

УДК 004.93

Е.Е. Пятикоп

Приазовский государственный технический университет, Мариуполь

ПАКЕТ ПРИКЛАДНЫХ ПРОГРАММ ОБРАБОТКИ ИЗОБРАЖЕНИЯ ТЕКСТА НА ОСНОВЕ НЕЧЕТКИХ КОГНИТИВНЫХ МОДЕЛЕЙ

В статье описывается структура пакета прикладных программ обработки изображения. Приводятся статические и динамические модели модулей пакета. Показаны результаты работы модуля локализации строк.

Ключевые слова: обработка изображения, зрительная система, модель, класс, метод, контур, локализация строк.

Введение

В области искусственного интеллекта актуальной задачей является проектирование систем автоматической обработки визуальной информации. Одним из направлений использования этих систем является распознавание текста. Анализ изображения с целью выявления рекламного текста является актуальной задачей в борьбе с «графическим спамом» [1 – 2]. Изобретательность создателей спама требует совершенствования технологий распознавания [3]. Наряду с классическими методами распознавания исследуется возможность применения знаний из когнитивной психологии о зрительном восприятии [4 – 7]. Понимание феномена зрения может повысить надежность и эффективность извлечения описания реального мира из его изображения.

Постановка задачи. Знания о первичной обработке зрительной информации автором формализованы с помощью нечетких множеств [8 – 9]. На основе этих моделей необходимо реализовать пакет прикладных программ обработки изображения текста.

Структура пакета

Пакет состоит из отдельных модулей: *ContourImage*, *ElementsImage*, *LocalString*, *RecognitionSymbol*. Функциональные возможности модулей следующие:

- модуль *ContourImage* реализует задачу выделения контура на изображении на основе нечеткой модели ганглиозных клеток;
- модуль *ElementsImage* представляет контур изображения в виде ориентированных фрагментов на основе нечеткой модели клеток первичной зрительной коры;
- модуль *LocalString* реализует метод локализации строк на изображении;
- модуль *RecognitionSymbol* демонстрирует возможность решения задачи распознавания (идентификации) объектов на примере символов с использованием моделей когнитивных знаний.

На рис. 1 показана структурная схема пакета.

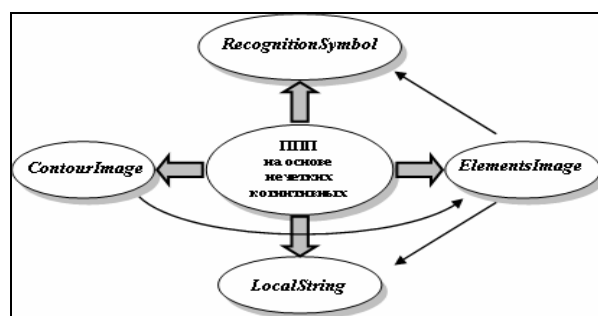


Рис. 1. Структурная схема пакета

Все модули реализованы с использованием объектно-ориентированных технологий построения программных продуктов, которые основываются на описании классов, их атрибутов и методов.

На рис. 2 показаны классы, основанные на нечетких моделях зрительной системы [8, 10]. Для удобства восприятия содержания классов описание их атрибутов, методов и интерфейсов приводится на русском языке. Эта статическая модель используется для работы первых трех модулей.

Между классами, показанными на рис. 2, используются такие интерфейсы:

- «Разбить на клетки», который представляет объект класса *Изображение* в виде объекта класса *Множества клеток* (ганглиозных) с использованием метода «Смоделировать клетку» класса *Клетка*;
- «Создание контура», который объект класса *Множества клеток* преобразует в объект класса *Контур* путем отсеивания части объектов класса *Клетка* с уверенностью *On* или *Off* меньше 0,5;
- Для работы модуля *ElementsImage* используются те же классы и методы, но добавляются еще интерфейсы:
- «Формировать детектор», который из объекта класса *Контур* выбирает по четыре смежных объекта класса *Клетка* и образует объект класса *Множество детекторов* из объектов класса *Детектор*;
- «Формировать колонки» из объекта класса *Множество детекторов* формирует объект класса *Колонка* только из объектов класса *Детектор* с одинаковым углом.

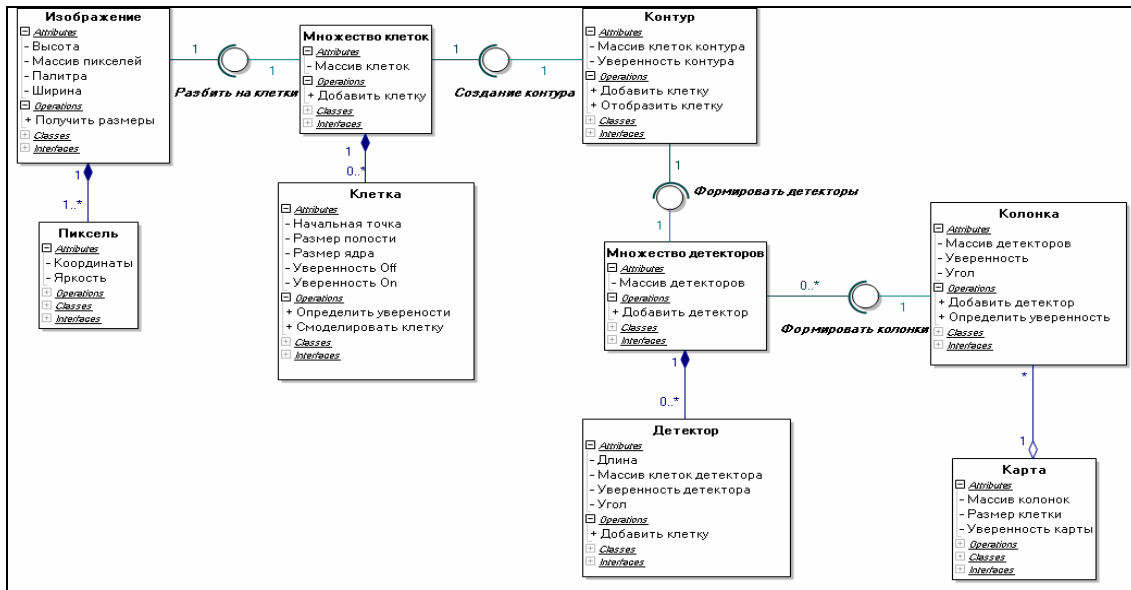


Рис. 2. Диаграмма классов, основанных на когнитивных моделях

Процесс деятельности модуля отображен на рис. 3, где показана диаграмма деятельности (активности) *ElementsImage*, а также отражена связь модулей *ContourImage* и *ElementsImage*.

Входными данными модуля *LocalString* является цифровое изображение текста со строками – графический файл формата Windows bitmap, с глубиной цвета 8 бит (256 градаций серого цвета).

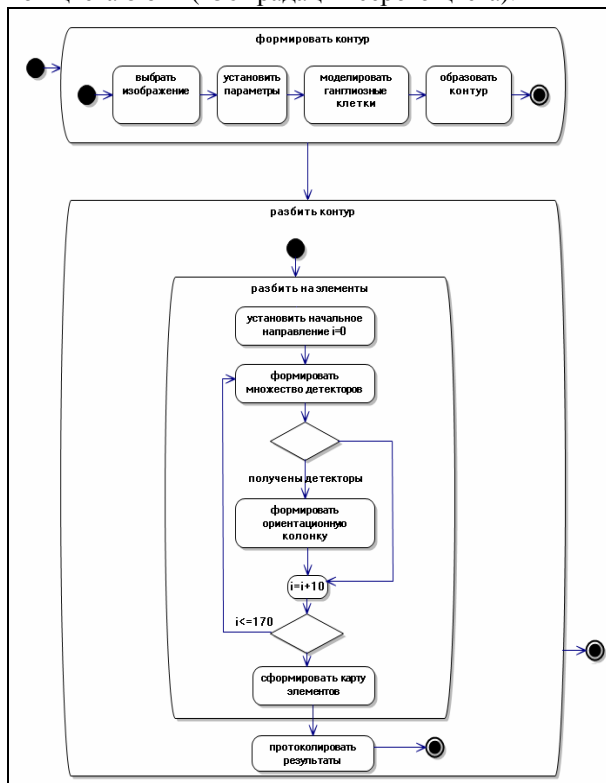


Рис. 3. Диаграмма деятельности модуля *ElementsImage*

Строки текста на изображении расположены под неизвестным углом наклона. Кроме этого неизвестен размер букв текста, т.е. высота строк. Таким образом, задача обработки изображения сводится к этапам:

- определение угла наклона строк текста;
- определение высоты строк текста;
- локализация предполагаемых строк или элементов строк (слов, словосочетаний).

Решение этой задачи выполняется с помощью метода, предложенного в работе [10]. Участие пользователя в этом модуле ограничивается только загрузкой изображения и запуском процесса локализации. Сам процесс локализации не требует взаимодействия с пользователем. Результат локализации визуально отображается на исходном изображении.

В данном модуле используется логическая модель классов компонента *ElementsImage*, которая показана на рис. 3. В дополнение к этим классам формируется новый класс *Множество Карт*, который состоит из множества объектов класса *Карта*. Отличительным атрибутом каждого объекта класса *Карта* является *Размер клетки*, в зависимости от которого определяется и уверенность карты. Это необходимо, поскольку в предложенном методе обрабатывается множество карт для разных размеров рецептивных полей клеток. Алгоритм метода показан на рис. 4 с помощью диаграммы деятельности.

Работа модуля *RecognitionSymbol* направлена на обработку графического файла формата BMP изображения символа. Первоначально с использованием модулей *ContourImage* и *ElementsImage* формируется представление изображения символа в виде множества ориентированных элементов (клеток зрительной коры). Далее из этих элементов формируется описание изображения символа с помощью множества узлов. Совокупность множества узлов и общих ориентированных клеток является представлением символа на изображении, которое сравнивается с базой прототипов. В результате работы модуля образуется упорядоченное множество функций принадлежности изображения символа к каждому из прототипов букв.

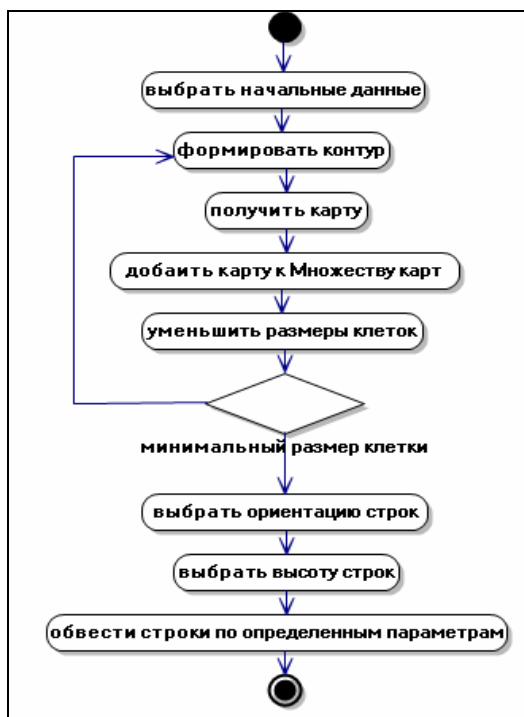


Рис. 4. Диаграмма деятельности LocalString

Статическая структура модуля показана на рис. 5 и содержит такие классы:

- класс Узел построен на основе нечеткой модели узла и состоит из двух объектов класса Детектор;
- класс Множество узлов представляет собой массив объектов класса Узел и формируется в соответствии с методом (интерфейсом) «Формировать узлы»;
- класс Общих детектор наследуется от класса Детектор, но появляется атрибут «Массив общих узлов», который показывает, что детектор является общим для некоторого не пустого множества узлов;

- класс Множество общих детекторов представляет собой массив объектов класса Общих детектор;
- класс Буква построен на основе нечеткой модели буквы, и формируется из Множества узлов и Множества общих детекторов.

Эксперименты

Для проведения экспериментов локализации строк текста было подобрано порядка 80 графических файлов формата BMP. Среди тестируемых файлов были фрагменты текста отсканированных газет, журналов, книг, а также спам-изображения и файлы, полученные как копии экрана. Высота строк на изображениях колебалась от 8 до 35 пикселей угол наклона от 0° до 175°.

Один из экспериментов, который проводился в рамках исследования метода локализации строк текста, направлен на изучение влияния зашумленности фона на результаты локализации. Необходимость исследования зашумленных изображений возникла в результате анализа изображений графического спама. Такие изображения содержат текст, но помимо этого имеют место мелкие элементы шума.

Также изображения с «грязным» фоном могут получиться при сканировании, если была неправильно выбрана чувствительность или порог яркости. В связи с этим было выбрано изображение с белым фоном, а затем для него были применены различные изменения фона. На рис. 6 показаны изображения с различным фоном.

Для всех изображений был применен метод локализации строк, т.е. были автоматически определены угол наклона текста и высота строк. Найденные значения, равные $\alpha=160^\circ$, $\omega=16$ идентичны для всех изображений, но с различной функцией уверенности.

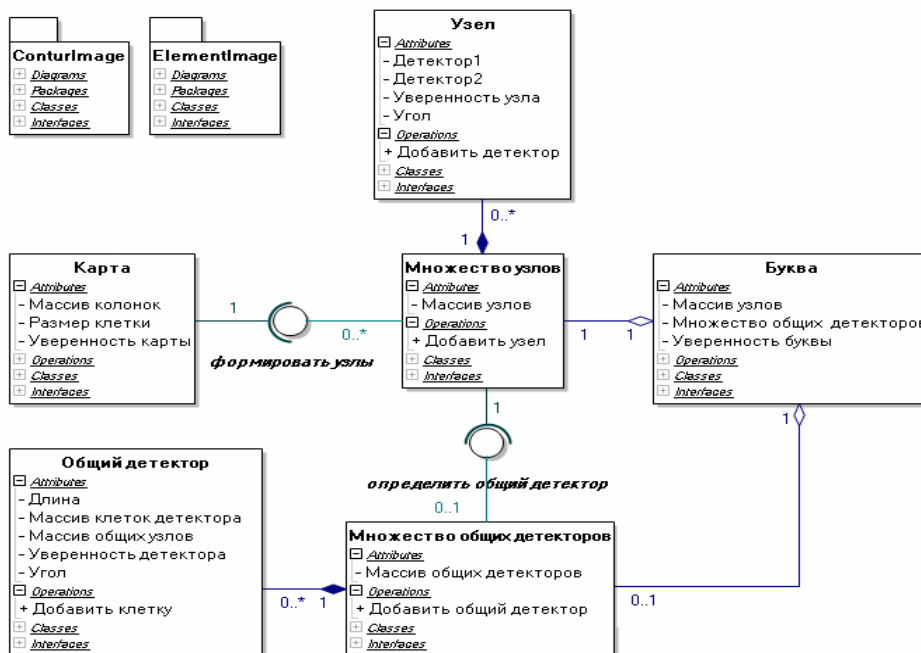


Рис. 5. Диаграмма классов модуля RecognitionSymbol

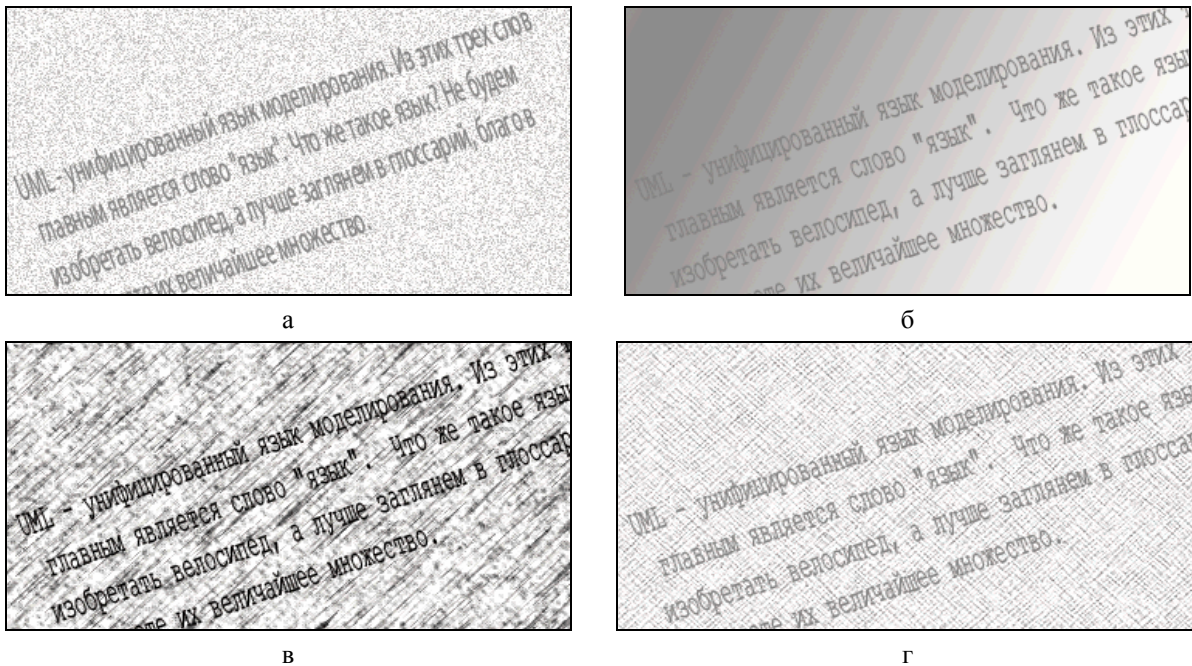


Рис. 6. Варианты изображений с зашумленным фоном

На рис. 7 приведены функции уверенности для каждого из изображений рис. 6, где последний столбец показывает уверенность локализации строк на изображении с белым фоном.

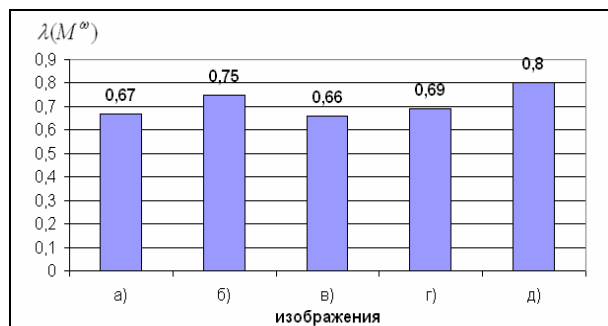


Рис. 7. Функции уверенности для изображений с зашумленным фоном

В результате разработки данного пакета были успешно реализованы нечеткие когнитивные модели зрительной системы, использование которых позволило решить задачу локализации строк текста на изображениях с зашумленным фоном под любым углом наклона и любой высоты [9]. Модули данного пакета имеют дальнейшее применение при разработке систем по борьбе с графическим спамом.

Список литературы

1. Рецидив графического спама. Спамтест [Электронный ресурс]. – Режим доступа: <http://www.spam-test.ru/news?id=207509462>.
2. Яфизов Р. Проактивные антиспам-технологии [Электронный ресурс] / Р. Яфизов // "Information Security/ Информационная безопасность" – 2008. – №4. – Режим доступа к статье: http://www.itsec.ru/articles2/Oborandeh/proaktivn_antispam_tehnologii.

3. Академики обсудили новые методы борьбы со спамом / Сnews|Бизнес [Электронный ресурс]. – Режим доступа к ресурсу: <http://www.cnews.ru/news/line/index.shtml?2007/04/05/244199>.

4. Ярмошевич Е.И. Функциональная спектральная пространственно-временная модель формирования изображений объектов зрительной системой человека [Электронный ресурс] / Е.И. Ярмошевич // Вестник ВГУ. Серия: системный анализ и информационные технологии. – 2008. – № 1. – С. 74-78. – Режим доступа к статье: www.vestnik.vsu.ru/pdf/analiz/2008/01/yarmoshevich.pdf.

5. Гладиллин С.А. Нейронная сеть, воспроизводящая выходной сигнал ганглиозной клетки [Электронный ресурс] / С.А. Гладиллин, Д.Г. Лебедев // Информационные процессы. – 2005. – Т. 5, № 3. – С. 258-264. – Режим доступа к статье: <http://www.jip.ru/2005/258-264.pdf>.

6. Шиффман Х.Р. Ощущение и восприятие: пер. с англ. З. Замчук / Х.Р. Шиффман. – СПб.: Питер, 2003. – 928 с.

7. Андерсон Дж. Когнитивная психология: пер. с англ. С. Комаров; 5-е изд. / Дж. Андерсон. – СПб.: Питер, 2002. – 496 с.

8. Каргин А.А. Моделирование функционирования релейных клеток зрительной системы в задачах интерпретации изображений / А.А. Каргин, И.А. Доценко, Е.Е. Пятикоп // Информационно-управляющие системы на железнодорожном транспорте. – 2007. – № 5,6. – С. 10-12.

9. Пятикоп Е.Е. Некоторые результаты компьютерных экспериментов локализация строк текста на основе когнитивных моделей восприятия изображения / Е.Е. Пятикоп // Вісник Донецького університету. Сер. А: Природничі науки. – 2008. – № 2, ч. 2. – С. 527-532.

10. Каргин А.А. Применение нечетких моделей когнитивных знаний восприятия изображения в задаче локализации строк текста / А.А. Каргин, Е.Е. Пятикоп // Искусственный интеллект. – 2009. – №2.

Поступила в редколлегию 26.06.2009

Рецензент: д-р техн. наук, проф. А.А. Каргин, Донецкий национальный университет, Донецк.

**ПАКЕТ ПРИКЛАДНИХ ПРОГРАМ ОБРОБКИ ЗОБРАЖЕННЯ ТЕКСТУ
НА ОСНОВІ НЕЧІТКИХ КОГНІТИВНИХ МОДЕЛЕЙ**

О.Є. П'ятикоп

У статті описується структура пакету прикладних програм обробки зображення. Приводяться статичні і динамічні моделі модулів пакету. Показані результати роботи модуля локалізації рядків.

Ключові слова: обробка зображення, зорова система, модель, клас, метод, контур, локалізація рядків.

THE PACKAGE OF APPLICATIONS OF IMAGE PROCESSING TEXT BASED ON FUZZY COGNITIVE MODELS

O.Ye. Pyatikop

The structure of application of treatment of image package is described in the article. Static and dynamic models over of the modules of package are brought. Job of the module of localization of lines performances are shown.

Keywords: treatment of image, visual system, model, class, method, contour, localization of lines.