



УДК 004.85

Тестирование алгоритма тернарной классификации окуломоторных событий (фиксации, саккады, smooth pursuit) I-BDT (Bayesian Decision Theory)

Жердев И.Ю.*

Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ),
г. Москва, Российская Федерация (аффилированный сотрудник)
ORCID: <https://orcid.org/0000-0001-6810-9297>
e-mail: ivan866@mail.ru

Рассмотрен байесовский классификатор движений глаз I-BDT [1], выделяющий 3 типа событий: фиксации, саккады, плавные следящие движения. Приведены показатели классификации по сравнению с известным алгоритмом детекции I-DT. Согласование по ряду фиксаций – 79 %, саккад – 56 %. Сравнение произведено на наборе данных общей длительностью 10,9 ч. Приведен пример каскадного фильтра, накладывающего пороговые ограничения на свойства полученных событий. Показаны особенности следящих движений глаз при классификации.

Ключевые слова: smooth pursuit, вероятностный, классификатор, детекция, следящие движения.

Для цитаты:

Жердев И.Ю. Тестирование алгоритма тернарной классификации окуломоторных событий (фиксации, саккады, smooth pursuit) I-BDT (Bayesian Decision Theory) // Моделирование и анализ данных. 2020. Том 10. № 2. С. 74–92. DOI:10.17759/mda.2020100206

Финансирование. Исследование выполнено при финансовой поддержке Российского научного фонда (РНФ) в рамках научного проекта № 18–18–00350.

***Жердев Иван Юрьевич**, аффилированный сотрудник, Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0001-6810-9297>, e-mail: ivan866@mail.ru



1. ВВЕДЕНИЕ

В современных условиях достаточно остро стоит проблема различий в результатах выдаваемых наборов окуломоторных событий между различными алгоритмами классификации сигнала (иногда их называют алгоритмами детекции). Поскольку случаются различия и в версиях одного и того же алгоритма из-за неточностей в трактовке тех или иных вычислений на этапах, варианты реализации могут по-разному классифицировать или численно оценивать разные виды параметров / характеристик алгоритма. Некоторые алгоритмы получают на вход калиброванные данные о траектории движения глаза, некоторым это не требуется – они учитывают только соотношения между отсчетами в сигнале, а не абсолютные величины (такovým является рассмотренный в этой статье).

Пороговые алгоритмы требуют точного указания величин, которые ограничивают теоретическое представление о характеристиках окуломоторных событий – фиксаций, саккад, постсаккадических осцилляций. Сравнить события, детектированные таким алгоритмом с исследованиями, проведенными с различными моделями айтрекеров, как минимум, рискованно [2]. Трекеры от различных производителей, формально используя один и тот же принцип роговичного блика, в основе имеют существенно отличные способы интерпретации этого сигнала [3], различный уровень шумов, дисперсию и амплитуду тремора, испытывают различное влияние физических условий дизайна эксперимента. Указанные особенности в записи сигналов сильно проявляются именно в динамике, на тех участках траектории движения глаза, где есть острые углы поворота, или большие перепады скорости.

Так, один трекер может имплицитно сглаживать траекторию глаза вследствие низкого разрешения исходного видеосигнала инфракрасной камеры, а другой – завышать мгновенную скорость вследствие высокой частоты регистрации, или неточно интерпретировать сигнал при проецировании зрительной оси на стимул в моменты начала/завершения саккад (см. исследование Hooge et al. относительно айтрекеров SMI), что приводит к регистрации недостоверных траекторий. При сравнении событий, детектированных пороговыми алгоритмами между этими трекерами, может получиться, что у одного количество плавных прослеживающих движений значимо больше.

Причиной дискрепансов также могут стать линейные различия в положении испытуемого относительно демонстрационной поверхности. Формально, фиксируя испытуемого на заданном расстоянии от экрана монитора или проектора, вы можете затем вычислить угловое расстояние до стимула и угловую скорость. Однако, если в другом эксперименте расстояние заметно меньше (иногда это делают с целью увеличить максимально возможную амплитуду саккад при исследовании высокоамплитудных движений глаз), то угловые расстояния увеличатся, но не пропорционально вашим изменениям дизайна, т.к. глазодвигательная система имеет собственные пороги и наиболее распространенные / привычные для испытуемых количественные характеристики глазодвигательной активности.

Несмотря на достаточное количество уже существующих программных реализаций классификаторов для выделения прослеживающих движений глаз [1, 4–7],



выбор их остается нетривиальной задачей. Это тем более актуально, когда исследователю требуется проводить исследования с различным экспериментальным дизайном. Сопоставление данных классификации между несколькими наборами данных не всегда представляется корректным, и может приводить к неверным выводам, особенно если временные характеристики (длительность, частота появления, величина тремора, качество калибровки) между ними сильно различаются, либо когда алгоритм рассчитанный на детекцию только двух типов событий применяется на данные, содержащие переходные типы событий, например, постсаккадические осцилляции [6], прослеживания, нистагмы. Исходя из вышесказанного, к классификации *окуломоторных* событий следует относиться как некоей приближенной оценке «потока» данных, а не строгой последовательности четко разграниченных событий, имеющих формальные признаки.

Целью данной работы является поиск алгоритма, способного классифицировать, кроме фиксации и саккад, прослеживающие движения глаз (более известные как *smooth pursuit*). Характеристики такого типа событий точно не известны [7], однако нам представляется логичным, что у них должен быть один неизменный признак – наличие фазы разгона, и фазы поддержания скорости в течение не менее 80 мс (см. Рис. 6; Табл. 2.2), потому что на этой величине часто проводят границу между длительностью саккад и фиксации. Мы считаем, что на данный момент не существует теоретических представлений о том, какой длительности должно быть минимально возможное плавное прослеживание. Минимальная теоретическая длительность саккады равна 22 мс, фиксации – 80 мс, прослеживания – 60 мс [4, 8]. В то же время, экспериментально показано, что максимально физиологически возможная скорость следящих движений глаз находится в области 30 гр/с [7]. Более быстрые цели требуют периодических корректирующих саккад и разбивают плавное следящее движение на несколько эпизодов. Мы считаем, что валидное прослеживание должно обладать следующими параметрами: иметь характерное «плато» в профиле скорости и длительность не менее 180 мс (плавно нарастить скорость, затем сохранять ее в течение этого периода, затем так же плавно снизить скорость до нуля, после чего перейти либо в фиксацию, либо в саккаду). Скорость прослеживания не обязательно должна быть постоянной, тем более что нередки случаи, когда объект прослеживания движется с переменной скоростью [9]. Напротив, скорость движения глаза постоянно корректируется в процессе слежения, с целью привести сетчаточное положение объекта к оптимуму. В отсутствие отдельной движущейся цели появление плавных следящих движений сомнительно (во всяком случае, достаточно длительных), но возможно.

На Рис. 1 представлены эмпирические данные по зависимости ускорения глаза при прослеживании цели от ее скорости. Видно, что в диапазоне от 0 до ± 30 °/с данные согласуются с предложенной Kashyap et al. моделью, тогда как за пределами этого скоростного диапазона наблюдается физиологический предел возможностей глазодвигательной системы (ускорение выходит на «плато»).

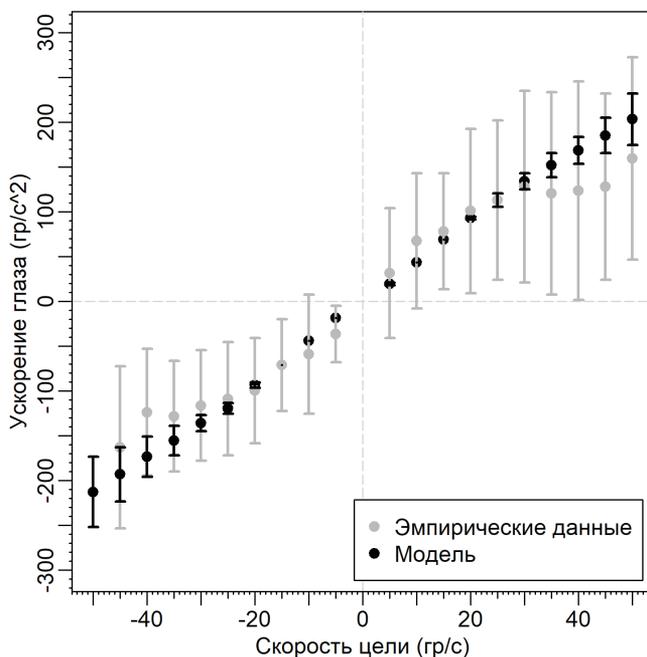


Рис. 1. Среднее ускорение глаза в зависимости от скорости движения цели в интервале от 80 до 180 мс после начала прослеживания. Усиками обозначено ст.откл. Эмпирические данные заимствованы из [10], данные модели – из [9]

Рассмотренный алгоритм классификации был сравнен (до и после фильтрации – см. далее) с классическим классификатором I-DT (реализация из утилиты *SMI Event Detector 3.0.18*) на наборе данных, включающем свободное рассматривание лица живого собеседника на видеомониторе. Записи в этом наборе не предполагают преобладание окуломоторных событий какого-то определенного класса выше обычного. Скорее, это стандартная окуломоторная активность в условиях, приближенных к натуральным. В эксперименте производилось 15 проб с каждым испытуемым (15 монологов разной направленности, отличия которых не имеют для данной статьи значения). Продолжительность каждой пробы – 1 мин. Средняя продолжительность записи – 18:42 мин. Общая длительность базы – 10:54:13 ч.

2. МЕТОДЫ

Классификатор I-BDT Данный классификатор [1] основан на теории Байеса, что означает, что он не получает на вход строго заданных величин, описывающих искомые детектированные события. Вместо этого он анализирует весь массив чисел (скорость движения), адаптирует на его основании внутренние гиперпараметры и затем разделяет поданные данные на три кластера (или на два – данный параметр задается), так, чтобы скорости в каждом кластере соответствовали представлению



о наиболее вероятном их распределении согласно Байесовской теории. Деление массива скоростей при движении глаза на 3 кластера происходит согласно гауссовой смеси распределений – т.н. нормальный Байесовский классификатор (EM-алгоритм кластеризации, первый этап – расчет мат. ожидания, второй – подбор кластеров с целью максимизации подобия / maximum likelihood).

Алгоритм классификатора работает с одним каналом данных (монокулярно), по сырым координатам и не требует калибровки айтрекера, нечувствителен к шумам и абсолютным значениям мгновенной скорости или пространственной дисперсии данных. Алгоритм не имеет заданных оператором минимальных или каких-либо других характеристик по длительности, амплитуде или скорости. При необходимости, вся дополнительная фильтрация и удаление артефактов выполняется в уже классифицированных группах событий. Алгоритм не производит детекцию шума или морганий, но выделяет отдельный четвертый кластер, скоростные характеристики которого не описываются полученной гауссовой смесью (помечая его меткой «noise»). Детекция шумов недифференцирована: только шумы с существенно отличным скоростным профилем (например, полный покой) будут распознаны как «шумы». В состав пакета с классификатором (исходный код на MATLAB; наша версия на языке Python) входит набор эталонных данных (в данной статье не рассмотрен), содержащий как классические эпизоды следящих движений за целью, так и контрольный набор записей (30 Гц, 4 серии, 6 испытуемых).

Детектированные примеры следящих движений обычно представляют собой бесформенные циклические движения вокруг некоего центра внимания (Рис. 2, справа), либо короткие, порядка 120–200 мс, плавно заворачивающие, скользящие движения. Но, в отличие от фиксаций, они имеют значимо большую амплитуду (протяженность в пространстве) и среднюю скорость, чем фиксации, а также обладают некой плавностью движения (которую можно увидеть при воспроизведении трека в динамическом режиме), в отличие от тремороподобных, «дрожащих» движений фиксаций.

Алгоритм позволяет выделять низкоамплитудные следящие движения там, где пороговые алгоритмы неминуемо относят сэмплы к фиксациям, как якобы эпизод повышенной пространственной дисперсии, а пороговые по скорости алгоритмы – как участок более плавной скорости (причины регистрации таких участков обсуждались выше). Он позволяет работать даже на тех записях, где нет следящих движений в классическом представлении как событий длительностью не менее 60 мс, с достижением глазом постоянной скорости [4], совпадающей со скоростью движения цели, и которое может, но не обязано, иметь по ходу компенсирующие саккады, фиксирующие позицию глаза на цели.

Важной особенностью данного классификатора является его инерционность (во времени). Требуется некоторое время, чтобы перейти от классификации участка записи как фиксации к следующему участку, как следящему движению. Обычно алгоритм запаздывает по сравнению с эталонной разметкой на 4–8 сэмплов (нами производился визуальный анализ на наборе данных, поставляемых вместе с классификатором). Однако, этот же эффект является и достоинством алгоритма, о котором



уже было сказано выше: спонтанные кратковременные пики и шумы не нарушают верной классификации. Мы считаем, что этот факт важнее для обнаружения искомым событий в длинных записях, чем точное разграничение фиксации и следящих движений (саккады переходящие в слежение, равно как и прослеживания, поправляемые периодическими саккадами, при этом разграничиваются достаточно строго).

И еще одно свойство алгоритма: он лучше всего работает на тех записях, в которых действительно содержатся искомые события. Это обусловлено тем фактом, что классификация всего распределения скорости по сэмплам должна произойти по завершению работы алгоритма независимо от содержания записи (в отличие от пороговых алгоритмов, которые проводят поиск строго заданных значений дисперсии или скорости), поэтому есть в записи прослеживания или нет, алгоритм все равно выделит ту часть распределения, которая максимально правдоподобно описывается гауссовой смесью, и обозначит ее как прослеживания. Авторы предоставляют версию данного алгоритма [1], которая лишена такой особенности (может выдавать классификации, не содержащие ни одного прослеживания), в которой начальное вычисление дисперсии происходит на основе минимума в заданном окне (обычно берется 500 сэмплов из начала записи), но такая реализация становится намного менее чувствительна к кратковременным и нечетким прослеживаниям, а результаты классификации следящих движений не совпадают с экспертной разметкой. Если же нужна детекция именно следящих движений, то более поздняя версия алгоритма расценивает минимальную дисперсию как смешение гауссиан.

Данные Классификатор в поставке от авторов получает данные в формате «journal» айтрекера *Dikablis Pro* (открытый текстовый формат). Наша реализация на Python 3 считывает данные айтрекера производства *SMI* из формата ‘Samples.txt’ (также открытый текстовый формат). Конвертация между этими форматами тривиальна, но не полная. Однако, алгоритм задействует только некоторые поля данных, поэтому отсутствующие в одном формате поля заполняются при конвертации пустыми значениями.

Фильтрация событий Поскольку рассмотренный классификатор помимо искомым окуломоторных событий выделяет множество шумов (согласно описательной статистике, порядка 35 % длительности всех классифицированных событий – интервалы мелких, длительностью в 1–2 сэмпла, участки саккад и прослеживаний – являются паразитными включениями, артефактами работы алгоритма), и при этом в нем не предусмотрено задание каких-либо порогов для обозначения теоретических границ по длительности или амплитуде для фиксации и саккад, то мы дополнительно выполнили фильтрацию полученных событий по ряду параметров с целью удаления тех из них, которые не попадают в теоретический диапазон для каждого типа.

Для фильтрации, а также при количественном сравнении двух алгоритмов между собой, предварительно был рассчитан ряд свойств для каждого из событий (см. Рис. 2), выделенных первым (I-DT) и вторым (I-BDT) классификатором. Среди них: длительность, средняя величина зрачка, позиция начала и конца события, средняя и пиковая скорость, позиция на пике скорости, дистанция, амплитуда. Также для облака то-



чек каждого события, независимо от его типа, была построена выпуклая оболочка и вычислена ее площадь. Амплитуда рассчитана как самая длинная хорда выпуклой оболочки. Она же является диаметром окружности, описанной вокруг облака точек, принадлежащих событию. Дистанция – сумма расстояний между отсчетами данных (в порядке их регистрации) в составе каждого выделенного события (приводится в пикселях). Площадь выпуклой оболочки приводится в пикселях квадратных.

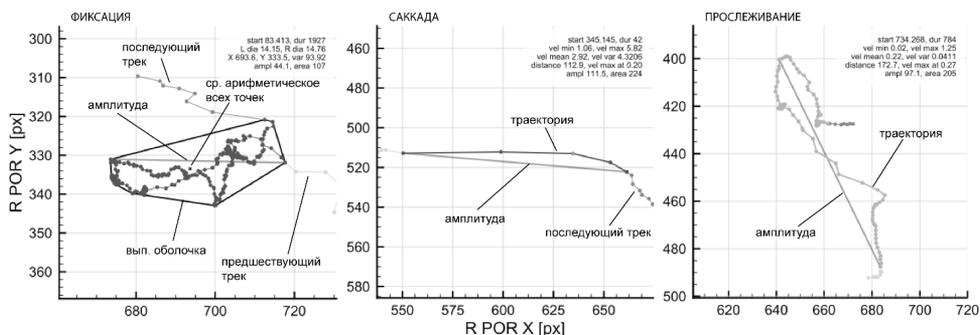


Рис. 2. Пример характеристик геометрических объектов, использованных при фильтрации невалидных событий

Слева – выпуклая оболочка, описанная вокруг облака точек фиксации длительностью 1927 мс. Посредине – траектория саккады длительностью 42 мс с нанесенной амплитудой и участками трека до и после саккады. Амплитуда события, и ее отношение к площади оболочки, позволяют корректно отфильтровать саккады от участков иужма (которые очень часто классифицируются пороговыми алгоритмами как последовательность саккад). Справа – один из многих вариантов траектории прослеживания, длительностью 784 мс.

Для фильтрации полученных на выходе классифицированных событий применялось логическое пересечение результатов всех фильтров в каскаде. Участки, на которых диаметр обоих зрачков равен 0, принимались за потерю сигнала и не классифицировались. Все статистические тесты и визуализации выполнялись в среде R 3.5.1. Предлагаемая батарея фильтров описывается следующими критериями (Табл. 1):

Таблица 1

Предлагаемые критерии фильтрации вывода классификатора I-BDT (N = 80)

	Фиксации		Саккады		Прослеживания	
	крит.	% удален	крит.	% удален	крит.	% удален
Длительность (мс)	≥ 80	65	> 18 < 80	65	≥ 80 < 2500	61
Ср. скорость (гр/с)			> 30	33	≤ 30	40
Мах скорость (гр/с)	≤ 30	23	< 800	21	< 400	8
Амплитуда (гр)	≤ 2.5	46	> 2.5 < 16	77	< 16	40



	Фиксации		Саккады		Прослеживания	
Дистанция-ампл. (пикс)			< 100	78		
Площадь вып. обол. (пикс кв)			ампл×60	62		
В комбинации, % от кол.-ва		72		92		67

3. РЕЗУЛЬТАТЫ КЛАССИФИКАЦИИ

В ходе визуального анализа нами не было найдено участков, соответствующих теоретическому представлению о выходе скорости движения глаза на плато, однако они были найдены данным классификатором. Средняя длительность прослеживания – 110 мс до фильтрации, 253 мс после фильтрации, при довольно большом ст. откл. 203 мс (Табл. 2.2). В этом проявляется нечеткая природа данного типа глаздвигательных событий на использованном нами наборе данных. Гистограмма распределения скоростей движения глаза (Рис. 3) показывает, что прослеживания занимают срединное положение по данному показателю между фиксациями и саккадами, но, в отличие от них, плавно распределены по всему диапазону. Это свидетельствует, что данный тип глаздвигательного события имеет широкий размах по скорости и гораздо более адаптивен, чем классическая фиксация или типичное скачкообразное движение.

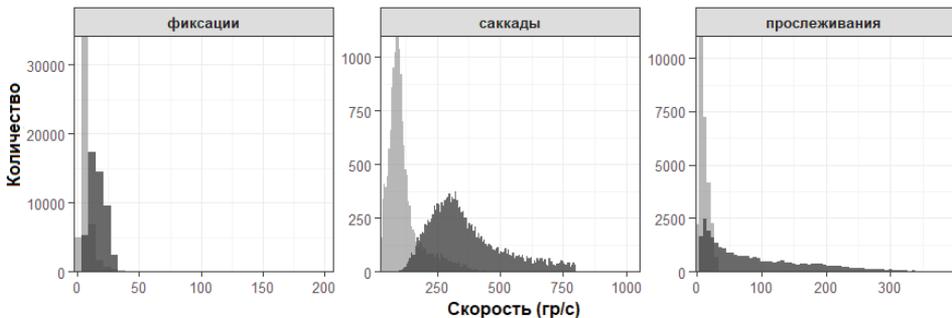


Рис. 3. Гистограмма распределения событий I-BDT по скоростям (после фильтрации).
Светлый тон – средняя скорость; темный тон – пиковая скорость

Всего в наборе данных классифицировано:

- 230132 событий суммарной длительностью 9:31:22 ч (I-DT; Табл. 2.1);
- 452500 событий суммарной длительностью 10:13:15 ч (классификатором I-BDT до фильтрации; Табл. 2.2 слева);
- 92387 событий суммарной длительностью 6:39:36 ч (классификатором I-BDT после фильтрации; Табл. 2.2 справа).

Табл. 2.1 показывает известную особенность алгоритма детекции I-DT – он центрирован на фиксациях. Корректными после классификации являются только показатели фиксаций. Длительность и скорость саккад выходит далеко за физиологически



возможные пределы человеческого глаза. Дистанция и амплитуда саккад также не может принадлежать к валидным скачкообразным движениям и, как правило, относится к участкам потери сигнала (трекинга).

Таблица 2.1

Характеристики событий алгоритма I-DT

		Фикс.	Сак.	Морган.	Потери
Количество	n	110737	119395	9146	12799
	%	44	47	4	5
Длительность (мс)	min	80	2.5	64	4.0
	max	5448	16629	674451	19307
	ср.	242	63	328	156
	ст. откл.	197	200	7156	674
Средняя скорость (гр/с)	min	1.06	.02	.09	0
	max	273	1612	847	1163
	ср.	9	53	260	185
	ст. откл.	7.4	71	149	161
Пиковая скорость (гр/с)	min	2.3	.03	0	0
	max	6196	13741	22676	14481
	ср.	99	352	3039	1445
	ст. откл.	169	715	1477	1517
Пройденная дистанция (пикс)	min	0	.01	0	0
	max	1031	270897	27894	532205
	ср.	73	724	1646	2729
	ст. откл.	46	3040	1816	14275
Амплитуда (гр)	min	.04	.02	.08	.03
	max	7.7	45.3	45.7	49
	ср.	.6	5.9	25.5	15.4
	ст. откл.	.2	6.2	11.8	12.5
Площадь вып. оболочки (пикс кв)	min	4.1	1.8	7.8	2.7
	max	609	4207	4272	4284
	ср.	54	498	2142	1304
	ст. откл.	18	541	1009	1070

Из Табл. 2.2 видно, что после фильтрации больше всего удаляется паразитных событий, классифицированных как саккады (83 % от начального количества) – высокоскоростные шумы и нарушения трекинга чаще всего похожи именно на них. Длительность всех трех типов событий после фильтрации приближается к теоретическим представлениям – 300 и более мс для фиксаций, около 35 мс для саккад и от 200 до 2500 мс для прослеживаний. Также нормализуются и показатели средней и пиковой скорости. Из данных по дистанции видно, как фильтрация удаляет невалидные события с чрезмерно большими пройденными расстояниями и амплитудой. Данные по площади выпуклой оболочки (convex hull) свидетельствуют, что до фильтрации в массиве данных существовали некоторые фиксации с нереалистично большими площадями, а среди саккад и прослеживаний были ошибочно классифицированы шумовые участки.



Таблица 2.2

**Описательная статистика событий, полученных классификатором I-BDT
(усреднено по всем экспериментальным ситуациям)**

		До фильтрации				После фильтрации		
		фикс.	сак.	просл.	потери	фикс.	сак.	просл.
Количество	n	173835	195585	83080	25128	48950	15852	27585
	%	36.5	41	17.5	5	53	17	30
Длительность (мс)	min	2.4	2.3	2.2	1.9	80.1	18.1	80.0
	max	10663	568	5140	674476	10663	78	2495
	ср.	132	24	110	98	336	34.8	253
	ст. откл.	278	24	175	4293	373	10.6	203
Средняя скорость (гр/с)	min	0	.02	.06	0	.7	30.0	.6
	max	3544	2547	3856	1788	205	1047	30
	ср.	42.4	107	58.8	270	7	117	11
	ст. откл.	314	947	455	1581	6.2	69.3	6.9
Пиковая скорость (гр/с)	min	0	0	0	0	2.3	68	2.3
	max	11811	14481	13741	22676	30	800	395
	ср.	57	608	191	1905	16.3	365	89.4
	ст. откл.	314	947	455	1581	6.0	147	78.6
Пройденная дистанция (пикс)	min	0	1.6	0	0	3.4	94.4	3.6
	max	3003	24804	36792	9579	1538	715	5948
	ср.	48	381	226	615	62.8	183	218
	ст. откл.	77.6	707	718	1025	61.6	80.0	334
Амплитуда (гр)	min	.01	.12	.01	.05	.05	2.5	.05
	max	6.5	45.8	43	45	2.5	16.0	13.2
	ср.	.51	6.5	1.8	28.5	.5	4.3	1.3
	ст. откл.	.4	7.1	2.5	9.6	.3	1.9	1.2
Площадь вып. оболочки (пикс кв)	min	.72	10.3	.7	4	4.4	189	4.6
	max	522	3723	3509	4272	260	1303	1122
	ср.	44	533	148	2382	48	331	115
	ст. откл.	34	582	201	816	26	150	98

Сравнение двух алгоритмов Для алгоритма I-DT были использованы следующие параметры: минимальная продолжительность фиксации 80 мс, максимальная дисперсия 40 пикс. На Рис. 4 представлены коробочные графики распределений важнейших свойств событий для обоих алгоритмов. Использован один и тот же набор данных. Из графика длительности видно, что алгоритм I-DT не захватывает длинные фиксации (более 300 мс) – вероятно, из-за шумов, которыми события могут прерываться, а для пороговых алгоритмов это основная причина разрывов. Тогда как классификатор I-BDT содержит полный спектр фиксаций вплоть до 400



мс и более (на графике не показаны). Из графика пиковой скорости видно, что I-DT содержит слишком быстрые фиксации и слишком медленные саккады, тогда как I-BDT четко относит эти два типа событий в нужный скоростной диапазон. Все фиксации находятся в интервале 10–30 гр/с, а саккады – в интервале 250–440 гр/с. На графике амплитуды снова видно, что пороговый I-DT выделил слишком широкий диапазон саккад, в которые попадает много шума с амплитудой от 8 гр и выше. Стоит отметить, что такой критерий фильтрации как разность дистанции и амплитуды позволил нам удалить 78 % паразитных событий (Табл. 1), а отношение амплитуды к площади выпуклой оболочки – 62 %. Оба они, как нетрудно понять, отсекают события, габариты которых выходят за пределы прямо направленной траектории саккад (чрезмерно криволинейные), но при этом не накладывают прямого ограничения на скалярную величину амплитуды.

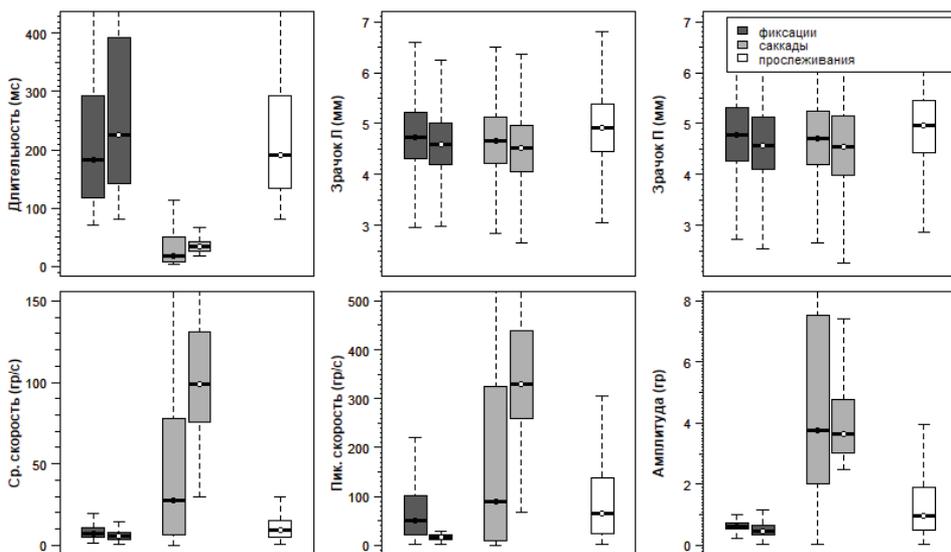


Рис. 4. Описательная статистика двух классификаторов, сгруппированная по событиям. Черная точка – (SMI) I-DT; белая точка – I-BDT (после фильтрации).

Границы коробок – межквартильный размах; усики – полуторный межквартильный размах; гор. черта – медиана

Из графиков на Рис. 5А (результаты детекции SMI I-DT) видно, что вывод данного алгоритма в части саккад не соответствует теоретическим положениям о свойствах этого типа окулomotorных событий. На графике распределения длительности видно (1 ряд, 1 столбец), что в массиве полученных данных много саккад длительностью 80 мс и более, а также менее 20 мс. Из графика дистанция-длительность (1 ряд, 4 столбец) видно, что в составе полученных событий есть много саккад с нереалистичной длительностью 500 мс. Амплитудные показатели некоторых фиксаций доходят до 3 гр при длительности всего 100 мс, что не может являться настоящей фиксацией (1 ряд,

5 столбец). Графики зависимости скорости от дистанции также показывают большое включение слишком быстрых фиксаций при слишком малой пройденной дистанции (типичный признак неотфильтрованных шумовых артефактов).

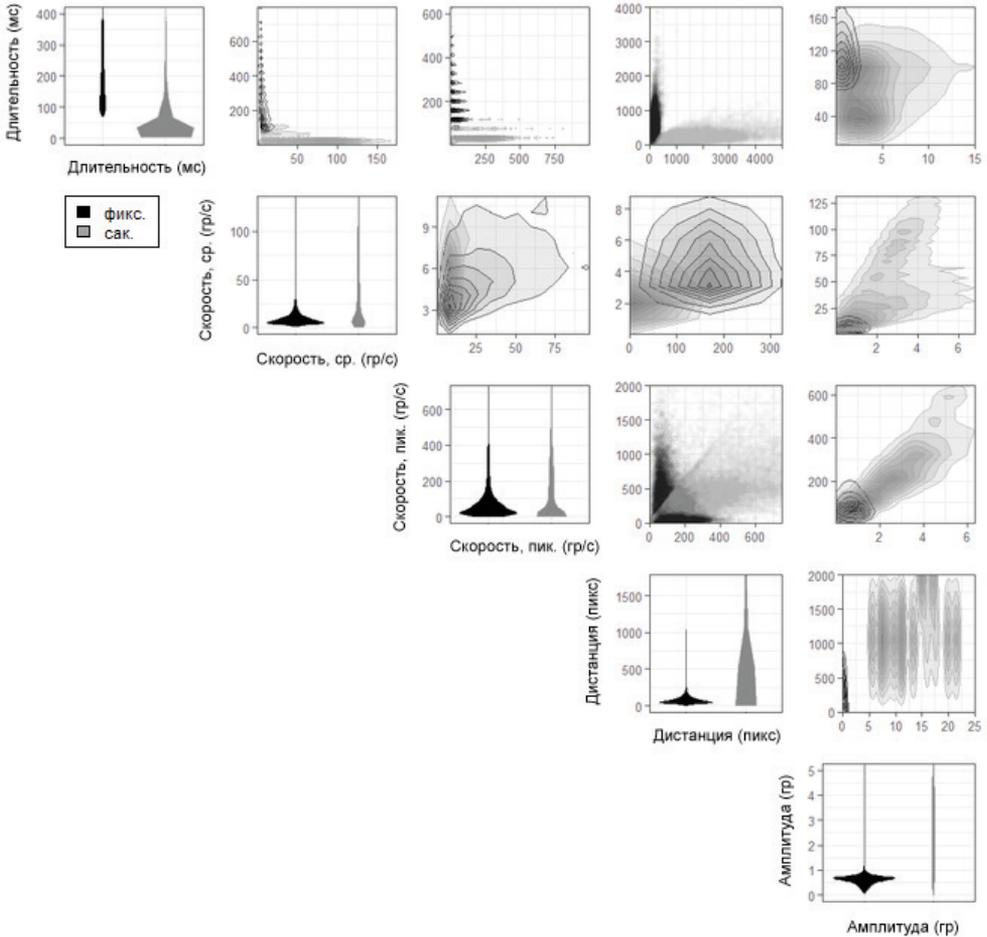


Рис. 5А. Парный график рассеяния основных характеристик событий алгоритма детекции (SMI) I-DT. Масштаб осей выбран произвольно в целях фокусировки на значимых деталях и не отражает фактические границы данных (не включает экстремальные значения и выбросы). Облако точек (scatter) показано на тех графиках, где пересечение контурных графиков нечитаемо

Из графика на Рис. 5Б видно, что скоростные характеристики классифицированных рассматриваемым алгоритмом (I-BDT) прослеживаний находятся в диапазоне между таковыми фиксаций и саккад. Амплитуда и дистанция также распределена в промежутке между статическим и быстрым движением глаза. Длительность же практически идентична фиксациям. В целом, после применения батареи фильтров



ни одна из приведенных статистических визуализаций не содержит данных, которые можно было бы считать артефактными.

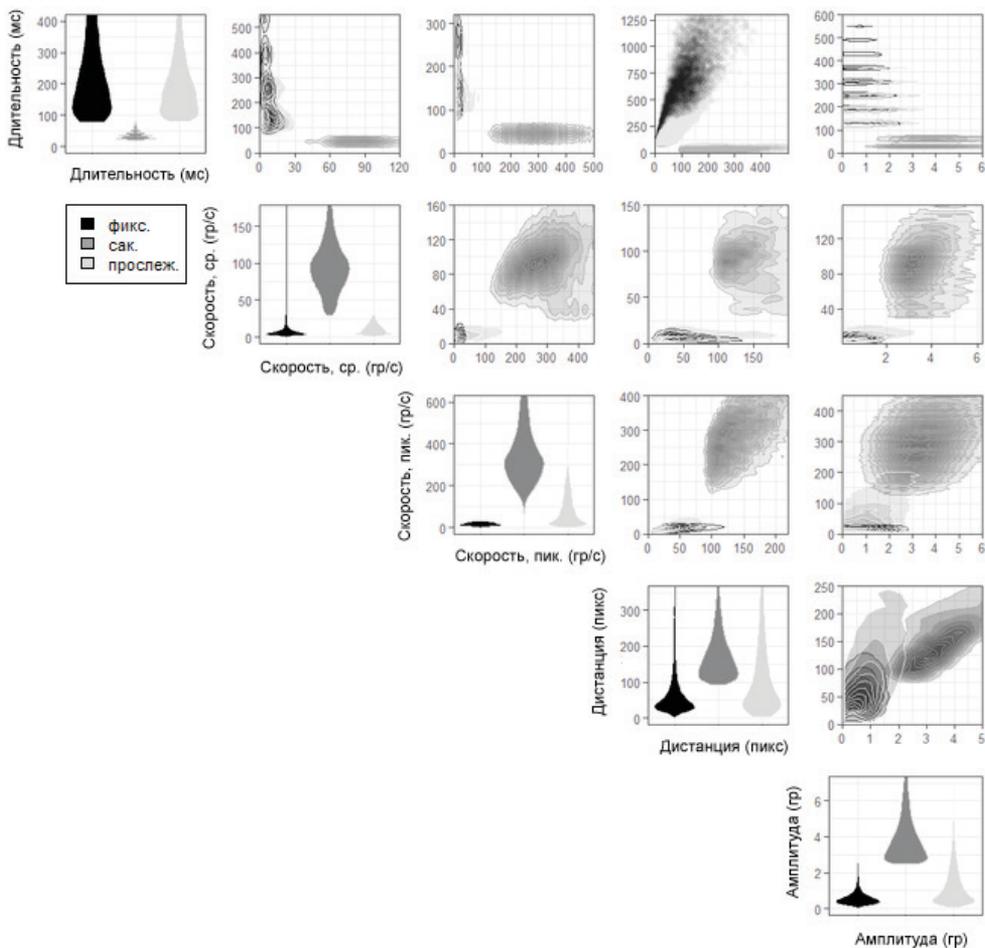


Рис. 5Б. Парный график рассеяния основных характеристик событий классификатора I-BDT (после фильтрации). Параметры расчета контурных графиков (детализация, количество изолиний) выбраны произвольно, индивидуально для каждого графика, в целях улучшения читаемости. График распределения на главной диагонали отражает форму распределения, но не соотношение между типами событий

Количественные метрики Матрица неточностей для двух классификаторов представлена в Табл. 3.1. Поскольку алгоритмы не полностью совпадают по выделяемым классам событий (I-BDT выделяет третий класс – прослеживания), предварительно оба массива были перекодированы и приведены к общему виду. Класс прослеживаний (в I-BDT) и класс шумов (в обоих классификаторах) был заменен на другой условный класс («остальные»). Такая схема перекодировки позволяет наиболее точно оце-



нить согласование общих классов, но компрометирует сравнение уникальных классов в массивах данных. Из таблицы видно, что процент согласований классификатора I-BDT с алгоритмом детекции I-DT в части фиксаций практически максимальный, тогда как большая часть несоответствий возникает в части саккад (62 % совпадает; 30 % классифицированных как саккады алгоритмом I-DT детектируются как фиксации). Абстрактный класс «остальные» не несет смысловой нагрузки, так как создан искусственно в целях сравнения двух множеств с неполностью пересекающимися категориями.

Таблица 3.1

Матрица неточностей для двух классификаторов (%)

I-BDT	(SMI) I-DT		
	фикс.	сак.	ост.
Фиксации	95	3.7	.8
Саккады	30	62	7.2
Остальные	60	19	20

Матрица неточностей была построена на основании доли совпадающих сэмплов (классифицированных обоими алгоритмами одинаково), и не отражает какую-либо пропорцию логических пересечений интервалов, в состав которых они входили. Стоит отметить, что несмотря на достаточно высокую долю совпадений по показателю точности, к Коэна ниже рекомендуемого значения 0,4 (Табл. 3.2), при котором можно бы было утверждать хорошее согласование результатов классификаций.

Таблица 3.2

Количественные метрики согласования между классификаторами

	Фикс.	Сак.	Ост.
Сбалансированная точность (balanced accuracy)	.79	.56	.81
Точность (precision)	.96	.63	.20
Полнота (recall)	.75	.13	.91
F1	.84	.22	.33
к Коэна (Cohen's kappa)		.31	

4. ОБСУЖДЕНИЕ

Классификатор I-BDT показал свою применимость на данном наборе данных и успешно выделил события трех типов согласно заявленным. Среди обнаруженных им следящих движений можно выделить даже мельчайшие участки плавного ускорения глаза. Однако, это чисто математический подход к кластеризации сигнала на основании скоростного профиля, и для получения набора событий с физиологически адекватными характеристиками необходимо проводить дальнейшую фильтрацию. Мы разработали батарею таких фильтров (в общей сложности 13 параметров, применяемых к 6-ти различным свойствам каждого события).

На Рис. 6 приведен среднестатистический скоростной профиль для трех типов детектированных данным классификатором событий. Виден практически плоский профиль



скорости фиксаций (усреднено 49 тыс. фиксаций), с медианной скоростью равной 5.5 гр/с; видна типичная параболическая траектория изменения скорости саккады с пиком более 220 гр/с (усреднено 16 тыс. саккад); на профиле прослеживаний виден участок разгона, затем плато с медианой около 10 гр/с (усреднено 28 тыс. прослеживаний). Большой размах скорости на первых самплах фиксаций и прослеживаний обусловлен тем, что классификатор не точно размечает начальную границу событий (она обычно запаздывает на 2–3 сампла), вследствие чего в начале этих классов событий возможен «захват» фрагмента предыдущего, более быстрого события или, чаще, участка шума. Данную особенность классификатора следует учитывать, если требуется разметить точные временные границы событий.

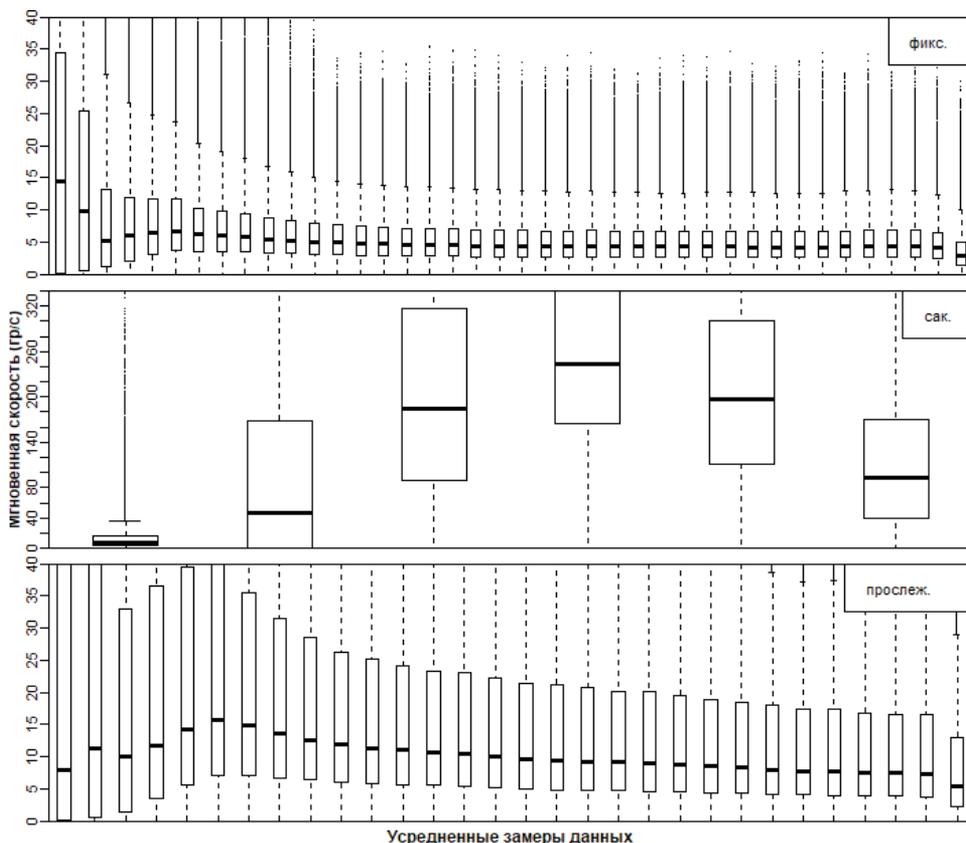


Рис. 6. Интерполированные профили скорости для трех типов событий, классифицированных I-BDT (после фильтрации)

Пояснение: все события одного класса приведены к общему количеству отсчетов методом узкополосной интерполяции, затем усреднены (кривая локально-взвешенной полиномиальной регрессии LOWESS), после чего для каждого отсчета (сампла) посчитана статистика, выраженная в виде коробок. Верхняя и нижняя границы коробок обозначают межквартильный размах, горизонтальная черта – медиану.



Благодаря тому, что алгоритм основан на относительном распределении скорости по выделяемым кластерам событий, а не строгом задании пороговых значений, он практически нечувствителен к шумам. При этом также возможна и классификация прослеживаний на зашумленных участках, если скоростной профиль данного участка лучше всего описывается смещением двух гауссиан, одна из которых описывает фиксации, другая – саккады.

Помимо описанной выше фильтрации событий, нами была предпринята попытка провести кластерный анализ в качестве фильтрации вывода классификатора, но она не дала нужных результатов. Предположительно потому, что вывод классификатора изначально получен от непрерывной функции (гауссова смесь) без применения каких-либо пороговых ограничений, а объективные параметры событий не проявились в данных – либо по причине низкой частоты регистрации, либо по причине их физического отсутствия (нет четко проходящей границы между длительностями / скоростями фиксаций и прослеживаний, настоящих фиксаций и артефактов записи).

Также в ходе работы нами была предпринята попытка сглаживания данных (кривой типа сплайн). Однако, никакого качественного улучшения классификации данная операция не дает. Временные границы плавных следящих движений при этом, ожидаемо, расширяются. Другими словами, сглаживание может приводить к детекции новых артефактных участков плавных следящих движений глаз. Исходя из этого, мы не рекомендуем сглаживать данные, даже если они сильно зашумлены.

5. ЗАКЛЮЧЕНИЕ

В ходе тестирования классификатора было проведено сравнение его результатов с результатами классического алгоритма детекции окуломоторных событий I-DT посредством выполнения обоих классификаторов на одинаковом наборе данных и подсчета одинакового набора свойств для каждого из результатов. Общее заключение по согласованию результатов классификации выполнено при помощи критерия к Козна и матрицы неточностей. Рассмотренный классификатор позволяет детектировать важный для исследований восприятия в натуральных условиях тип событий – плавные прослеживающие движения (smooth pursuit). В отличие от полной версии алгоритма, представленный упрощенный вариант имеет большую чувствительность к следящим движениям в тех данных, где фактически отсутствуют длительные по времени, широкоамплитудные прослеживающие движения, или присутствует большое количество шума. Классификатор обладает высокой чувствительностью к окуломоторным событиям трех типов – фиксациям, саккадам, и плавным следящим движениям. Он позволяет выделять даже низкоамплитудные (1–1.5 градуса) эпизоды прослеживаний, которые пороговыми алгоритмами детекции неминуемо включаются в состав фиксаций как якобы участки с повышенным разбросом, либо в состав коротких саккад как участки с низкой скоростью скачка. Однако, вывод данного классификатора требует дополнительной фильтрации.

Алгоритм основан на абстрактной математической модели (гауссово распределение) и не предполагает для событий никаких фильтров, параметризующих вывод



классификатора в соответствии с теоретическим представлением о них. Поэтому мы рекомендуем дополнительно производить фильтрацию событий на основании ряда характеристик. В числе таких типичных фильтров, как минимальная длительность фиксации, максимальная скорость саккады, для саккад мы дополнительно вводим разность дистанции и амплитуды, и порог по площади выпуклой оболочки (описанной вокруг облака точек), что позволяет отделить паразитные шумы – без грубого деления массива данных на интервалы по произвольно заданной величине.

Литература

1. *Santini T., Fuhl W., Kübler T., et al.* Bayesian identification of fixations, saccades, and smooth pursuits // ACM ETRA. Charleston, 2016. P. 163–170. DOI:10.1145/2857491.2857512
2. *Nyström M., Andersson R., Holmqvist K., et al.* The influence of calibration method and eye physiology on eyetracking data quality // *Behav. Res. Met.* 2013. Vol. 45. № 1. P. 272–288. DOI:10.3758/s13428-012-0247-4
3. *Hooge I., Holmqvist K., Nyström M.* The pupil is faster than the corneal reflection (CR): Are video based pupil-CR eye trackers suitable for studying detailed dynamics of eye movements? // *Vis. Res.* 2016. Vol. 128. P. 6–18. DOI:10.1016/j.visres.2016.09.002.
4. *Larsson L., Nyström M., Ardö H., et al.* Smooth pursuit detection in binocular eye-tracking data with automatic video-based performance evaluation // *J. Vis.* 2016. Vol. 16. № 15. P. 20. DOI:10.1167/16.15.20
5. *Startsev M., Agtzidis I., Dorr M.* 1D CNN with BLSTM for automated classification of fixations, saccades, and smooth pursuits // *Behav. Res. Met.* 2019. Vol. 51. P. 556–572. DOI:10.3758/s13428-018-1144-2
6. *Zemblys R., Niehorster D.C., Komogortsev O., et al.* Using machine learning to detect events in eye-tracking data // *Behav. Res. Met.* 2018. Vol. 50. P. 160–181. DOI:10.3758/s13428-017-0860-3
7. *Komogortsev O.V., Karpov A.* Automated classification and scoring of smooth pursuit eye movements in the presence of fixations and saccades // *Behav. Res. Met.* 2013. Vol. 45. P. 203–215. DOI:10.3758/s13428-012-0234-9
8. *Komogortsev O. V., Gobert D. V., Jayarathna S., et al.* Standartization of automated analyses of oculomotor fixation and saccadic behaviors // *IEEE Trans. Biomed. Eng.* 2010. Vol. 57. № 11. P. 2635–2645. DOI:10.1109/tbme.2010.2057429
9. *Kashyap H.J., Detorakis G., Dutt N., et al.* A recurrent neural network based model of predictive smooth pursuit eye movement in primates // *IEEE IJCNN.* Rio de Janeiro, 2018. P. 5353–5360. DOI:10.1109/IJCNN.2018.8489652
10. *Xivry J.J.O. de, Coppe S., Blohm G., et al.* Kalman Filtering Naturally Accounts for Visually Guided and Predictive Smooth Pursuit Dynamics // *J. Neurosci.* 2013. Vol. 33. № 44. P. 17301–17313. DOI:10.1523/JNEUROSCI.2321-13.2013



Assessment of Bayesian Ternary Gaze Classification Algorithm (I-BDT)

Ivan Y. Zherdev*

Moscow State University of Psychology & Education,
Moscow, Russia (affiliated researcher)

ORCID: <https://orcid.org/0000-0001-6810-9297>

e-mail: ivan866@mail.ru

I-BDT eyetracking data ternary classification (fixations, saccades, smooth pursuit) algorithm [1] is investigated. Comparison with well-known Identification / Dispersion Threshold (I-DT) algorithm is held (accuracy, precision, recall, F1 measure). A novel approach for additionally filtering the algorithm output by distance/amplitude, area of convex hull is described.

Keywords: smooth pursuit, classification, Bayesian decision theory, eye movements.

For citation:

Zherdev I.Y. Assessment of Bayesian Ternary Gaze Classification Algorithm (I-BDT). *Modelirovanie i analiz dannykh = Modelling and Data Analysis*, 2020. Vol. 10, no. 2, pp. 74–92. DOI:10.17759/mda.2020100206 (In Russ., abstr. in Engl.).

Funding. The reported study was funded by Russian Scientific Foundation (RScF), project number 18–18–00350.

References

1. Santini T., Fuhl W., Kübler T., et al. Bayesian identification of fixations, saccades, and smooth pursuits. *ACM ETRA*. Charleston, 2016. pp. 163–170. DOI:10.1145/2857491.2857512
2. Nyström M., Andersson R., Holmqvist K., et al. The influence of calibration method and eye physiology on eyetracking data quality. *Behav. Res. Met.* 2013. Vol. 45, no 1, pp. 272–288. DOI:10.3758/s13428-012-0247-4
3. Hooge I., Holmqvist K., Nyström M. The pupil is faster than the corneal reflection (CR): Are video based pupil-CR eye trackers suitable for studying detailed dynamics of eye movements? *Vis. Res.* 2016. Vol. 128, pp. 6–18. DOI:10.1016/j.visres.2016.09.002.
4. Larsson L., Nyström M., Ardö H., et al. Smooth pursuit detection in binocular eye-tracking data with automatic video-based performance evaluation. *J. Vis.* 2016. Vol. 16, no 15, pp. 20. DOI:10.1167/16.15.20
5. Startsev M., Agtzidis I., Dorr M. 1D CNN with BLSTM for automated classification of fixations, saccades, and smooth pursuits. *Behav. Res. Met.* 2019. Vol. 51, pp. 556–572. DOI:10.3758/s13428-018-1144-2
6. Zemblys R., Niehorster D.C., Komogortsev O., et al. Using machine learning to detect events in eye-tracking data. *Behav. Res. Met.* 2018. Vol. 50, pp. 160–181. DOI:10.3758/s13428-017-0860-3

***Ivan Y. Zherdev**, former jr. research fellow; affiliated researcher, Moscow State University of Psychology & Education, Moscow, Russia, ORCID: <https://orcid.org/0000-0001-6810-9297>, e-mail: ivan866@mail.ru



7. Komogortsev O.V., Karpov A. Automated classification and scoring of smooth pursuit eye movements in the presence of fixations and saccades. *Behav. Res. Met.* 2013. Vol. 45, pp. 203–215. DOI:10.3758/s13428-012-0234-9
8. Komogortsev O. V, Gobert D. V, Jayarathna S., et al. Standartization of automated analyses of oculomotor fixation and saccadic behaviors. *IEEE Trans. Biomed. Eng.* 2010. Vol. 57, no 11, pp. 2635–2645. DOI:10.1109/tbme.2010.2057429
9. Kashyap H.J., Detorakis G., Dutt N., et al. A recurrent neural network based model of predictive smooth pursuit eye movement in primates. *IEEE IJCNN*. Rio de Janeiro, 2018. pp. 5353–5360. DOI:10.1109/IJCNN.2018.8489652
10. Xivry J.J.O. de, Coppe S., Blohm G., et al. Kalman Filtering Naturally Accounts for Visually Guided and Predictive Smooth Pursuit Dynamics. *J. Neurosci.* 2013. Vol. 33, no 44, pp. 17301–17313. DOI:10.1523/JNEUROSCI.2321-13.2013