

УДК 004.9

ББК 32.97

## **СИСТЕМА ОТСЛЕЖИВАНИЯ НАПРАВЛЕНИЯ ВЗГЛЯДА С ИСПОЛЬЗОВАНИЕМ ДОСТУПНОЙ ВИДЕОАППАРАТУРЫ**

И.К. Малин, А.В. Крапивенко

### **Аннотация**

Технологии отслеживания направления человеческого взгляда (eye-tracking) используются в различных системах проектирования компьютерных интерфейсов. Время фиксации глаза и плотность траектории позволяют делать выводы о том, каким элементам рассматриваемого объекта или изображения уделяется наибольшее внимание. Также эта технология потенциально способна дать ответ на вопрос, какие дефекты, возникающие при сжатии видеоизображения с потерей качества, наиболее заметны. В данной статье рассказывается о создании системы отслеживания направления взгляда с использованием доступной видеоаппаратуры, что позволяет расширить область использования таких систем.

### **Ключевые слова**

технологии отслеживания взгляда; eye-tracking; видеосжатие; интерфейс; эргономика; видеоаппаратура.

### **Сферы применения систем отслеживания зрачка**

Основную часть информации об окружающем мире человек воспринимает с помощью зрения. Элементы зрительной системы – глаз, нерв и зрительный анализатор головного мозга – тесно связаны между собой, поэтому изучение траектории движения глазного яблока позволяет делать выводы о процессе распознавания зрительных образов и мыслительных процессах человека в целом. Отслеживание направления взгляда также дает возможность строить принципиально новые интерфейсы взаимодействия между человеком и

техническими средствами. В связи с этим развитие систем, позволяющих определять положение точки зрения (eye-tracking), является актуальной научно-технической проблемой.

Сферу применения систем eye-tracking'a можно разделить на две основные части. Во-первых, это системы, которые осуществляют пассивный сбор и анализ информации о движении глаза в течение эксперимента. К этой области можно отнести основное на настоящий момент коммерческое использование eye-tracking'a – оценку визуального удобства интерфейсов приложений и веб-страниц (т.н. usability). Объективная информация о том, на какие элементы интерфейса чаще всего обращает внимание пользователь во время работы с приложением, может быть дополнена стандартным методом анкетирования. В медицине eye-tracking используется для диагностики заболеваний нервной системы – при некоторых из них движения глаз больного и здорового человека отличаются.

Из перспективных направлений, eye-tracking возможно будет применяться для идентификации личности, так как характер движений глаз индивидуален, как почерк, но практически невозможен для имитации. Кроме того, анализ траекторий может использоваться для диагностики заметности визуальных артефактов, возникающих в системах видеобработки при высоких степенях сжатия алгоритмов с потерей качества.

Вторую крупную область применения технологии представляют интерактивные системы, позволяющие осуществлять управление различными устройствами с помощью глаз. К примеру, компанией Canon был разработан опытный экземпляр видеокамеры, где точка фокусировки определяется по тому, куда смотрит оператор. Eye-tracking используется в шлемах пилотов, когда руки заняты другими элементами управления. Также, для парализованных людей этот метод взаимодействия с компьютером иногда является единственно возможным [1].

### **Технологические разновидности систем eye-tracking'a**

Исследования движений глаза проводились уже в конце XIX - начале XX веков, в частности, офтальмологом Луи Эмилем Жавалем, психологами Эдмундом Хью и Гаем Баксвеллом. Одной из наиболее известных в этой области является работа русского психолога Альфреда Ярбуса «Роль движений глаз в процессе зрения» [2]. Был исследован характер движения точки ясного зрения при чтении текстов и разглядывании объектов и изображений, установлено, что глаза движутся не непрерывно, а резкими скачками (саккадами) с некоторым временем фиксации между ними, в течение которого совершаются значительно менее амплитудные микросаккадические движения в окрестностях точки взгляда. Также была установлена достаточно очевидная связь между когнитивной

деятельностью человека и траекторией глаза. В зависимости от того, что именно являлось важным для рассматривающего человека на изображении, менялась плотность и форма траектории в различных его участках.

Современные системы eye-tracking'a различаются по способу получения данных о положении глаза. Во-первых, это контактные системы, которые требуют установки датчиков непосредственно на органы человека. Вышеописанные опыты Ярбуса проводились именно с помощью контактного механического eye-tracker'a – установленная на глаз присоска соединялась с пишущим устройством. В настоящее время контактные системы также используются, причем двух видов. Линза, установленная на глаз, может иметь зеркальную поверхность, в этом случае о движениях глазного яблока судят по отражению света от нее. Также иногда в линзу монтируют миниатюрную катушку индуктивности и помещают испытуемого в высокочастотное магнитное поле, что позволяет вычислить положение глаза с высокой точностью.

Во-вторых, это бесконтактные оптические системы. В этом случае производится цифровая видеозапись движения одного (выбранного) глаза. В настоящее время съемка чаще всего ведется на специальную аппаратуру в инфракрасном диапазоне, что позволяет подсветить глаз исследуемого человека, не ослепляя его [3]. Также о положении взгляда судят по взаимному расположению лучей направленного инфракрасного освещения, отраженного от различных частей глазного яблока – так называемый метод изображений Пуркинье.

Однако, для реализации вышеперечисленных технических решений необходимо специальное оборудование и лабораторные условия. В связи с этим актуальной является задача создания системы отслеживания направления взгляда, их не требующей. Цифровые видеокамеры, включая таковые в фотоаппаратах и мобильных телефонах, широкодоступны в данный момент. Поэтому в качестве регистрирующего устройства системы был выбран неконтактный оптический способ. Объектом фокусировки человеческого зрения в эксперименте являлись специальные тестовые изображения на экране. Описание процесса построения такой системы и первые результаты её опытной эксплуатации приведены в данной работе ниже.

### **Локализация точки ясного взора в экранных координатах**

Задачу определения точки взгляда по видеозаписи движения глаза можно разделить на две основных подзадачи. Во-первых, это определение положения зрачка в каждом кадре (подзадача распознавания). Во-вторых, это сопоставление положения зрачка на видео точкам

на экране, в которые действительно смотрит пользователь (подзадача проецирования). Рассмотрим последовательно алгоритмы решения этих подзадач.

С точки зрения физики, зрачок представляет собой почти абсолютно черное тело, то есть отношение отраженного и поглощенного световых потоков близко к нулю. Это обеспечивает сильный контраст между областями зрачка и радужной оболочки в кадре. В зависимости от условий освещения, настроек баланса белого у видеокамеры и цвета радужки глазного яблока испытуемого, контраст является различным на разных цветовых каналах. В нашем случае, в условиях электрического освещения лампой накаливания и правильно выбранного баланса белого при рассмотрении кадра в цветовой системе RGB наибольший контраст достигается на красном канале. Именно его будем использовать в качестве исходного изображения.

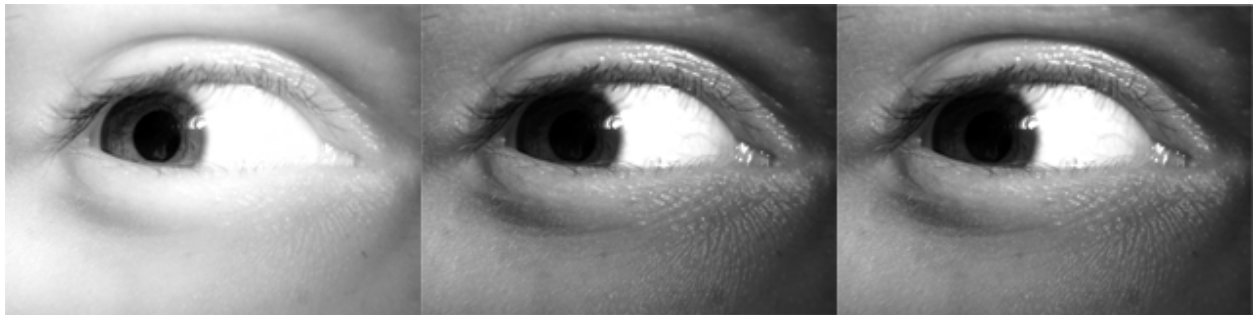


Рис 1. Слева направо: R-, G- и B-каналы видеокadra.

В качестве следующего шага алгоритма, бинаризуем полутоновое изображение путем вычисления функции инвертированного порога, взяв в качестве аргумента красный канал кадра.

Если рассматривать изображение  $I$  как функцию яркости от координат  $I(x, y)$ , то функционал инвертированного порога можно определить как

$$Tinv(I, T) = I': I'(x, y) = \begin{cases} I_{max}, & I(x, y) \leq T \\ 0, & I(x, y) > T \end{cases}$$

где  $T$  – значение порога,  $I_{max}$  – максимально возможная яркость точки в данном формате изображения. Область определения функции – пространство полутоновых изображений. Область значений функции – пространство бинарных изображений того же размера.

Отдельной проблемой стоит выбор значения порога  $T$ . Глаз может быть освещен неравномерно, в этом случае яркость пикселей как самого зрачка, так и пикселей окружающей его радужки будет различаться для разных направлений взгляда. Чувствительность распознавания зрачка в зависимости от порога  $T$  продемонстрируем на примере:



Рис 2. Направление взгляда – влево-вверх. Слева порог бинаризации равен 100, справа 30.

Очевидно, что когда человек смотрит в левый верхний угол, порог должен быть равен 100 – тогда зрачок правильно отсекается по яркости, в то время как в случае порога в 30 выбирается только часть нужной области.

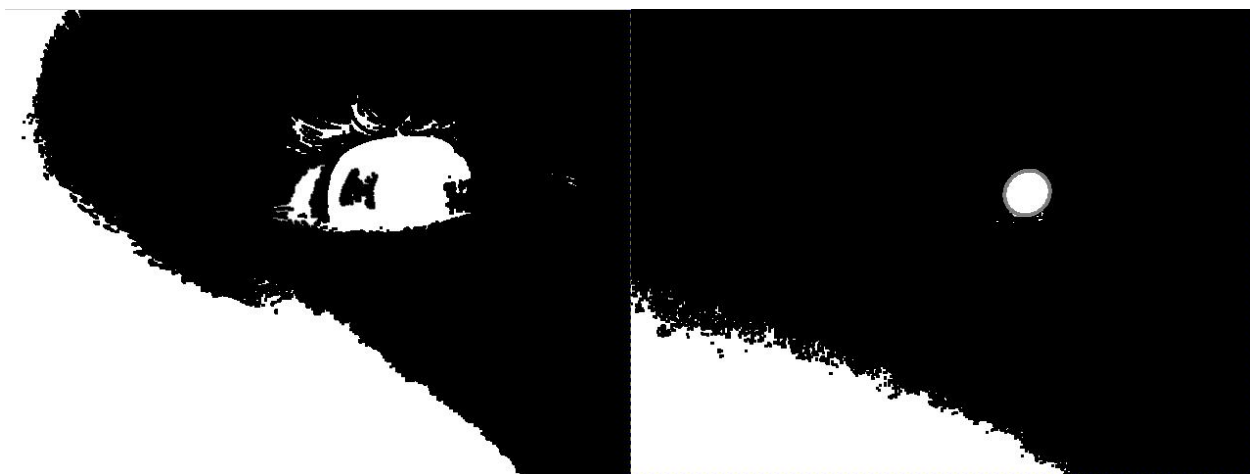


Рис 3. Направление взгляда – вправо-вниз. Слева порог бинаризации равен 100, справа 30.

Здесь ситуация обратная. Порог должен быть равен 30, тогда область зрачка отсекается верно. Так как изначально мы не знаем распределение средней освещенности по областям, будем использовать метод динамической пороговой бинаризации, используя априорные знания об отслеживаемом объекте. Здесь речь идет об известной нам антропометрической информации, что зрачок представляет собой эллипс, а также его размерах и местоположении на предыдущем кадре. Перебирая возможные значения порога для нового кадра, останавливаемся на том, для которого размер и положение зрачка оказываются максимально близки к аналогичным величинам на предыдущем кадре.

Затем применим к изображению морфологическую операцию замыкания изображения. В качестве ядра операции будем использовать квадрат размером 3 на 3

пиксела. Это позволит уменьшить количество отдельных замкнутых контуров, объединяя их в большие по размеру и сглаживая шум.

На полученном изображении выделим замкнутые контуры. Для задачи обнаружения замкнутых контуров на бинарном растровом изображении существуют решения, предлагаемые специалистами в области компьютерного зрения и обработки видеосигналов. С таким алгоритмом можно ознакомиться, например, в работе Сузуки и Эйба [4].

Затем аппроксимируем выделенные контуры эллипсами. Наилучшая аппроксимация контура эллипсом определяется методом наименьших квадратов. Данный подход изложен в работе Фитцгиббона, Пилу и Фишера [5].

Далее, среди множества полученных эллипсов требуется отобрать именно тот, который представляет собой границу зрачка. Для этого на этапе калибровки системы будем определять характерные размеры зрачка и затем выбирать эллипс, размеры которого наиболее близки к характерным, методом наименьших квадратов.

После выполнения последовательности этих шагов для каждого кадра, мы получим список координат центра эллипса, но пока – в системе координат отснятого видеоролика. Поэтому необходимо установить соответствие между координатами точки ясного зрения в системе видеоролика и экранными координатами демонстрируемого тестового изображения. Для этого перед началом эксперимента требуется провести геометрическую калибровку системы. В зависимости от выбранного алгоритма отображения координат, используется калибровка по различному количеству точек. Как правило, это углы рассматриваемого экрана, его центр, а также середины его краев. Чем большее количество точек используется при калибровке, тем точнее получается отображение. В работе Гленструпа и Энгела-Нильсена [1] описывается метод калибровки на основе нейронных сетей. В этом случае нейронная сеть обучается за довольно длительное время (приводятся интервалы порядка получаса), в течение которого пользователь следит за перемещающимся по экрану курсором, однако достигается высокая точность отображения, а также большая степень устойчивости к перемещению головы.

Следует отдельно отметить проблему смещения головы при калибровке. Обычно требуется неподвижность головы во время всего эксперимента. Для этого используются рамки, упоры для подбородка и т.н. bite-bar'ы (неподвижные валики, которые человек прикусывает и таким образом фиксирует свою голову). На основе накопленного опыта первой эксплуатации системы, также отметим исключительную важность качественного освещения при съемке. При недостаточном освещении (либо малой светочувствительности

видеоаппаратуры) пиксели как зрачка, так и радужки имеют нулевую яркость и неотличимы друг от друга. Если источник освещения расположен неудачно, на поверхность глазного яблока падают тени от носа, надбровных дуг и скул. В связи с этим источник освещения следует размещать по возможности ближе, корректируя его яркость, чтобы не ослеплять исследуемый глаз. Улучшение условий съемки существенно повышает точность системы.

Введение в систему дополнительной априорной информации также позволяет увеличить ее производительность. Если в кадр, помимо глаза, попадает часть лица, нет смысла проводить бинаризацию и поиск контуров в районе щеки или лба. Поэтому в системе должна присутствовать возможность ограничить регион, в котором следует искать зрачок (иначе поиск будет вестись по всему кадру). Также знание интервалов яркости пикселей зрачка для данного освещения и специфики используемых видеокамер позволяет снизить количество перебираемых значений порога.

### **Выводы и рекомендации**

Прежде всего следует отметить, что на основе проведенных экспериментов обоснована принципиальная возможность создания экономичных систем eye-tracking'a с приемлемой для большинства задач точностью (~5%) на основе широкодоступного оборудования. Это позволяет существенно увеличить объемы практического применения таких технологий.

На данный момент современные системы eye-tracking'a чаще всего используются экспертами в области машинных интерфейсов или исследователями психофизиологических особенностей зрения. Остальные же специалисты, прежде всего в области работы с цифровым видео, в силу сложившихся стереотипов недооценивают потенциал использования систем eye-tracking. Однако предложенный в статье метод распознавания и проецирования движения зрачков может стимулировать развитие данной области, прежде всего в сфере локализации дефектов, возникающих при сжатии видеоизображения с потерей качества.

### **Библиографический список**

1. *Glenstrup A. J., Engell-Nielsen T.* Eye-controlled media: present and future state. University of Copenhagen DIKU, June 1995
2. *Ярбус А. Л.* Роль движений глаз в процессе зрения. М.: Наука, 1965, 195 с.
3. *Duchowski A. T.* Eye Tracking Methodology: Theory and Practice. Springer, 2007, 22

4. *Suzuki S., Abe K.* Topological structural analysis of digital binary images by border following. Computer Vision, Graphics and Image Processing 30, 1985

5. *Fitzgibbon A., Pilu M., Fisher R. B.* Direct Least Square Fitting of Ellipses. Pattern analysis and machine intelligence, vol. 21, no. 5, may 1999

### **Сведения об авторах**

**Малин Иван Константинович**, Московский авиационный институт (государственный технический университет)

студент шестого курса. [ivan.malin@gmail.com](mailto:ivan.malin@gmail.com)

**Крапивенко Андрей Викторович**, Московский авиационный институт (государственный технический университет)

доцент кафедры «Вычислительная математика и программирование», к.ф.-м.н. 364-75-58, [kav@elias.ru](mailto:kav@elias.ru)