в себя следующие действия:

- определение типов используемых структурных элементов;
- определение множества параметров, идентифицирующих результаты распознавания данного типа структурного элемента, т.е. атрибутов гипотезы;
- создание алгоритмов распознавания структурных элементов;
- спецификация множества параметров распознавания для каждого типа структурных элементов;
- определение специализированных отношений.

На основании анализа задачи идентификации структуры печатных документов, проведенного в первой главе, в качестве структурных элементов выбраны следующие:

- горизонтальная или вертикальная разделительная линия;
- цепочка символов из заданного множества;
- ключевое слово;
- фраза – последовательность ключевых слов или цепочек символов, образующая текстовую область с примерно одинаковой высотой строки и межстрочным расстоянием.

Введение фразы в качестве структурного элемента продиктовано соображениями удобства – это очень частый тип текстовой области, используемый в описании печатных документов, многократно определять который в виде составного элемента было бы трудоемко.

Все гипотезы данных типов структурных элементов имеют в качестве атрибутов прямоугольник, задающий их расположение на изображении документа. Для текстовых структурных элементов, т.е. ключевого слова, фразы и цепочки символов, можно получить строку распознанного текста и некоторые параметры шрифтового оформления.

Для текстовых терминальных элементов в описании может быть указан широкий набор параметров распознавания, представленных ниже.

Для цепочки символов:

- основной набор символов, например - «0123456789»;
- символы, которые графически могут походить на заданные, например - «ОЗЧТ» – такие символы допускаются с дополнительным штрафом при оценке гипотезы;
- максимальное расстояние между символами.

Для ключевого слова:

- варианты ключевого слова, например – «Сумма» или «почтой|телеграфом|электронный»;

- допустимое количество ошибок – реализовано нечеткое сравнение, каждая ошибка учитывается при оценке гипотезы, а данный параметр используется для разумного отсека.

Для фразы:

- «пословное» регулярное выражение, для задания текста фразы, например – «Банк плательщика|плат|пл-щика»;
- допустимое количество ошибок на слово относительно его длины – каждое слово фразы сравнивается нечетко;
- максимальное расстояние между словами;
- минимальное и максимальное количество строк во фразе;
- среднее расстояние между строками.

Алгоритмы выделения текстовых структурных элементов должны учитывать возможные ошибки распознавания, поэтому применяются алгоритмы нечеткого поиска с построением конечного автомата на основе текста, заданного для данного структурного элемента.

Для удобства и повышения читаемости описания вводится более простая форма задания некоторых отношений:

Выше: Элемент, Диапазон

Ниже: Элемент, Диапазон

Левее: Элемент, Диапазон

Правее: Элемент, Диапазон

Именно эти отношения чаще всего используются при описании геометрической структуры печатных документов, так как определяют положение текущего элемента относительно других элементов описания. Фрагмент описания формы платежного поручения представлен ниже:

...

Фраза *Банк_плательщика*

Текст: "БАНК ПЛАТЕЛЬЩИКА|ПЛ-ЩИКА";

Наказание_за_пропуск: 0;

Необязателен: 0.94, 0.94;

отношения:

Выше: *Разделитель_Банк_плат*;
 Ниже: *Разделитель_плат_получ - 2 * см*;
 Выше: *Разделитель_плат_получ.Bottom*
 + *высота_строки / 2*;
 Левее: *Большой_верт_разделитель*;
конец_составного_элемента

...

Базовая вычислительная модель системы предполагает, что выделение элементов использует только параметры, указанные в описании. Полученные гипотезы оцениваются и отфильтровываются с помощью отношений. Однако, так как наиболее распространенные отношения в языке описания печатных документов достаточно просты (Выше, Ниже, Левее, Правее), можно попытаться сузить область поиска элементов.

Локализация определяется как множество значений атрибутов структурного элемента, при котором выполняются все его отношения с другими элементами:

$L_t[R(t_1, \dots, t_n, t)]$ - локализация отношения $R(t_1, \dots, t_n, t)$

$R(t_1, \dots, t_n, t)$ - отношения для текущего элемента t с выделенными ранее элементами t_1, \dots, t_n

$L_t[R(t_1, \dots, t_n, t)] = \{ t \mid R(t_1, \dots, t_n, t) > 0 \}$

Так как отношения формируются из примитивных отношений с помощью логических операторов, то необходимо уметь вычислять локализацию для логических операторов:

$$\underline{L}[A(t) \wedge B(t)] = \{ t \mid (A(t) \wedge B(t)) > 0 \} = \{ t \mid A(t) > 0 \} \wedge \{ t \mid B(t) > 0 \} =$$

$$= \underline{L}[A(t)] \wedge \underline{L}[B(t)]$$

$$\underline{L}[A(t) \vee B(t)] = \{ t \mid (A(t) \vee B(t)) > 0 \} = \{ t \mid A(t) > 0 \} \vee \{ t \mid B(t) > 0 \} =$$

$$= \underline{L}[A(t)] \vee \underline{L}[B(t)]$$

$$\underline{L}[\neg A(t)] = \{ t \mid (\neg A(t)) > 0 \} = \neg \{ t \mid A(t) > 0 \} = \neg \underline{L}[A(t)]$$

Вычисление локализации для примитивных предикатов основывается на том, что их оценка формируется кусочно-линейной функцией.

Таким образом, методика инверсии отношений при компиляции описания образа заключается в следующем:

- вводится тип данных – локализация;
- компилируется исходный текст отношений;
- при этом логические операции заменяются на операции с локализацией;
- примитивные предикаты заменяются на соответствующие функции, вычисляющие отношения.

В приложении представлена полная грамматика языка структурных описаний, полный список параметров, атрибутов и отношений, фрагменты описаний формы платежного поручения и справки о доходах физического лица в ГНИ.

Выводы

1. Разработана система идентификации структуры изображений печатных документов, позволяющая учесть все необходимые геометрические свойства документа.
2. Разработан метод структурного распознавания образов, не накладывающий ограничений на типы используемых структурных элементов и вид отношений, связывающих структурные элементы образа.
3. Создан язык структурных описаний, предназначенный для задания структуры и свойств распознаваемого образа.
4. Разработана методика инверсии отношений – методика априорного вычисления параметров распознаваемых структурных элементов печатных документов, позволяющая оптимизировать процесс создания описания, повысить скорость и вероятность правильного распознавания.
5. Разработанная система используется для распознавания форм платежных поручений в ряде государственных и коммерческих банков, в том числе в 9 отделениях Сбербанка России. Создан прототип описания справки о доходах физического лица для Государственной налоговой инспекции.

Публикации:

1. Зуев К.А.. Бинаризация изображений печатных документов // Компьютерная хроника, 1996 г., №2, стр. 95-99.

2. Зуев К.А., Технология анализа документа // Компьютерная хроника, 1996 г., №3, стр. 53-59.
3. Zuyev K.A., Table Image Segmentation // Proceedings of the Forth International Conference on Document Analysis and Recognition (ICDAR), 1997, Ulm, Germany, Vol. II, pp 705-708.
4. Зуев К.А., Система распознавания штрих-кода // Тезисы VII Международной научно-технической конференции «Оптические, радиоволновые, тепловые методы и средства контроля природной среды, материалов и промышленных изделий», 1997 г., г. Череповец, стр. 146.
5. Зуев К.А., Язык структурных описаний // Международная научная конференция "Перспективные технологии автоматизации", 1998 г., г. Вологда.
6. Зуев К.А., Метод выделения полей при распознавании форм // Труды МГУЛ, 1999 г., вып. 302.
7. Зуев К.А., Язык структурных описаний для систем распознавания образов // Труды МГУЛ, 1999 г., вып. 302.
8. Zuyev K.A., Structural Approach to Form Image Understanding // to be published in Proceedings of the Fifth International Conference on Document Analysis and Recognition (ICDAR) 1999.