

Промежуточный отчёт по обработке результатов эксперимента по распознаванию эмоций

Пузырев Дмитрий

31 марта 2017 г.

Пролог. Описание и методика эксперимента

В рамках работы над обучающимся алгоритмом, способным определять по видеопотоку, насколько внимательно человек следит за лекцией на экране, был проведён эксперимент для получения обучающей выборки и её анализа.

Испытуемым предлагалось просмотреть 4 коротких лекции научно-популярного содержания с видеохостинга Youtube (<https://www.youtube.com/watch?v=M75JQA8FaeM>), записать свою реакцию на веб-камеру и после этого пройти небольшой тест. По каждому видео было 3 вопроса: 2 с выбором ответа и 1 с кратким ответом. Также каждый участник эксперимента мог оставить свои общие впечатления о просмотренном и пожелания по эксперименту.

Далее в отчёте будут предоставлены основные зависимости между основными визуальными паттернами и "интересностью" конкретного видео для респондента. Также будут даны численные оценки значений данных параметров для определения заинтересованности человека.

В работе были исследованы влияние различных параметров моргания глаз, положения головы, движения бровей и губ на то, насколько заинтересован человек в просмотре данного видео и насколько хорошо он справляется с тестом на усвоение материала.

I. Моргание

Была поставлена задача выяснить, как изменяются характеристики моргания во время просмотра различного по содержанию видеоконтента. Таковых характеристик выделено две: время между двумя морганиями и длина одного моргания.

Для замера нужных данных я выбрал видео одной из испытуемых. В анкетировании были показаны следующие результаты:

- Видео 1 - 3 правильных ответа
- Видео 2 - 2 правильных ответа
- Видео 3 - 0 правильных ответов
- Видео 4 - 3 правильных ответа

Исходя из результатов тестирования можно предположить, что наибольшую концентрацию респондент проявила в первом и в четвёртом видео, второе видео заинтересовало её несколько меньше, а третье не заставило сконцентрироваться. Теперь проанализируем полученные результаты.

Видео 1

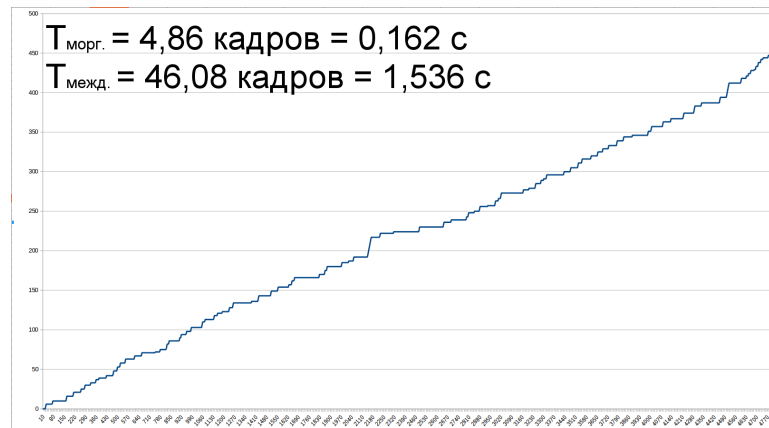


Рис 1. Функция частоты морганий видео 1

Из построенного выше графика количества морганий от времени видно, что он хорошо приближается линейной функцией. Это происходит потому, что он состоит из коротких "ступенек" то есть моргания достаточно короткие, частые и имеют между собой небольшой интервал. Были рассчитаны среднее время одного моргания (4,96 кадра), среднее время между морганиями (46,08 кадра) и частота моргания (35 в минуту). Скоростные параметры моргания при просмотре видео будут значительно выше, чем в обычной жизни, что можно связать с тем, что ожидаемое действие на видео на некоторое время запаздывает и заставляет нас моргать чаще.

Видео 2

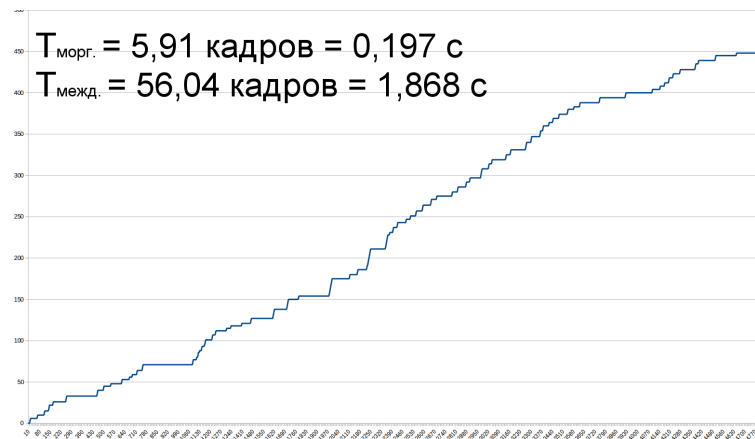


Рис 2. Функция частоты морганий видео 2

Этот график гораздо меньше похож на линейную функцию, так как увеличились как промежутки между морганиями, так и длина одного моргания (5,91 кадр и 56,04 кадра соответственно). Средняя частота моргания — 29 в минуту.

Видео 3

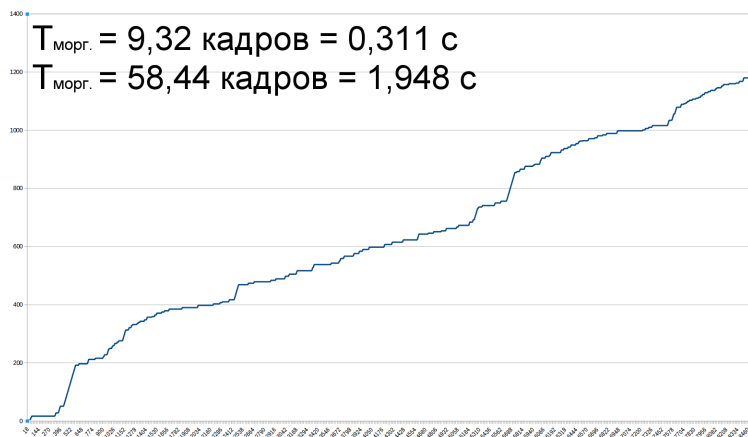


Рис 3. Функция частоты морганий видео 3

Третье видео было несколько длиннее предыдущих, и оно было наименее интересным респонденту. По графику видно, что он стал моргать значительно дольше (9,32 кадра) и делать большие промежутки между морганиями (58,44 кадра). Средняя частота моргания также снизилась — 27 в минуту.

Видео 4

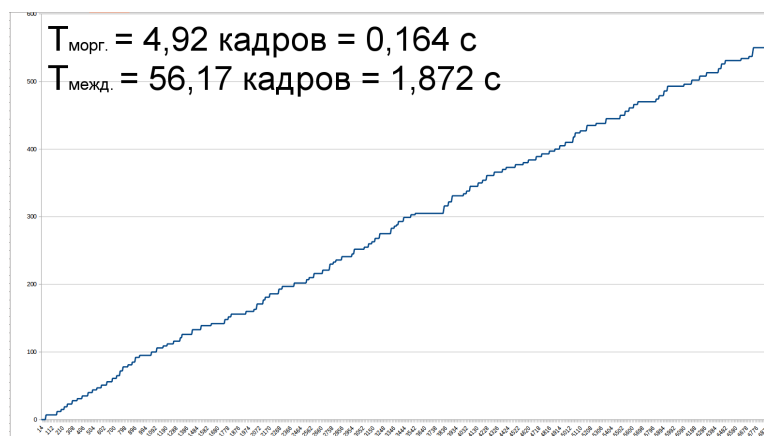


Рис 4. Функция частоты морганий видео 4

Это видео, в отличие от предыдущего, вызвало интерес, также мы видим, что график похож на линейную функцию. Сделав замеры, получили 4,92 кадра как среднюю длину моргания и 56,17 кадров в среднем между морганиями с частотой моргания 30 в минуту.

Исходя из полученных данных, можно выявить ряд корреляций между параметрами моргания и внимательностью, проявленной зрителем при просмотре:

- **длина одного моргания:** большая "внимательность" наблюдается при сокращении длины текущего моргания; моргание в среднем больше 6 кадров можно рассматривать как признак невнимательности
- **время между морганиями:** чем меньше промежутки между морганиями, тем чаще зритель "обновляется чтобы синхронизировать происходящее на экране, что может говорить о большей заинтересованности; однако количественные оценки данного параметра достаточно сложны в виду влияния на него темпа самого видео и индивидуальных характеристик зрения
- **частота моргания:** как и в случае с предыдущим параметром, увеличивающаяся частота говорит об увеличении темпов синхронизации восприятия с видео; уменьшение частоты идентифицирует попытку смотрящего сконцентрироваться на данном

моменте, что является сигналом того, что, возможно, материал является слишком сложным; в силу схожих с предыдущим пунктом причин численные оценки затруднительны.

II. Положение головы

Следующим макропараметром, на который я обратил внимание в рамках исследования, является то, как перемещается голова зрителя при просмотре.

Для фиксирования перемещений и построения графиков была использована программа Tracker. По горизонтальной оси отложено время в секундах, на вертикальной оси — смещение центра головы (нос) относительно его изначального положения в условной единице, равной 2 пикселям при записи в качестве 720р. Ниже показан процесс фиксирования текущего положения.

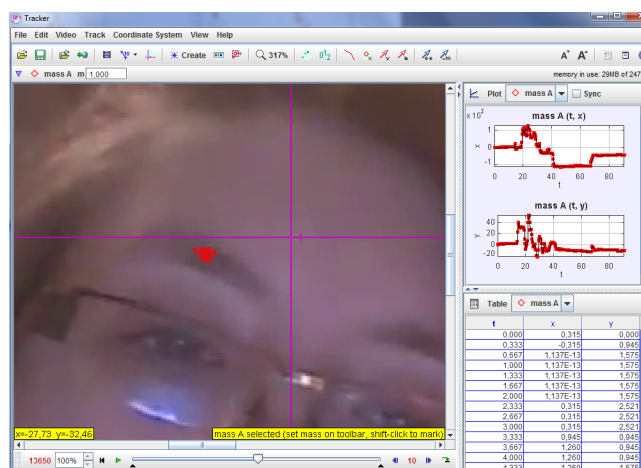


Рис 5. Отслеживание траектории в программе Tracker

Также графики были статистически проанализированы по 5 параметрам: максимальная и минимальная координата, среднее значение координаты, среднеквадратичное отклонение и стандартная ошибка.

Видео 1

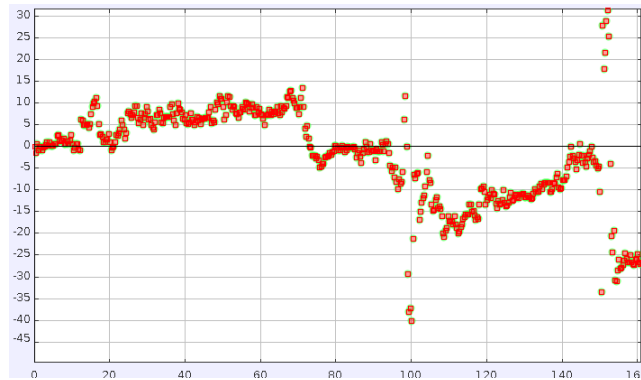


Рис 6. Смещение по координате x видео 1

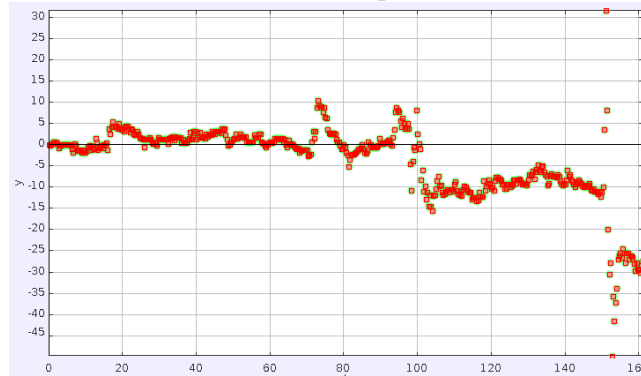


Рис 7. Смещение по координате y видео 1

3,150E1	3,175E1
-4,000E1	-4,975E1
-2,239E0	-3,826E0
1,089E1	8,636E0
4,949E-1	3,925E-1
484	484

Рис 8. Статистика смещений видео 1

Во время просмотра видео смещения положения головы по обоим координатам незначительны, не считая пары "рывков" в середине и конце видео. Значения среднеквадратичного отклонения и стандартной ошибки не слишком велики. В среднем от начала видео голова сместилась слегка вниз и влево.

Видео 2

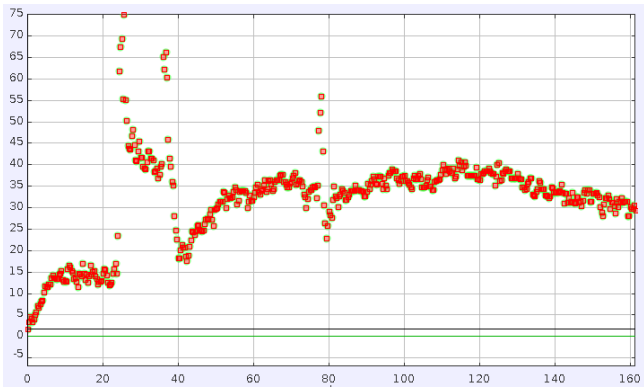


Рис 9. Смещение по координате x видео 2

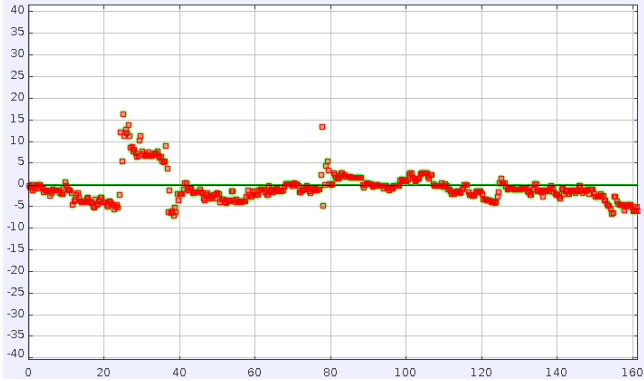


Рис 10. Смещение по координате y видео 2

7,515E1	1,632E1
1,752E0	-6,952E0
3,182E1	-6,448E-1
1,034E1	3,405E0
4,693E-1	1,546E-1
485	485

Рис 11. Статистика смещений видео 2

В то время как вертикальное положение головы остаётся практически неизменным (кроме 25-40 и 80 секунды), имеется значительное смещение влево с резкими поворотами в начале и середине. Статистически отклонения почти незаметны для оси y и достаточно малы, однако заметны, для оси x.

Видео 3

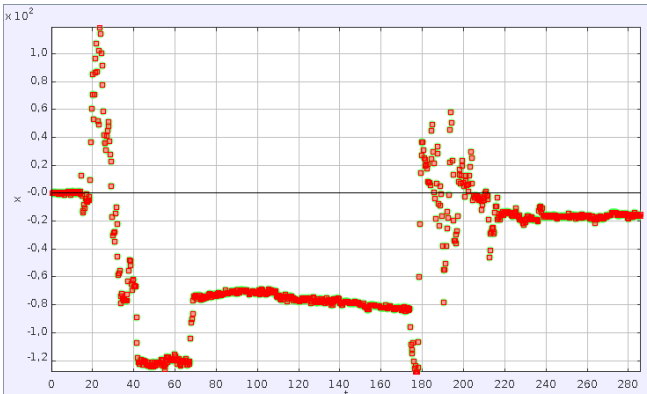


Рис 12. Смещение по координате x видео 3

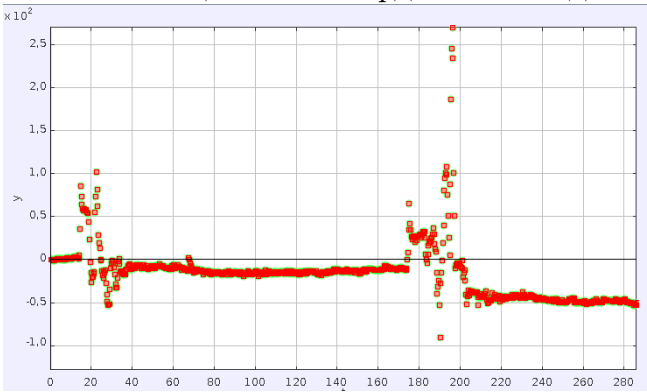


Рис 13. Смещение по координате y видео 3

1,191E2	2,704E2
-1,281E2	-9,038E1
-4,526E1	-1,563E1
4,605E1	3,116E1
1,571E0	1,063E0
859	859

Рис 14. Статистика смещений видео 3

Во время данного видео голова не только кратковременно вертится по обоим координатам, но и "перескакивает" из одного положения в другое, стабильно находясь в нём какой-то период времени, и потом его меняя. Наиболее сильные скачки до 200 единиц наблюдаются на 20-40 секунде и 180-200 секунде. Среднеквадратичное отклонение большое для обеих осей, 46 единиц по x и 31 по y.

Видео 4

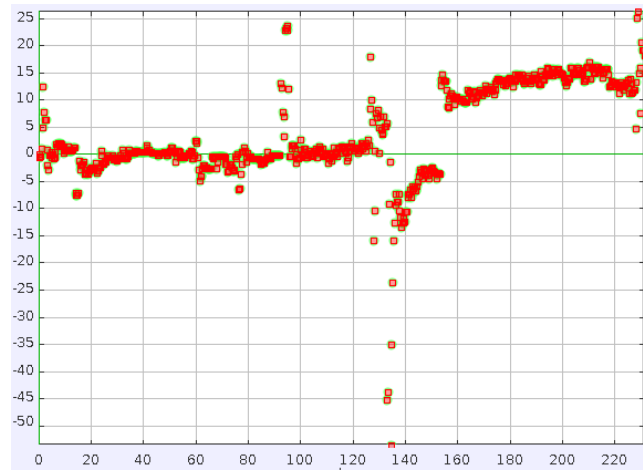


Рис 12. Смещение по координате x видео 4

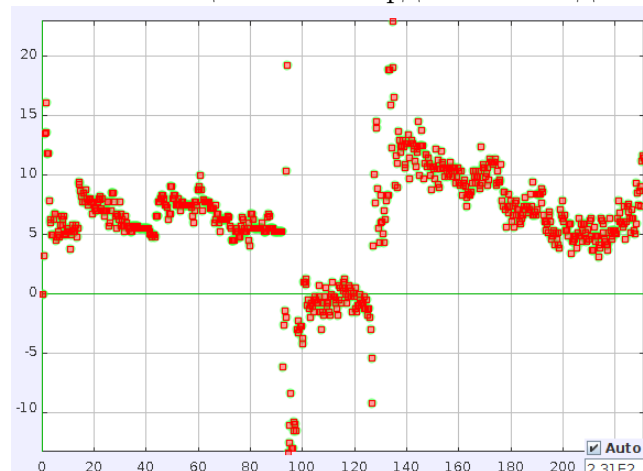


Рис 13. Смещение по координате y видео 4

2,644E1	2,296E1
-5,341E1	-1,331E1
3,939E0	6,074E0
8,473E0	4,393E0
3,216E-1	1,668E-1
694	694

Рис 14. Статистика смещений видео 4

Положение головы в целом стабильное, кроме 2 "рывков" на 120-140 секунде и 90-100 секунде (для оси y). Среднеквадратичные отклонения незначительны (8 для x и 5 для y).

Исходя из наблюдаемых результатов, можно выделить следующие параметры, "моментальные" и "статистические":

"Моментальные" параметры:

- **Резкое смещение (>30 пикселей) вправо-влево относительно предыдущего положения:** невнимательность, очень вероятно, что человек отвлекся на что-то внешнее
- **Резкое смещение (>30 пикселей) вверх-вниз относительно предыдущего положения:** невнимательность, человек также мог отвлечься (например, на часы или телефон) или чувствует себя некомфортно; также отклонение головы вверх может означать то, что человек зевнул, что также не говорит о сосредоточенности на видео.

- **Небольшое смещение (10-20 пикселей) вниз относительно предыдущего положения:** внимательность, человек, слегка наклонивший голову вниз, более настроен к получению информации

"Статистические" параметры:

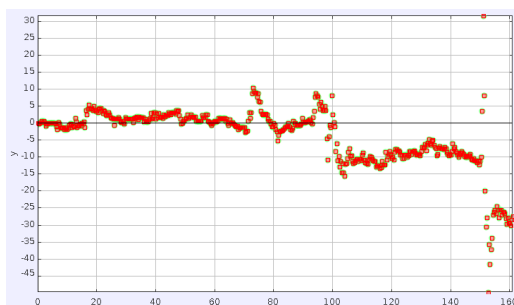
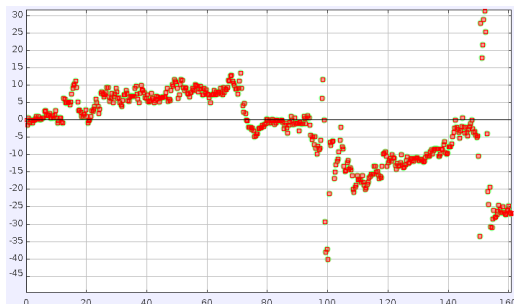
- **Значительно отличающаяся от нуля ($>20-25$ пикселей) координата среднего положения по обеим осям:** невнимательность, сигнализирует о том, что человек часто менял положение головы (как на видео 3), значит, ему скучно или некомфортно.
- **Высокие значения (больше 40 пикселей по оси x и больше 30 по y) среднего квадратичного отклонения:** невнимательность, голова часто "вертелась" человек отвлекался
- **Небольшое (10-20 пикселей) смещение вниз среднего относительно начального положения:** внимательность, аналогично "моментальному" параметру

III. Положение бровей

Ещё один параметр, который говорит о степени вовлечённости в происходящее — это брови.

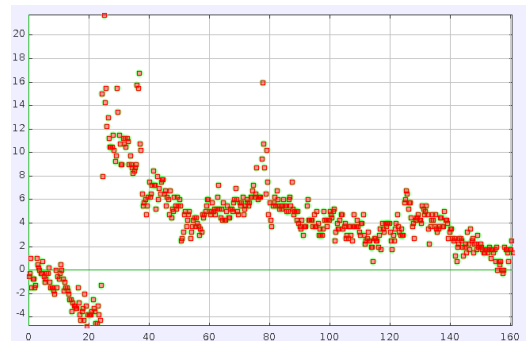
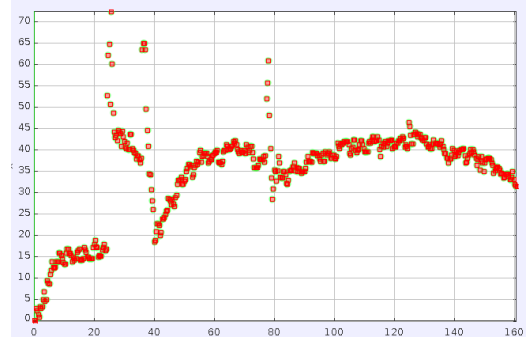
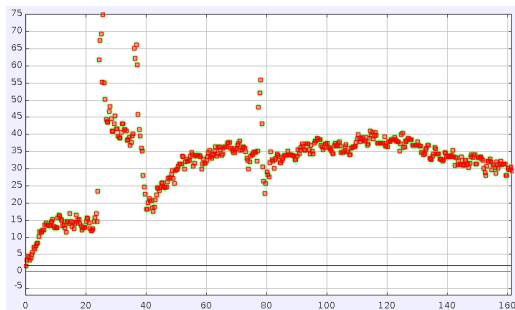
Методика измерения положения останется та же, что и в предыдущем опыте. Исходя из того, что перемещение бровей будет наверняка схожа с перемещением головы в целом, будем давать сравнения с предыдущими графиками и анализировать разницу. Для упрощения будем следить только за правой (для испытуемого) бровью, считая, что обе брови передвигаются одинаково.

Видео 1



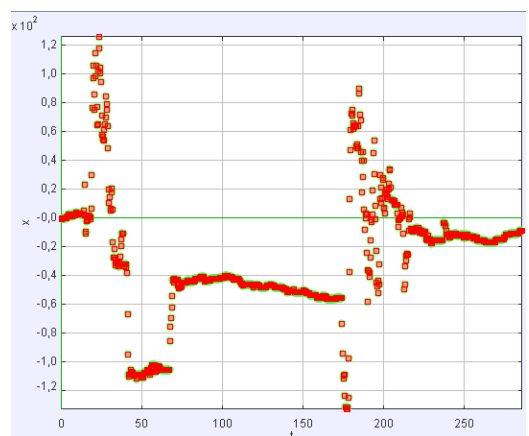
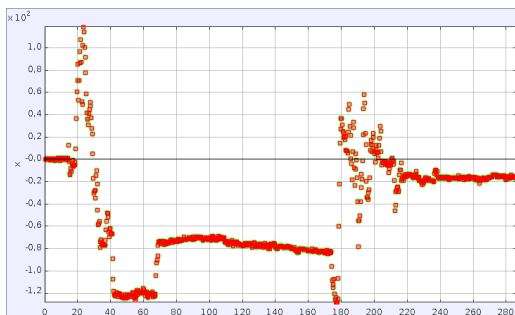
Проанализируем различия по осям x и y . Видим, что в обоих случаях графики в целом повторяют друг друга, но практически с начала видео координата x чуть-чуть (5 - 10 пикселей) сместилась вниз, что говорит о небольшом сдвиге брови в центр. По оси y начиная с середины ролика бровь сдвинута немного вниз, но ближе к концу видео (140 секунда) сместилась выше среднего.

Видео 2

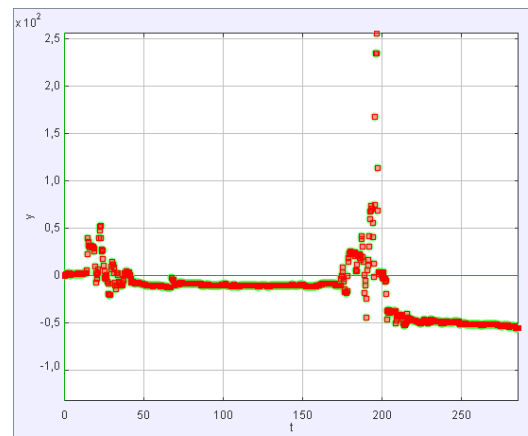
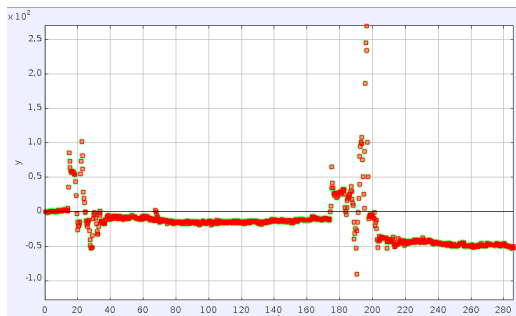


Здесь наблюдаем идентичную картину для оси x и несколько приподнятые брови (примерно 10 пикселей) начиная с 50 секунды и до конца видео.

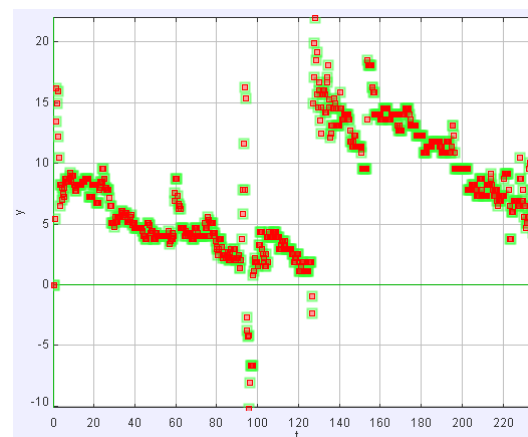
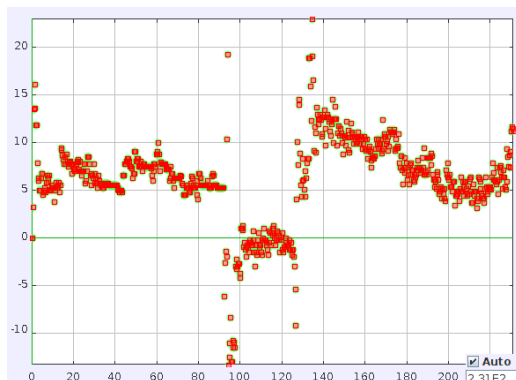
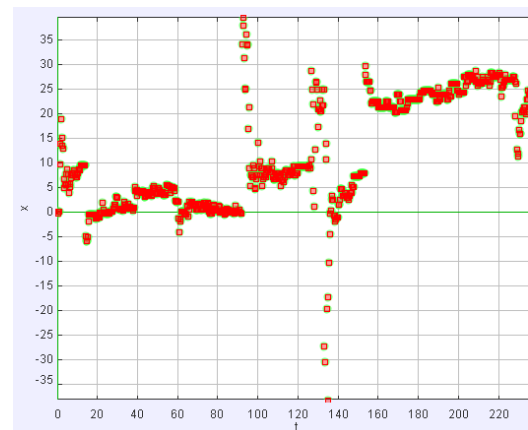
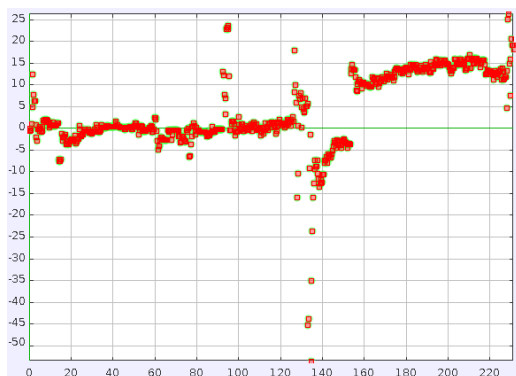
Видео 3



Смещения в координатах существенно различаются в начале и середине видео (60-80 пикселей по оси y и 30-50 пикселей по оси x), что объясняется тем, что голова была наклонена. В конце ролика различий почти нет.



Видео 4



По оси x разница как незначительно увеличивалась в пользу смещения бровей (секунды 100-120, 140-220), так и в пользу смещения головы (20-40, 60-100), все разы — примерно 10 пикселей. Для оси y брови сместились ниже на секундах 40-80, выше — на 140-200.

Сопоставив ожидания с результатами, можно выделить следующие параметры:

- **Большая разница смещений (>40 пикселей) по любой из осей:** невнимательность, скорее всего, имел место наклон головы, сигнализирует о скуке
- **Небольшая разница смещений (10-30 пикселей) вправо в пользу бровей:** брови "насупились" человек пытается что-то понять, возможно, лекция слишком сложная

- **Небольшая разница смещений (10-30 пикселей) вниз в пользу бровей:** брови идут вниз, человек пытается сконцентрироваться, вероятно, следует облегчить материал
- **Небольшая разница смещений (10-30 пикселей) вверх в пользу бровей:** внимательность: брови идут вверх, зрителю нравится, что он видит

IV. Положение бровей

Последнее, чему будет посвящено внимание в отчёте — мимике губ и их влиянию на восприятие материала.

Во время проведения данного опыта я сконцентрировал внимание на измерении расстояния между верхней и нижней губами и между уголком (правым) и центром губы, а также вертикальным положением этого уголка. Далее будут представлены результаты для каждого видео.

Видео 1

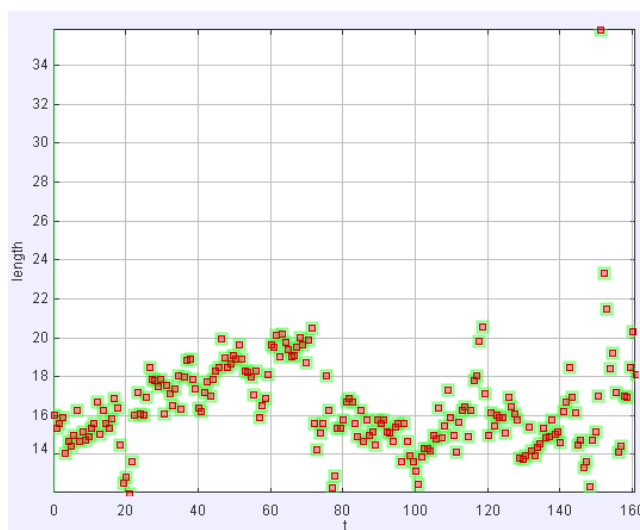


Рис 23. Расстояние между верхней и нижней губой видео 1

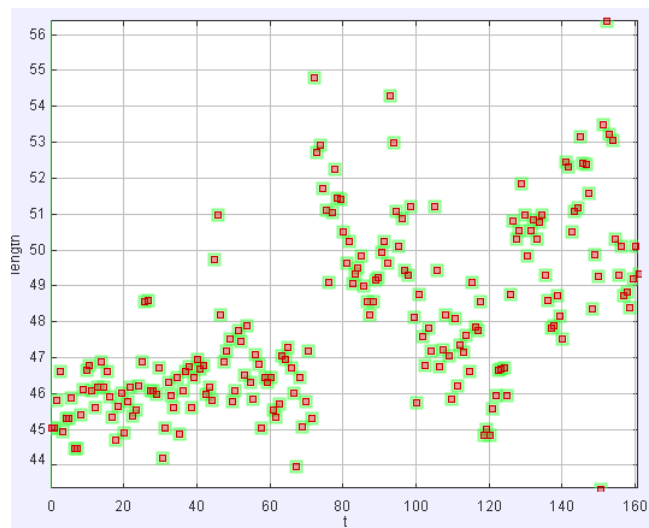


Рис 24. Расстояние от уголка губы до центра видео 1

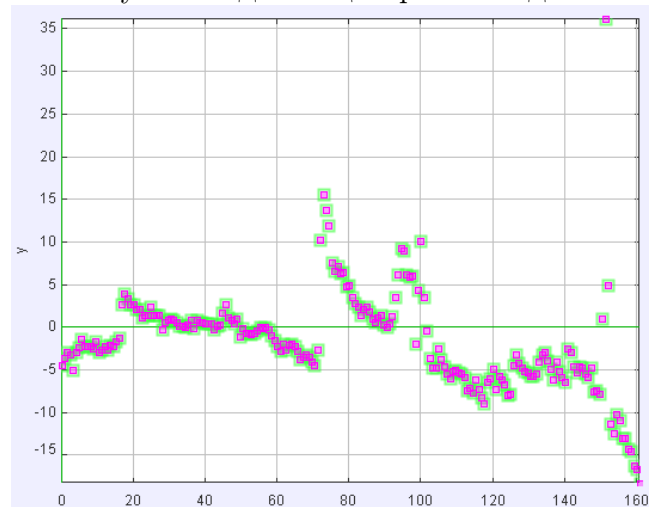


Рис 25. Смещение по оси y уголка губы видео 1

Расстояние между верхней и нижней губой стабильно кроме небольшого падения в середине и колебания в конце (зритель зевнул). В середине ролика уголок сдвинулся вверх и расстояние до центра увеличилось — испытуемый улыбнулся. В конце, напротив, уголки опустились.

Видео 2

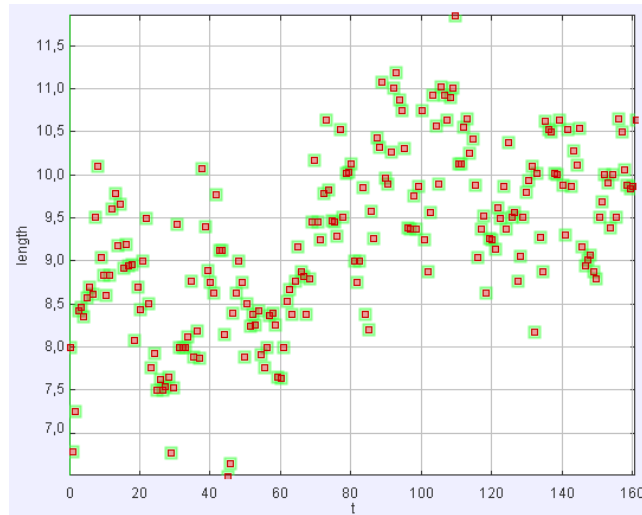


Рис 26. Расстояние между верхней и нижней губой видео 2

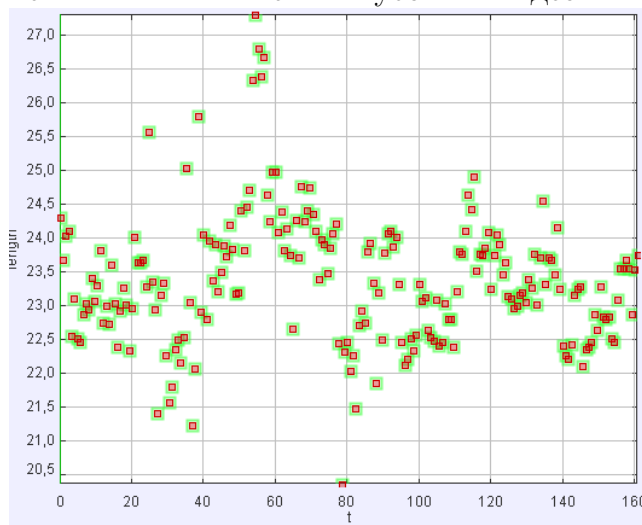


Рис 27. Расстояние от уголка губы до центра видео 2

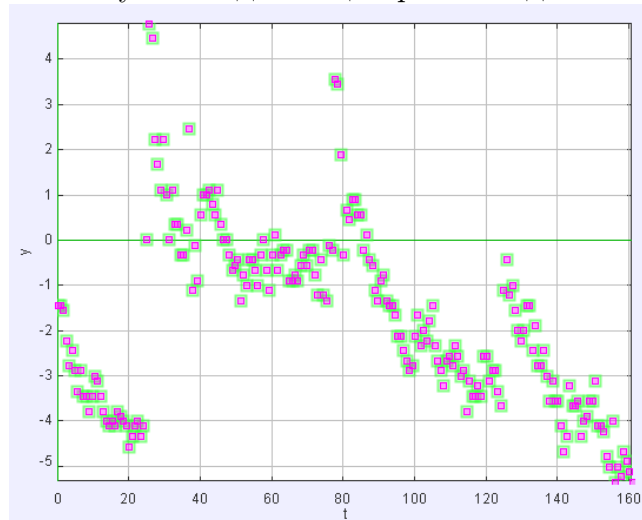


Рис 28. Смещение по оси y уголка губы видео 2

Промежуток между губами изменяется крайне незначительно. Уголки губ почти не двигались вертикально. Изменения мимики губ практически отсутствовали на протяжении ролика.

Видео 3

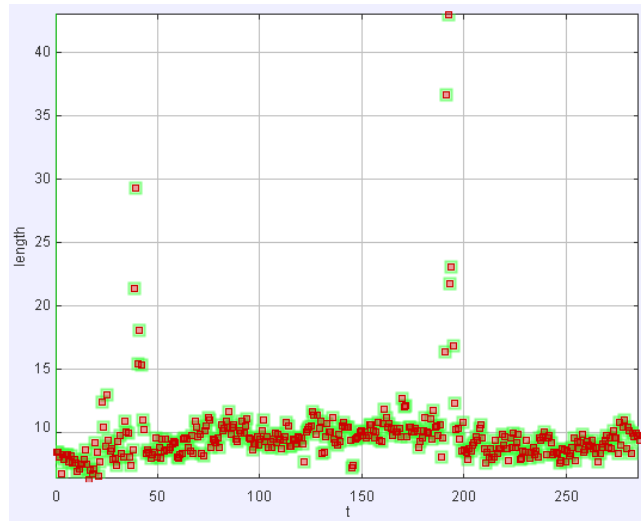


Рис 29. Расстояние между верхней и нижней губой видео 3

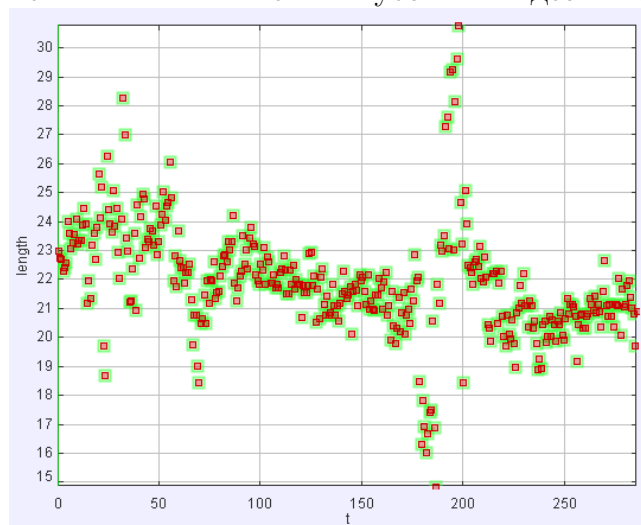


Рис 30. Расстояние от уголка губы до центра видео 3

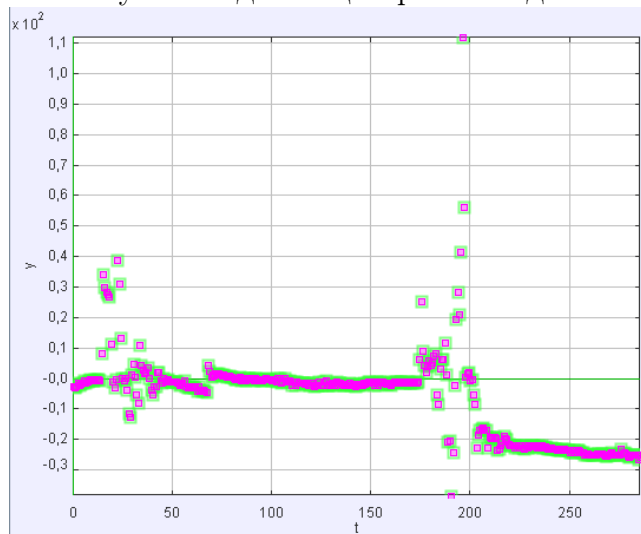


Рис 31. Смещение по оси y уголка губы видео 3

На первом графике можно идентифицировать 2 пика около 50 и 200 секунд. Они соответствуют 2 зевкам, при которых расстояние между губами существенно увеличилось. "Зевки" можно также отличить на 2 следующих графиках: уголки губ значительно ото-

двинулись в сторону.

Видео 4

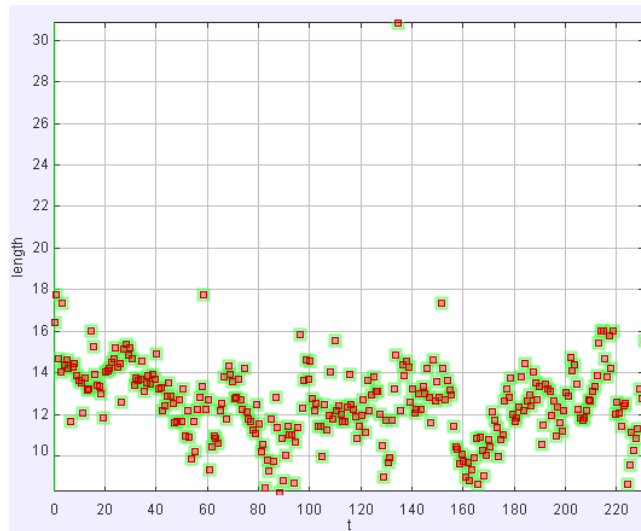


Рис 29. Расстояние между верхней и нижней губой видео 4

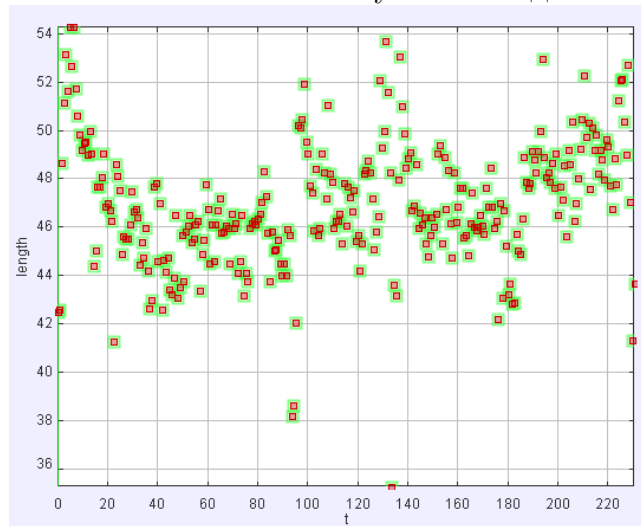


Рис 30. Расстояние от уголка губы до центра видео 4

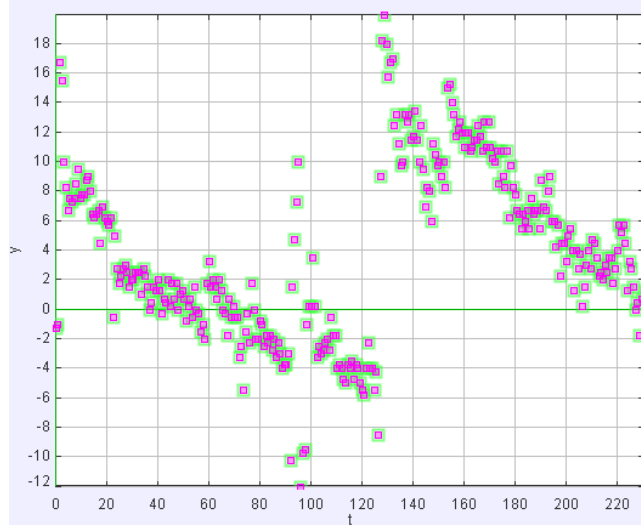


Рис 31. Смещение по оси y уголка губы видео 4

Расстояние между губами практически неизменно на протяжении ролика. Расстояние от угла до центра довольно велико, при это на 140 секунде видим, что уголки поднимаются вверх — улыбка.

Исходя из проанализированных графиков выделим следующие параметры:

- **Большая расстояние (>40 пикселей) между верхней и нижней губой:** невнимательность, слушатель, вероятнее всего, зевнул
- **Увеличение расстояния от края до центра + поднятие уголка вверх:** улыбка, зрителю нравится материал, который он видит, хотя, возможно, он был чем-то отвлечён
- **Увеличение расстояния от края до центра + уголка вниз:** уголки опускаются, зрителю скучно, его не устраивает просмотренный материал

Эпилог. Итоги и выводы

По итогам исследования было выделено 16 параметров, помогающих по набору кадров определить реакцию зрителя на обучающее видео. Большинству параметров приведена количественная (в пикселях) привязка, объяснено, как он интерпретируется с точки зрения эмоций и "языка тела" и какая ситуация в экспериментальном видео отвечала данному параметру.