

ОБ ИСПОЛЬЗОВАНИИ ПЕРЕМЕННОЙ РАЗРЕШАЮЩЕЙ СПОСОБНОСТИ В СИСТЕМАХ ОБРАБОТКИ ВИЗУАЛЬНОЙ ИНФОРМАЦИИ

В.Г. Калмыков

Институт проблем математических машин и систем НАН Украины

e-mail: vl.kalmykov@gmail.com

В современных информационно-технологических системах, в частности, в СППР, в процессе принятия решения предполагается совместная работа зрительного восприятия человека – оператора и средств машинного зрения.

Для оптимального взаимодействия человека – оператора и средств машинного зрения в рамках одной системы должны быть учтены физиологические особенности зрительного восприятия.

Исследования в области физиологии зрительного восприятия и создания информационных технологий автоматической обработки визуальной информации (иначе говоря, технического – компьютерного зрения) объективно являются взаимосвязанными областями человеческой деятельности. Действительно, предметом обеих научных дисциплин является исследование механизмов зрительного восприятия. Если предметом физиологии зрения является исследование процессов зрительного восприятия человека и животных, то одним из предметов компьютерных наук является создание и исследование средств технического зрения. Прогресс в одной из этих областей должен был бы инициировать прогресс в другой области.

К сожалению взаимного междисциплинарного обмена результатами не происходит, по-видимому, из-за очень слабого взаимодействия между специалистами, представляющими разные научные направления, которое, в свою очередь, определяется существенно различными методами, характерными для перечисленных научных направлений.

Многие известные исследователи в области распознавания образов и обработки изображений вообще не упоминают о системе зрительного восприятия как о естественном прототипе технического зрения [1], [2].

Другие авторы используют в качестве прототипов упрощенные модели [3], в которых не отображены важные, уже известные элементы механизмов зрительного восприятия.

Обычно предполагается [2], что исходное изображение, подлежащее обработке, задано в аналоговой форме. Как правило, изображение ограничено прямоугольником приемлемых для обработки размеров. Рассматриваемое исходное изображение обязательно должно полностью помещаться в этом прямоугольнике и, по возможности, заполнять его.

Первой из операций, которую выполняют над обрабатываемым изображением, является дискретизация и квантование, причем эти параметры системы технического зрения выбираются обычно из практических соображений.

Предполагается, что разрешающая способность системы технического зрения выбрана так, что не возникают ненужные подробности при слишком большой разрешающей способности (например, не искажаются контурные линии объектов) и не пропадают существенные детали объектов, как это бывает при слишком малой разрешающей способности.

Таким образом, изображение можно представить в виде двумерного массива $V(N,M)$ шириной N и высотой M , причем каждому элементу этого массива $v(n,m)$ соответствует либо значение яркости элемента изображения – пиксела с координатами n,m для однотонного изображения, либо значения яркостей базовых цветов, например, красного, синего, зеленого для цветного изображения.

Разработано и успешно применяются большое количество методов и алгоритмов для обработки изображений, представленных в виде вектора или матрицы. Большинство этих методов успешно применяются для обработки полутоновых изображений, если удач-

но выбраны параметры дискретизации и квантования для данного класса изображений. В то же время необходимо отметить следующее. При обработке изображений методами статистического распознавания, например, при определении меры сходства двух изображений как $f(V_1, V_2)$, используются значения яркостей всех пикселей, образующих изображение. Однако пиксели в поле изображения обычно относятся либо к объекту, либо к фону. В таком случае результат обработки зависит не только от значений яркости пикселей интересующего объекта, но и от значений яркости пикселей фона,

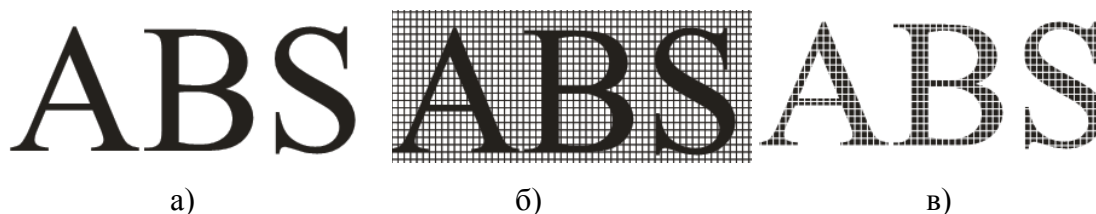


Рис.1 Примеры произвольного текста: а) фон нейтральный, б) фон – решетка совпадает с цветом текста, в) на текст наложена решетка, совпадающая с цветом фона.

что во многих случаях неприемлемо. В качестве примера можно привести изображение произвольного текста на однотонном фоне и на фоне некоторой произвольной сетки (рис.1).

Текст на рис. 1а может быть распознан как статистическими, так и структурными методами распознавания. Такой же текст, но наложенный на произвольную решетку (рис.1б) представляет, с точки зрения распознавания, гораздо более сложную задачу. В случае попытки использования статистических методов результат вычисления сходства с эталонными изображениями будет искажен за счет наличия в поле изображения пикселей, относящихся к фону, отображающих произвольно наложенную решетку, причем пространственное расположение линий решетки меняется от изображения к изображению. В случае же попытки использования структурных методов на изображениях вида 1б невозможно получить контуры объектов. Вместо контуров объектов будут выделены контуры клеток решетки. В настоящее время внимание специалистов привлекают задачи распознавания текстур, частным случаем которых является и наложенная на текст решетка. Однако методы распознавания текстур, как правило, значительно превосходят по сложности задачи распознавания текста, лиц и т.п., что затрудняет или исключает возможность их применения при решении подобных задач. Похожая, но несколько другая ситуация возникает при произвольном наложении на текст решетки цвета, совпадающего с цветом фона – рис.1в. В данном случае при использовании статистических методов результат вычисления сходства с эталонными изображениями будет искажен за счет наличия в поле изображения пикселей, относящихся к объектам, отображающих произвольно наложенную решетку, причем пространственное расположение линий решетки, как и ранее, меняется от изображения к изображению. Так же как и для рис. 1б попытки использования структурных методов на изображениях вида рис. 2 невозможно получить контуры объектов – вместо контуров объектов будут выделены контуры клеток решетки.

В то же время зрение человека справляется с подобными задачами совершенно незаметно для нас, по-видимому, на подсознательном уровне. Поэтому актуальным является изучение результатов нейрофизиологических исследований зрительного восприятия и использование их при создании средств технического зрения [5].

Некоторые исследованные свойства нейронных структур зрительной системы [4] позволяют сделать вывод, что разрешающая способность зрения, изменяется в течение зрительного акта от очень низкой до предельно возможной.

Возможно, именно это свойство зрительной системы позволяет ответить на вопрос, почему изображения знаков на рис. 1а, 1б, 1в воспринимаются зрительной системой чело-

века практически одинаково уверенно, в то время как для современных средств обработки



Рис.2 Изображения рис. 1, причем дискретизация выполнена при низком разрешении.

визуальной информации задача распознавания изображений, таких как рис. 1б, 1в, неразрешима, при этом рис. 1а распознается уверенно. Можно предположить, что именно обработка наблюдаемых изображений с низкой разрешающей способностью в начале зрительного акта, когда "не видны" мелкие детали текстуры дает возможность уверенного зрительного восприятия изображений знаков на различном фоне. Если это так, тогда можно предположить, что если некоторая программа распознавания текстов успешно обработает изображения знаков (рис. 1а), отсканированные с низкой разрешающей способностью, то она также успешно обработает и изображения (рис. 1б, 1в), отсканированные с такой же – низкой разрешающей способностью.

Проведем простой демонстрационный эксперимент с широко известной программой обработки и распознавания текстов FineReader.

В качестве исходных используем изображения на рис.1. Результат распознавания исходных изображений: рис.1а – успешно, остальные – отказ от распознавания по причине невозможности найти (определить) объект в поле изображения.

На рис.2 представлены те же самые изображения, причем дискретизация выполнена с существенно меньшим разрешением. Результат распознавания всех этих изображений – успешно.

Результат проведенного простого демонстрационного эксперимента позволяет сделать обоснованное предположение, что восприятие – распознавание изображений в зрительной системе происходит на всех этапах зрительного акта, в том числе и на начальном этапе, когда изображение "расфокусировано", но зато и мешающие факторы, как например, в рассмотренном случае, в виде текстурных помех не оказывают влияния. Использование этого результата в средствах обработки визуальной информации позволяет совершенно естественным образом достигнуть инвариантности относительно масштаба объектов, выполнять распознавание объекта на фоне, искаженном высокочастотными помехами, также выполнять распознавание объектов, например, письменных знаков, элементы которых представляют мозаику или образованы элементами различных текстур.

Литература

1. М.Шлезингер, В.Главач Десять лекций по статистическому и структурному распознаванию. – Київ: Наукова думка, 2004. – 535 с.
2. Т. Павлидис Алгоритмы машинной графики и обработки изображений. М Радио и связь 1986г. 400 с.
3. У. Прэтт Цифровая обработка изображений: Пер. с англ.—М.: Мир, 1982.—Кн.1—312 с.
4. Подвигин Н.Ф. Динамические свойства нейронных структур зрительной системы Ленинград: Наука, 1979. – 158 с.
5. Калмыков В.Г. Переменная разрешающая способность как фактор развития средств обработки визуальной информации // International Journal "Information Technologies and Knowledge". – Sofia в печати.