

УДК 57.087

Автоматизированное определение направления взгляда по видеоизображению лица

*Пилипенко М.Н., студент
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Биомедицинские технические системы»*

*Научный руководитель: Латышева Е.Ю., ассистент
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана
bmt-1@bmstu.ru*

Зрение – орган чувств, воспринимающий до 70 % всей информации об окружающем мире. В повседневной жизни человек постоянно просматривает и анализирует пространство, выделяя полезную информацию. Предмет восприятия меняется дважды в секунду, таким образом, изменяется и направление взгляда.

Изучение природы движения глаз и изменения направления взгляда насчитывает не одно столетие. Проблема изучалась крупнейшими естествоиспытателями с точки зрения механики (И. Кеплером, И. Ньютоном, И.М. Сеченовым), а также в рамках различных психологических течений: физиологической психологии, когнитивной психологии, деятельностного подхода и др. Связь движения глаз и процесса познания и восприятия произведений искусства отмечали Л. да Винчи, О. Роден. Одним из первых исследователей, систематизировавших знания о движениях глаз, был А.Л. Ярбус (1965). Им отмечено существование особых маршрутов осмотра изображения лица. Согласно полученным им окулограммам (отображениям изменения направления взгляда), движения глаз носят циклический характер: взор перемещается по одним и тем же деталям изображения в одной и той же последовательности.

Характер перемещения глаз выражает динамику внимания, мышления, воображения, представления, направленность говорит об интересе, намерении, отношении человека. Поэтому системы, которые используют информацию об изменении направления взгляда, могут быть крайне разнообразны:

1. Системы диагностики состояния человека:
 - определение психоэмоционального (напряженного) состояния для выявления асоциального поведения;

- контроль состояния операторов, выполняющих работу в тяжелых условиях труда.
- 2. Системы управления:
 - роботами (в частности, манипулятором типа «мышь»).
- 3. Системы для проведения маркетинговых исследований:
 - определение заинтересованности в билбордах;
 - определение качества рекламы в Интернете.
- 4. Тренажеры для обучения (например, моделирование управлением самолета).

При разработке системы, в которой оценивается направление взгляда, первоначально необходимо разработать требования, связанные с характеристиками движения глаза.

Окуломоторика – это двигательная активность глаз, которая характеризуется многообразием видов и переходов. Глаз человека осуществляет движения благодаря шести наружным мышцам, закрепленным в глазной впадине. Центр вращения глазного яблока расположен внутри глаза на зрительной оси. Расстояние между вершиной роговицы и центром вращения глаза приблизительно равно 13,5 мм.

Все движения поддаются классификации в зависимости от амплитуды, скорости, траектории, выполняемой функции движений: тремор, дрейф, микро- и макросаккады, прослеживающие, вергентные, торзионные движения и нистагм. В то же время все движения разделяются на микро- и макродвижения. Микродвижения связаны преимущественно с сохранением местоположения глаз, макродвижения – с изменением. Подробные характеристики различных видов движений представлены в таблице 1.

Классификация движений глаз

	Движение	Амплитуда	Частота	Длительность	Скорость	Примечание
	1	2	3	4	5	6
Микродвижения	Тремор	20-40 "	До 250-270 Гц	–	–	Естественный двигательный фон окулоmotorной активности, не поддается произвольному контролю.
	Дрейф	–	–	30-5000 с	0-40 °/с	Медленное, плавное перемещение глаза, прерываемое микроскачками. При фиксации на дрейф приходится 97 % времени. Дрейф создает условия для приема и переработки оптической информации.
	Микросаккады	2-50 '	–	10-20 мс	3-12 °/с	Быстрые движения, появляются во время фиксаций.
Макро-движения	Макросаккады	От 40-50' до 50-60 °, в естеств. усл. 20 °	2-3 Гц	при 20° и максимальной скорости – 70 мс	Максимальная скорость при 20° – 450 °/с	Резкие изменения позиции глаза. Возникают при смене точек фиксации (во время рассматривания, поиска объекта).

	1	2	3	4	5	6
Макродвижения	Прослеживающие движения	$\pm 60^\circ$ по горизонтали $\pm 40^\circ$ по вертикали	–	Начинаются с задержкой 150-200 мс от начала движения объекта Кончаются с задержкой 300 мс после остановки объекта	$5 \text{ } ^\circ/\text{с} - 12 \text{ } ^\circ/\text{с}$	Возникают при движении объекта в поле зрения. Обеспечивают сохранение изображения фиксируемого объекта в зоне наилучшего видения.
	Вергентные движения	–	–	Начинаются с задержкой 150-200 мс после появления объекта и продолжаются около 200 мс	–	Сведение или разведение оптических осей глаз.
	Торзионные (ротационные) движения	До 10°	–	–	–	Вращательные движения для компенсации наклонов головы.
	Нистагм	–	–	–	–	Чередование саккад и плавных прослеживающих движений для компенсации нарушений зрительной и вестибулярной систем или сохранения эффективности восприятия элементов движущейся среды.

Разработка систем, использующих информацию о направлении взгляда, предполагает определение движений глаз, связанных с познавательной деятельностью человека. Такими движениями являются макродвижения, а именно, прослеживающие движения. Для них характерны амплитуды $\pm 60^\circ$ по горизонтали и $\pm 40^\circ$ по вертикали, скорости $5^\circ/\text{с} - 12^\circ/\text{с}$. При этом длительность движений зависит от времени прослеживания объекта. В то же время при осуществлении просмотра объекта происходят фиксации с макросаккадами, амплитуда которых в нормальных условиях составляет 20° , а длительность – 70 мс.

В настоящее время существуют различные методики измерения и оценки характеристик движений глаз человека. Чаще всего используется видеоокулография и электроокулография, в лабораторных условиях распространены фотооптический, фотоэлектрический, электромагнитный методы (Табл. 2).

Таблица 2

Классификация методов измерения движений глаз по принципу действия

Метод	Измеряемые движения	Обеспечиваемая точность	Взаимодействие с человеком	Принцип
Видеосъемка (видеоокулография)	макродвижения	$1-2^\circ$	бесконтактное	Видеосъемка, покадровый анализ, наложение траекторий на экспонируемый объект
Электроокулография	макродвижения	$1,5^\circ$	контактное	Перемещение глаза сопровождается изменением разности потенциалом между роговицей и сетчаткой, а следовательно, и изменением потенциала в тканях, прилегающих к глазницам
Фотооптический метод	все виды движений	$2-5''$	контактное	Узкий пучок света, направленный на глазное яблоко, отражается от установленного на нем миниатюрного зеркала и поступает на вход фоторегистрирующего устройства

Фото-электрический метод	макродвижения	1 °	бесконтактное	Преобразование отраженного от роговицы пучка света в электрический сигнал. На фотоэлектрический датчик проецируется изображение переднего участка глаза, имеющего резкий перепад отражающей способности. При перемещении глаз меняется количество отраженного света, а следовательно, величина фототока или фотосопротивления
Электромагнитный метод	микродвижения, макродвижения	20-30 '	контактное	Изменение напряженности ЭМП при изменении расстояния между излучателем и приемником. Поворот глаз (на которые крепятся присоски-катушки, создающие ЭМП) преобразуется в напряжение на приемных катушках

При разработке системы автоматизированного определения направления взгляда будем использовать самый простой с точки зрения технологии метод видеосъемки, который будет регистрировать макродвижения, отражающие отношение человека к наблюдаемому объекту или пространству, без использования контактных, вызывающих дополнительные неудобства средств, которые могут вносить искажения.

После определения характеристик регистрируемых движений необходимо разработать требования к камере, а именно требования к разрешающей способности и частоте кадров.

Разрешающая способность определяется исходя из размера различаемого смещения. С учетом амплитуд движений глаз $\pm 60^\circ$ по горизонтали и $\pm 40^\circ$ по вертикали и радиуса кривизны $R=13,5$ мм, допустим, необходимо различать движение в $\varphi = 10^\circ$.

$$x = R * \varphi = 13,5 \text{ мм} * \frac{2\pi * 10^0}{360^0} = 2,35 \text{ мм.}$$

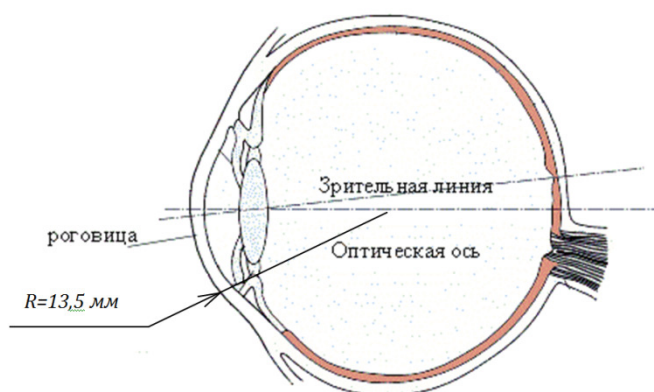


Рис. 1. Оптическая ось глаза

По теореме Котельникова для того, чтобы можно было различить элемент, необходимо минимум два отсчета. Таким образом, шаг дискретизации по обеим осям изображения равен:

$$T_{\text{дискр}} = \frac{2,35 \text{ мм}}{2 \text{ точки}} = 1,175 \frac{\text{мм}}{\text{точка}}$$

С учетом округления:

$$T_{\text{дискр}} = 1 \frac{\text{мм}}{\text{точка}}$$

Пространственная частота дискретизации равна:

$$\vartheta_{\text{дискр}} = \frac{1}{T} = 1 \frac{\text{точек}}{\text{мм}}$$

В пределах нормы размеры области глаз составляют: высота 2 см, ширина до 3 см. Возьмем коэффициент запаса $k=1,25$, тогда размеры области, которая должна быть отражена на изображении, равны:

$$1,25 * (20 \text{ мм} \times 30 \text{ мм}) = 25 \text{ мм} \times 37,5 \text{ мм}.$$

Тогда, исходя из полученного значения пространственной частоты дискретизации, получаем, что минимальные размеры области глаз на изображении должны быть равны:

$$(25 \text{ мм} \times 37,5 \text{ мм}) * 1 \frac{\text{точек}}{\text{мм}} = 25 \text{ точек} \times 37,5 \text{ точек}.$$

В системе оценки психоэмоционального состояния по динамике изображений лица используются цифровые изображения, поэтому отсчетами являются пиксели, т.е. минимальные размеры глаз на изображении равны 25 пикс. \times 38 пикс.

При разработке требований к частоте кадров видеоизображения устанавливается минимальная частота кадров, обусловленная длительностью 70 мс движения глаз. Таким образом, максимальная частота микровыражений равна:

$$f_{\text{глаз}} = \frac{1}{T} = \frac{1}{0,070} = 14,3 \text{ Гц}.$$

По теореме Котельникова частота дискретизации должна быть больше максимальной частоты в два или более раза:

$$f_{min} = 2 * f_{глаз} = 28,6 \text{ Гц},$$

т.е. минимальная частота кадров видеоизображения с учетом округления равна 30 кадров/с.

В дальнейшем при разработке макета автоматизированной системы необходимо также учесть максимально возможную частоту кадров, так как в случае ее превышения информация не будет успевать обрабатываться и возникнет временная растущая задержка. Максимальная частота кадров обуславливается скоростью передачи и обработки данных и может быть рассчитана только при определенных выбранных технических средствах системы.

Алгоритмы определения направления взгляда

Существуют два основных подхода к определению направления взгляда по видеоизображению: с использованием опорных точек, с использованием текстурных признаков лица (рис. 2). Опорные точки могут быть получены с помощью калибровки, рассчитаны относительно геометрических признаков лица или блика, создаваемого инфракрасным датчиком.



Рис. 2. Классификация алгоритмов определения направления взгляда

При разработке системы автоматизированного определения направления взгляда был разработан алгоритм определения направления взгляда с использованием опорных точек уголков глаз (Рис. 3). При определении опорных точек использовались библиотеки OpenCV, EMGU [3]. Для нахождения центра зрачка использовался алгоритм, описанный в [4].

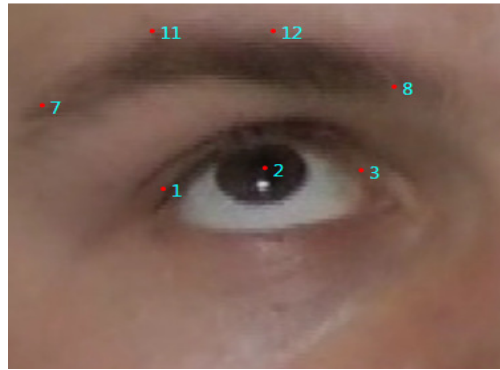


Рис. 3. Расстановка опорных точек

Направление взгляда определялось в двух системах координат: декартовой и полярной – с целью выявления наиболее оптимальной.

В декартовой системе координат направление взгляда рассчитывается относительно точек уголков глаза (Рис. 4).

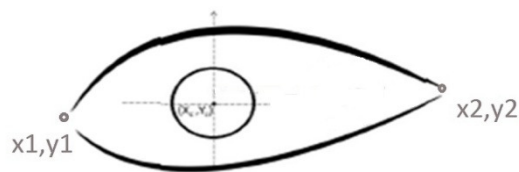


Рис. 4. Определение направления взгляда

$$e = \frac{x_i - x_1}{x_2 - x_i}$$

$$s = \frac{y}{y_1}$$

где (x_i, y_i) – рассчитанные координаты центра глаз, (x_1, y_1) – внутренний угол глаза, (x_2, y_2) – внешний угол глаза.

$$\text{Горизонтальное состояние} = \begin{cases} \text{вправо, если } e \geq 1.35, \\ \text{прямо, если } 0,65 < e < 1.35, \\ \text{влево, если } e \leq 0,65. \end{cases}$$

Соответственно, определяется 3 возможных направления взгляда: прямо, влево, вправо.

Для тестирования использовались 142 изображения испытуемого 1 и 158 изображений испытуемого 2 с тремя различными направлениями взгляда. Результаты тестирования работы алгоритма приведены в таблицах 3, 4.

Таблица 3

Автоматическое определение направления взгляда по изображениям испытуемого 1 в декартовой системе координат

Определено автомати- чески Истинное	прямо	влево	вправо	Количество изображений с данным направлением
прямо	45	0	1	46
влево	3	43	0	48
вправо	3	0	43	48

Таблица 4

Автоматическое определение направления взгляда по изображениям испытуемого 2 в декартовой системе координат

Определено автомати- чески Истинное	прямо	влево	вправо	Количество изображений с данным направлением
прямо	35	0	22	57
влево	24	16	0	40
вправо	17	0	44	61

Вероятность правильного определения направления взгляда по изображениям испытуемого 1 составляет 95,1 %, в то время как для 2-ого – 60,1 %. Снижение качества распознавания объясняется индивидуальными особенностями строения лица, а также низкой точностью расстановки точек уголков глаза.

Использование полярной системы координат обосновано анатомией глаза: глазное яблоко имеет сферообразную форму. В качестве центра полярной системы координат берется пересечение прямых, образованными точками 1, 7 и 3, 8 (Рис. 2, Рис. 4).

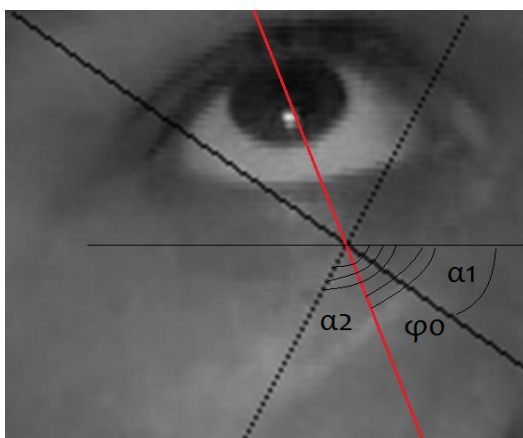


Рис. 5. Нахождение центра полярной системы координат

$$\text{Горизонтальное состояние} = \begin{cases} \text{вправо, если } \frac{\varphi_0 - \alpha_1}{\alpha_2 - \alpha_1} < 0,3, \\ \text{прямо, если } 0,3 \leq \frac{\varphi_0 - \alpha_1}{\alpha_2 - \alpha_1} < 0,7, \\ \text{влево, если } \frac{\varphi_0 - \alpha_1}{\alpha_2 - \alpha_1} \geq 0,7. \end{cases}$$

Для тестирования использовались 142 изображений испытуемого 1 и 158 изображений испытуемого 2 с тремя различными направлениями взгляда (Табл. 5, Табл. 6).

Таблица 5

Автоматическое определение направления взгляда по изображениям испытуемого 1 в полярной системе координат

Истинное \ Определенное автоматически	прямо	влево	вправо	Количество изображений с данным направлением
	прямо	36	0	
влево	7	39	2	48
вправо	8	1	39	48

Автоматическое определение направления взгляда по изображениям испытуемого
2 в полярной системе координат

Определенное автомати- чески Истинное	прямо	влево	вправо	Количество изображений с данным направлением
прямо	35	0	22	57
влево	24	16	0	40
вправо	17	0	44	61

Вероятность правильного определения направления взгляда по изображениям испытуемого 1 составляет 80,3 %, испытуемого 2 – 78,4 %.

Выводы

Разработка автоматизированной системы определения направления взгляда предполагает определение движений глаз, связанных с познавательной деятельностью человека, такими являются макродвижения, а именно, прослеживающие движения и фиксации. Проведен литературный анализ существующих методик регистрации глаз. Исходя из параметров движений, выбрана методика и сформулированы требования к блоку регистрации видеоизображения.

Разработан и протестирован алгоритм определения направления взгляда по контрольным точкам в декартовой и полярной системах координат. При переходе из декартовой системы координат в полярную вероятность распознавания направления взгляда по изображениям испытуемого 1 снизилась на 14,8 %, испытуемого 2 – увеличилась на 18,3 %. Это можно объяснить индивидуальными особенностями испытуемых.

В дальнейшем планируется увеличить размер тестовой выборки с целью оценки стабильности вероятностных характеристик, а также использовать другие алгоритмы определения направления взгляда для достижения лучших результатов автоматизированного определения направления взгляда.

Список литературы

1. Ярбус А.Л. Роль движений глаз в процессе зрения. М.: Наука, 1965. 166 с.

2. Барабанщиков В.А. Окуломоторные структуры восприятия. М.: Институт психологии РАН, 1997. 384 с.
3. OpenCV image processing library. Available at: http://www.emgu.com/wiki/index.php/Main_Page, accessed 01.09.2014.
4. Пилипенко М.Н. Определение центров зрачков на видеоизображении // Молодежный научно-технический вестник. МГТУ им. Н.Э. Баумана. Электрон. журн. 2013. № 6. Режим доступа: <http://sntbul.bmstu.ru/doc/586663.html> (дата обращения: 01.08.2014).