

Обзор принципов первичной обработки изображений человеческим глазом применительно к вопросам компьютерного зрения.

Ю.В. Слынько

Московский физико-технический институт (государственный университет)

slynko_yv@pisem.net

Общепризнанно, что человеческий глаз является самым эффективным и быстрым обработчиком изображения. Современные алгоритмы компьютерного зрения по всем показателям уступают алгоритмам, заложенным в человеческом мозге и глазе.

В наше время ведутся активные исследования принципов обработки изображений человеческим мозгом, однако существенных результатов, проливающих свет на используемые человеком алгоритмы, до сих пор не существует. Хорошо известно, как глаз устроен физически, какие клетки к какой длине волны чувствительны и т.д. Также известны некоторые свойства человеческого восприятия, выраженные чаще всего в так называемых «обманах зрения». Однако хоть сколько ни будь полного описания обработки человеком изображения не существует – поэтому при разработке алгоритмов компьютерного зрения редко удается опираться на функционирование человеческого зрительного аппарата.

В то время как алгоритмы, применяемые мозгом, нам не известны, наука многое может сказать о предварительной (первичной) обработке изображений мозгом человека и млекопитающих. Эти сведения изложены, например, в [1] и [2].

Условно первичную обработку изображения человеком можно разделить на три этапа:

1. Механический – управление направлением оптической оси глаза.
2. Окологлазная обработка – обработка изображений внутри глаза и в близлежащих нейронах (наружном коленчатом теле).
3. Обработка в коре головного мозга.

Механическая обработка

Человеческий глаз отнюдь не является неподвижным. Среди его движений можно выделить два принципиально различных класса:

1. Небольшие колебательные движения с частотой ~ 100 Гц и амплитудой $20''$ – $40''$. Такая амплитуда означает, что изображение на сетчатке глаза смещается на размер одного фоторецептора.
2. Направление взгляда смещается вдоль контуров и останавливается в наиболее интересных объектах.

Такие движения глаза говорят о том, что на сетчатку попадает не произвольное изображение, а такое, которое наилучшим образом может быть обработано. Более того, существует обратная связь между мозгом, который обрабатывает изображение, и глазом, который смотрит так, чтобы давать мозгу оптимальное изображение для обработки.

Окологлазная обработка

Человеческий глаз посылает в кору головного мозга не просто сигналы, снятые с фоторецепторов. Изображение предварительно обрабатывается нейронами глаза и наружного коленчатого тела (НКТ).

Удалось установить, что нейроны НКТ формируют изображения разного масштаба. Масштаб изменяется от грубого до точного приблизительно за четверть секунды, а затем скачком возвращается опять к грубому. Таким образом, НКТ строит пирамиду изображений (выражаясь языком компьютерного зрения) разного масштаба.

Обработка в коре головного мозга

Из наружного коленчатого тела информация подается на кору головного мозга. Сведений о дальнейшей ее обработке на данный момент очень мало. Но кое что существенное можно сказать.

Известно, что некоторые нейроны головного мозга реагируют на конкретные изображения. Например, существуют нейроны, которые возбуждаются при появлении в определенном месте поля зрения точки, линии, угла или отрезка. При этом для линии и отрезка существенно не только их положение, но и ориентация. При смещении объекта или изменении его ориентации на незначительную величину начинает возбуждаться уже другой нейрон. Более того, существуют нейроны, реагирующие на такие сложные фигуры, как круги и треугольники, а также на движущиеся в определенной области поля зрения предметы.

Таким образом, можно сказать, что мозг выделяет на изображении характерные участки и особенности. Однако это отнюдь не означает, что мозг выделяет только такую информацию. Известно также, что мозг производит некоторое преобразование, сходное с преобразованием Фурье. Существуют нейроны, реагирующие на появление периодических конструкций в поле зрения – а поскольку такие нейроны существуют для разных пространственных частот и ориентаций, то можно говорить о том, что происходит преобразование Фурье по системе периодических функций.

Также необходимо отметить, что мозг обрабатывает только изменения в изображении. Такие изменения, в частности, обеспечиваются высокочастотными колебаниями глаза. Однако если к глазу приклеить некоторое изображение, то через некоторое время мозг перестанет воспринимать такое изображение.

Что касается дальнейшей обработки, то о ней сказать что либо определенное трудно. Но есть один факт, который необходимо отметить. При восприятии изображения человеческий мозг прежде всего оценивает общий характер картины, а только затем (и в соответствии с этой оценкой) детали. Таким образом, уже на начальном этапе мозг принимает некоторые решения об изображении и строит дальнейшую обработку в соответствии с этими решениями, чего не удастся достичь на компьютере, т.к. первые решения об изображении можно получить только после продолжительной обработки.

Таким образом, видно, что предварительная обработка изображений в зрительной системе человека во многом сходна с основными принципами, используемыми в компьютерном зрении. Ясно также, что количество одновременно используемых методов велико, и соответственно велико и количество обрабатываемой информации. И что самое главное, все эти методы работают в человеческом глазе в рамках одного алгоритма, т.е. существует система объединения самой разнородной информации, такой как, например, координаты отрезков и Фурье преобразование изображений.

Литература

1. В. Демидов «Как мы видим то, что видим», <http://galactic.org.ua/Biblio/Bibl.htm>.
2. Д. Хьюбел. «Глаз, мозг, зрение», Москва, Мир, 2003.