

Отслеживание направления взгляда в реальном времени без использования специальной видеоаппаратуры

Иван Малин

Факультет прикладной математики и физики
Московский Авиационный Институт, Москва, Россия
ivan.malin@gmail.ru

Аннотация

В данной статье рассматривается задача отслеживания направления взгляда в реальном времени с использованием видеокамер бытового уровня в видимом диапазоне. Приводится анализ и сравнение метода Даугмана и преобразования Хаффа для обнаружения радужки. Предлагается способ определения точки взгляда по положению центров радужек с использованием однослойного перцептрона. Приводятся оценки точности полученной системы.

Ключевые слова: eye-tracking, обнаружение радужки, метод Даугмана, метод Хаффа, перцептрон

1. ВВЕДЕНИЕ

В настоящее время системы отслеживания направления взгляда используются для решения широкого круга задач. К сферам применения относятся маркетинговые, медицинские и психологические исследования, а также построение бесконтактных интерфейсов для человеко-машинного взаимодействия. В работе (Hansen & Pese, 2005) приводится исчерпывающий обзор современных подходов к отслеживанию направления взгляда. Большинство систем, используемых в настоящее время в исследовательских и коммерческих целях, требуют дополнительного оборудования, например, активной подсветки, ИК- и стереокамер. Таким образом, задача отслеживания направления взгляда в реальном времени с использованием одной видеокамеры, работающей в видимом диапазоне, остается актуальной. Целью данной работы является исследование возможности построения системы, решающей задачу определения области на экране, в которую смотрит пользователь, и отвечающей следующим требованиям:

- Отсутствует необходимость использования специализированной аппаратуры
- Функционирование осуществляется в реальном времени
- Процедура калибровки проста, отсутствует необходимость ввода данных о камере или сцене, таких как фокусное расстояние камеры, взаимного расположение камеры, пользователя и экрана и т. д.

Общая идея всех методов определения точки взгляда по видеопоследовательности глаз заключается в выделении на изображении некоторого вектора признаков и построения отображения пространства признаков на координаты рассматриваемого объекта, чаще всего - плоского экрана. В зависимости от конкретной реализации, используемыми признаками могут быть координаты центров радужек и зрачков, отражений на роговице от активной подсветки и пр. (Ji & Yang, 2001), (Ebisawa & Satoh, 1993), (Morimoto, Koons,

Amir, & Flickner, 2000). Многомерным вектором признаков может являться само изображение глаза, отмасштабированное до некоторого стандартного размера (Xu & Machin, 1998). В данной работе в качестве признаков используются координаты центров радужек глаз, так как они не требуют активной подсветки и могут быть эффективно выделены при обычных условиях съемки. В целях быстродействия радужки ищутся как окружности, несмотря на то, что в общем случае их изображение представляет эллипс. Раздел 2 посвящен сравнению и исследованию применимости двух наиболее популярных методов обнаружения окружностей, в разделе 3 описывается построенная система и результаты измерений её точности.

2. МЕТОДЫ ОПРЕДЕЛЕНИЯ ПОЛОЖЕНИЯ РАДУЖКИ

2.1 Оператор Даугмана и преобразование Хаффа

Рассмотрим и сравним два метода поиска окружностей на полутоновом изображении. Первый из них – метод Даугмана (Daugman, 2004), ставший стандартом для определения границ радужки в задачах биометрии. Идея метода заключается в поиске наиболее выраженного кругового контура. Для этого ищутся параметры окружности (x_0, y_0, r) – координаты центра и радиус, в которых достигается максимум оператора Даугмана:

$$\max_{x_0, y_0, r} \left| \frac{1}{2\pi r} G_\sigma(r) * \frac{\partial}{\partial r} \oint_{(x_0, y_0, r)} I(x, y) ds \right| \quad (1)$$

Символ * означает свертку, а $G_\sigma(r)$ является сглаживающей функцией, в данном случае гауссианом со среднеквадратическим отклонением σ . Т.к.

$$\frac{\partial}{\partial r} \oint_{(x_0, y_0, r)} I(x, y) ds = \oint_{(x_0, y_0, r)} \left(\frac{\partial I(x, y)}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial I(x, y)}{\partial y} \frac{\partial y}{\partial r} \right) ds$$

(1) можно переписать в ином виде:

$$\max_{x_0, y_0, r} \left| \frac{1}{2\pi r} G_\sigma(r) * \oint_{(x_0, y_0, r)} \langle \nabla I(x, y), c \rangle ds \right| \quad (2)$$

Здесь $\langle \nabla I(x, y), c \rangle$ обозначает скалярное произведение вектора градиента яркости изображения $\nabla I(x, y)$ и единичного вектора c , направленного от центра окружности x_0, y_0 к текущей точке интегрирования (положению элемента ds).

Таким образом, для дискретного цифрового изображения процесс поиска окружности методом Даугмана сводится к вычислению градиента изображения и его свертке с маской

радиально ориентированных единичных векторов. Точка, в которой достигается максимальное значение свертки, является центром окружности. Применительно к задаче поиска радужки следует учитывать, что она имеет меньшую яркость, чем белки глаз, поэтому в операторе (2) нужно убрать скобки модуля. Сложность такого алгоритма составляет $O(M \cdot N \cdot R^2)$, где $R = r_{max} - r_{min}$ – величина диапазона возможных радиусов искомой окружности.

Другим общепризнанным методом для обнаружения окружностей является преобразование Хаффа и его различные модификации, в частности, градиентное преобразование. Эта группа методов успешно используется в том числе для обнаружения радужки (Тоенниес, 2002). Ниже описана модификация, использованная в данной работе при проведении сравнения с методом Даугмана.

Для определения положения и радиуса окружности используется трехмерный аккумулятор $A(x_0, y_0, r)$. Каждая точка изображения $I(x, y)$ добавляет вес $\frac{1}{r} |\nabla I(x, y)|$ к значению аккумулятора A в точке $(x - \frac{I_x(x, y)}{|\nabla I(x, y)|} r, y - \frac{I_y(x, y)}{|\nabla I(x, y)|} r, r)$ для всех значений r в устанавливаемых границах $[r_{min}, r_{max}]$. Иными словами, каждая точка изображения голосует с весом $\frac{1}{r} |\nabla I(x, y)|$ за расположение центра окружности на расстоянии r от нее в направлении антиградиента. Алгоритмическая сложность процедуры голосования составляет $O(M \cdot N \cdot R)$.

Следует отметить, что в большинстве работ, использующих круговое преобразование Хаффа, отсутствует коэффициент $\frac{1}{r}$ при вычислении веса, с которым голосует точка изображения. Однако в этом случае большим окружностям будут соответствовать большие значения в аккумуляторе, т.к. количество голосующих пикселей на границе окружности линейно зависит от ее радиуса. Введение коэффициента $\frac{1}{r}$ осуществляет необходимую нормировку (как и в операторе Даугмана).

Также необходимым этапом является постобработка аккумулятора – сглаживание значений путем свертки с трехмерным гауссианом. Это связано с тем, что малая погрешность в оценке направления градиента в точке может дать значительное смещение координат центра, за который эта точка голосует, если радиус достаточно велик. Учитывая сепарабельность гауссовского фильтра, свертку можно осуществить за $O(M \cdot N \cdot R)$ операций. Таким образом, алгоритмическая сложность данного метода составляет также $O(M \cdot N \cdot R)$.

2.2 Сравнение методов

Описанная реализация метода Даугмана проигрывает методу Хаффа в скорости работы, однако является более устойчивой к ошибкам определения направлений градиента – в силу интегрального характера оператора (2).

Для тестирования использовались синтезированные изображения окружностей. В каждой точке изображения к вычисленному с помощью оператора Собеля градиенту добавлялся случайно ориентированный вектор случайной длины $r \sim |N(0; \sigma)|$. График на рисунке 1 показывает зависимость среднеквадратического отклонения координат центра окружности в зависимости от параметра σ/G_{max} , где G_{max} – максимальное значение модуля градиента исходного синтезированного изображения. На рисунке 2 показано сравнение времени работы методов в зависимости от размера изображения. Для тестирования использовались

синтезированные изображения с отношением сторон 1:1 с изображениями окружностей, радиус которых составлял от 0.2 до 0.3 стороны изображения.

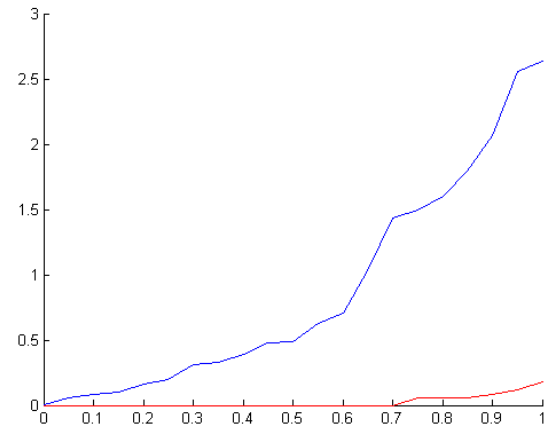


Рисунок 1. Зависимость среднеквадратической ошибки определения положения центра окружности (в пикселях) от σ/G_{max} . Красным обозначен метод Даугмана, синим – Хаффа.

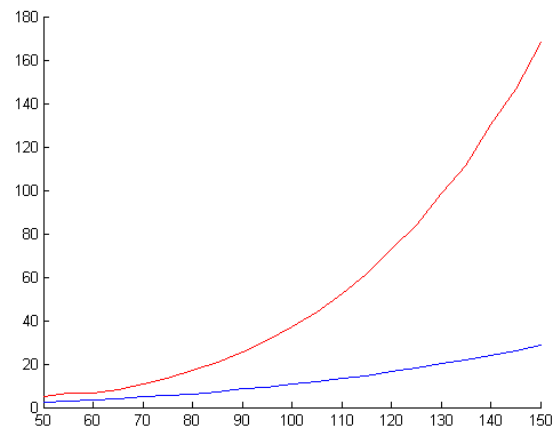


Рисунок 2. Зависимость времени работы методов в миллисекундах от линейного размера изображения в пикселях. Красным обозначен метод Даугмана, синим – Хаффа.

3. ПРЕДЛАГАЕМАЯ СИСТЕМА

3.1 Общая структура

Как было указано во введении, одной из целей данной работы являлась разработка системы реального времени. В связи с этим за основу алгоритма нахождения центра радужки был взят градиентный метод Хаффа. Следует учитывать, что поиск окружности на всем кадре затратен по времени, а также ведет к большей вероятности ошибок детектирования. Поэтому зона поиска радужек предварительно сужается путем нахождения региона глаз с помощью метода Виолы-Джонса (Viola & Jones, 2001). Также для исследования потенциальной возможности обеспечения пользователю свободы движения головой, осуществлялся поиск налобного маркера круглой формы (см. рисунок 3).

Следующим этапом является построение отображения $M(x) = y$, где x – вектор выделенных на кадре признаков, а $y \in R_2$ – координаты предполагаемой точки взгляда на экране. В данном случае $x \in R_6$ и представляет собой набор координат центров радужек обоих глаз и налобного маркера на кадре. В качестве модели отображения M был выбран однослойный перцептрон. Обучение перцептрона осуществляется методом обратного распространения ошибки (Хайкин, 2006). Обучающей выборкой являются пары (x_i, y_i) , полученные в результате процесса калибровки, когда пользователю предлагается последовательно посмотреть на несколько меток, расположенных в узлах регулярной сетки на экране.

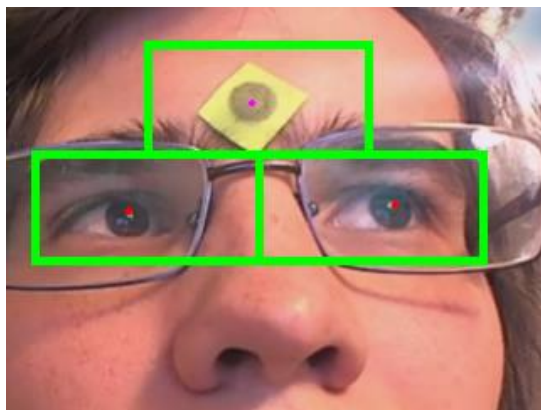


Рисунок 3. Найденные области поиска окружностей и их полученные центры.

3.2 Результаты тестирования

Процедура тестирования была проведена для четырех пользователей, среди которых были лица как с темной, так и со светлой радужкой. Эксперимент проводился в двух вариантах – в условиях естественного дневного света и освещения лампой накаливания. Рассматривалось два случая движений головы – без фиксации (однако с указанием пользователю по возможности держать ее неподвижно) и с фиксацией подбородка на подставке. Расстояние от глаз до экрана составляло около 0,45 м, сам экран имел размеры около 0,41 м по горизонтали и 0,26 м по вертикали. В процессе тестирования пользователь последовательно смотрел на 25 меток на экране, расположенных в узлах регулярной прямоугольной сетки размером 5 на 5, целиком покрывающей экран. В дальнейшем 13 точек с четными номерами были использованы в качестве обучающей выборки для построения функции отображения координат, а оставшиеся 12 – в качестве тестирующей выборки для оценки точности системы. Съемка осуществлялась веб-камерой Logitech HD Pro Webcam C920.

Среднеквадратическая ошибка определения координат точки взгляда приведена в таблице 1. Визуализация вычисленных точек взгляда представлена на рисунках 4 и 5. Таким образом, среднеквадратическая ошибка на тестирующей выборке в наилучшем случае составила около 30 пикселей по вертикали и 50 по горизонтали на мониторе с разрешением 1440 на 900.

Таблица 1. Среднеквадратическая ошибка определения координат точки взгляда (в долях от линейных размеров экрана).

	Обучающая выборка	Тестирующая выборка
Без фиксации, без налобного маркера	0.13	0.20
Без фиксации головы, с налобным маркером	0.021	0.036
С фиксацией головы	0.017	0.033

	σ_x	σ_y	σ_x	σ_y
Без фиксации, без налобного маркера	0.13	0.07	0.20	0.24
Без фиксации головы, с налобным маркером	0.021	0.027	0.036	0.039
С фиксацией головы	0.017	0.025	0.033	0.035

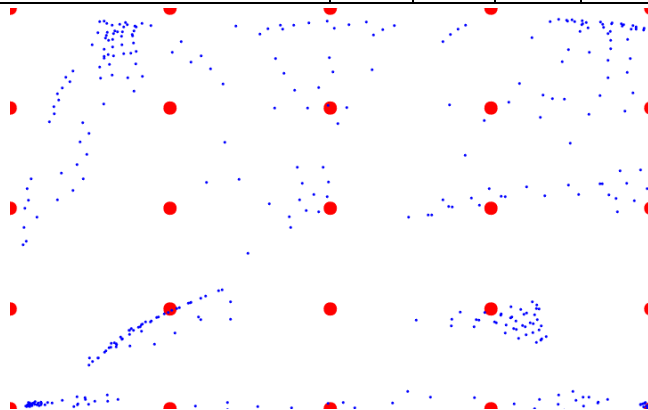


Рисунок 4. Пример некорректного отображения координат без использования налобного маркера и фиксации головы.

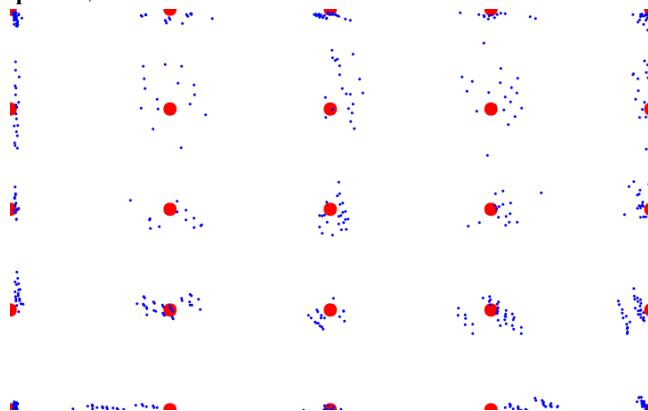


Рисунок 5. Пример отображения координат при использовании налобного маркера без фиксации головы.

4. ЗАКЛЮЧЕНИЕ

Предлагаемый метод позволяет создать систему отслеживания направления взгляда, работающую в режиме реального времени и использующую только широкодоступную видеоаппаратуру бытового уровня. Использование одного налобного маркера как источника дополнительной информации о положении головы позволяет обеспечить некоторую малую свободу движений пользователя. Однако для обеспечения свободы амплитудных движений требуется дополнительная информация. Вопрос об оценке требуемого количества маркеров и калибровочных точек для обеспечения свободы движений головы, а также о структуре и размерах слоев перцептрона, также остается открытым.

5. ССЫЛКИ

- Daugman, J. (January 2004 г.). How Iris Recognition Works. *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY*, 14(1).
- Ebisawa, Y., & Satoh, S. (1993). Effectiveness of pupil area detection technique using two light sources and image difference method. *5th Annual Int. Conf. of the IEEE Eng. in Medicine and Biology Society*, (стр. 1268–1269).
- Ji, Q., & Yang, X. (2001). Real time visual cues extraction for monitoring driver vigilance. *Comput Sci 2095*, 107.
- Morimoto, C., Koons, D., Amir, A., & Flickner, M. (2000). Pupil detection and tracking using multiple light sources. *IVC 18 (4)*, (стр. 331–335).
- Toennies, K. (2002). Feasibility of Hough-transform-based iris localisation for real-time-application. *Pattern Recognition, 2002. Proceedings on 16th International Conference, 2*, стр. 1053 - 1056.
- Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *ACCEPTED CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION 2001*.
- Xu, L., & Machin, D. (1998). A novel approach to real-time non-intrusive gaze finding. *British Machine Vision Conference*.
- Хайкин, С. (2006). *Нейронные сети*. Москва: Вильямс.