

СЧИТАЙТЕСАМИ:

эффективная проверка видео с
участковых веб-камер

<http://считайтесами.рф>

<http://schitaytesami.org>

schitaytesami@gmail.com

15.02.2012

Цели проекта

- **Цель-минимум (этап 1):** для нескольких сотен подозрительных участков
 - Точно оценить расхождение **явки по протоколу** и **количества бюллетеней, опущенных в урну, согласно видео**
 - Выявить подозрительные события, связанные с опусканием бюллетеней.
- **Цель-максимум (этап 2):** приближенная оценка расхождений для максимально-возможного количества участков (вплоть до 90000) с приемлемой точностью

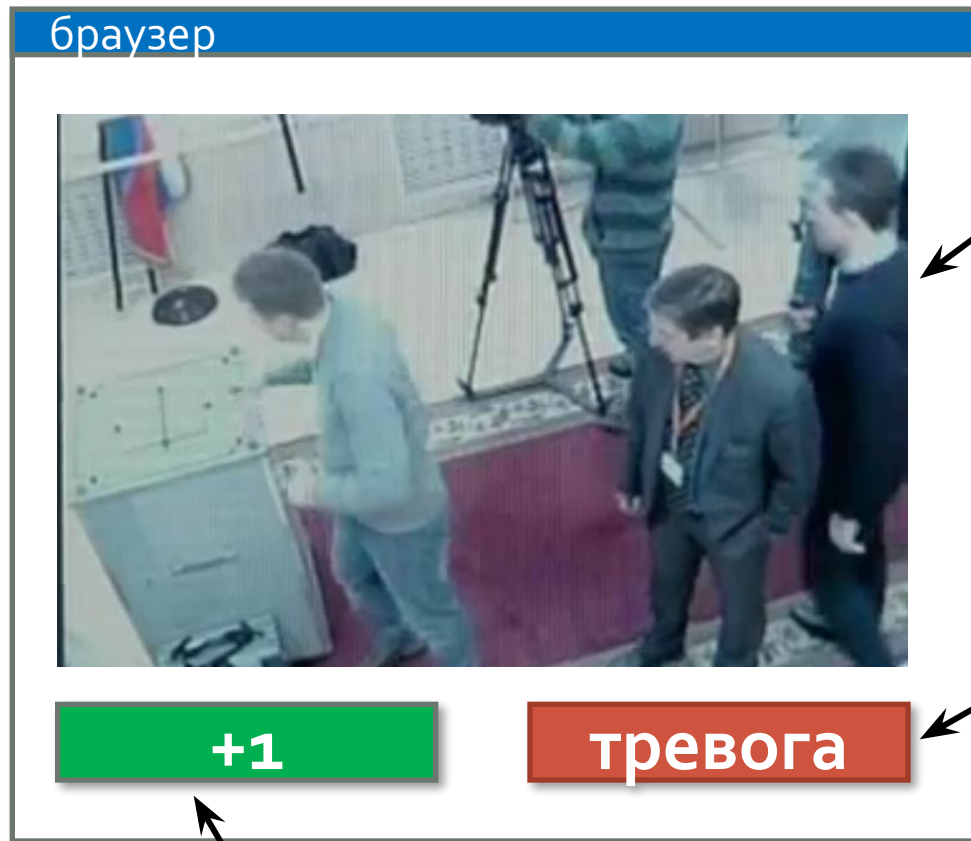
Данная презентация описывает способ достижения цели-минимум (этап 1)

Почему это важно

- Наиболее сильные искажения в результаты голосования вносят следующие нарушения:
 - Переписывание протокола на стадии прохождения через ТИК
 - Вброс бюллетеней до начала или после окончания голосования
 - Вброс в процессе голосования (один человек одновременно вбрасывает стопку бюллетеней).
- Все эти нарушения приведут к **расхождению** между явкой по протоколу и количеством "актов опускания бюллетеней" на видео. Фиксация таких расхождений и есть цель проекта.
- Ряд нарушений ("карусели", многократное вбрасывание по одному бюллетеню одним и тем же человеком) зафиксированы не будут. Однако по степени влияния они уступают нарушениям, перечисленным выше.

Система для разметки

На этапе 1, в основе системы лежит "ручной" подсчет количества опусканий. Разметка осуществляется в браузере. Схематично интерфейс выглядит так.



Окно, для проигрывания записи видео с подозрительного УИК

Кнопка для использования в нештатных ситуациях (например, очевидный вброс стопки)

Кнопка, нажимаемая единожды при каждом опускании бюллетеня. Может быть продублирована клавишей клавиатуры или кликом непосредственно по окну с видео

Обе кнопки фиксируют временной момент каждого нажатия для последующей обработки.

Ускорение видео

Проблема: для обработки каждого участка необходим просмотр 12-ти часового видео.

Решение: технологии компьютерного зрения позволят сжать **12 часов** до **0.5-2 часов** (в зависимости от числа проголосовавших) за счёт ускоренной перемотки частей видео, где никто не подходит к урне.

Пример обработки видео:

<http://www.youtube.com/watch?v=AEjQI77-OIs>



Ускоренная перемотка предпочтительна вырезанию, поскольку позволяет пользователю корректировать ошибки компьютерного зрения.

Распределение и контроль качества

Проблема: даже после сжатия, просмотр 0.5-2 часов видео на каждый УИК – неподъемная задача для ограниченного круга людей

Решение: каждое видео разрезается на небольшие куски и выгружаются для разметки на сайте в Интернете. С помощью социальных сетей большое количество добровольцев привлекается к разметке этих частей.

Проблема: возможно снижение качества из-за участия ленивых/невнимательных разметчиков или злоумышленников.

Решение: каждый кусок видео размечается 3 или более раз случайно выбранными участниками. Разметки, произведенные разными людьми объединяются в одну. Пропуски и очевидно-неправильные разметки (от злоумышленников) автоматически выявляются и исключаются на стадии объединения.

Производительность системы

Чтобы обработать каждый участок необходимо порядка:

1 час x 3 человека = **3 человека-часа**

↑
Ожидаемая
продолжительность
видео после сжатия

↑
Каждую часть видео нужно
покрыть троекратной разметкой

Количество добровольцев и количество времени, проводимого за разметкой оценить сложно. Всё же, предположим, что будет вовлечено 10000 участников, и каждый участник потратит в среднем полчаса. Таким образом, "трудозатраты" составят:

10000 человек x 0.5 часа = **5000 человеко-часов**

Таким образом, на первом этапе можно будет обработать до 1500 участков.

"Приватность"

Чтобы избежать претензий, связанных с нарушением privacy, всё видео будет предобработано перед показом разметчикам. А именно, всё пространство кадра, за исключением непосредственной окрестности прорези на урне, будет "размыто" до полной и очевидной неузнаваемости лиц.

Пример:



До предобработки



После предобработки

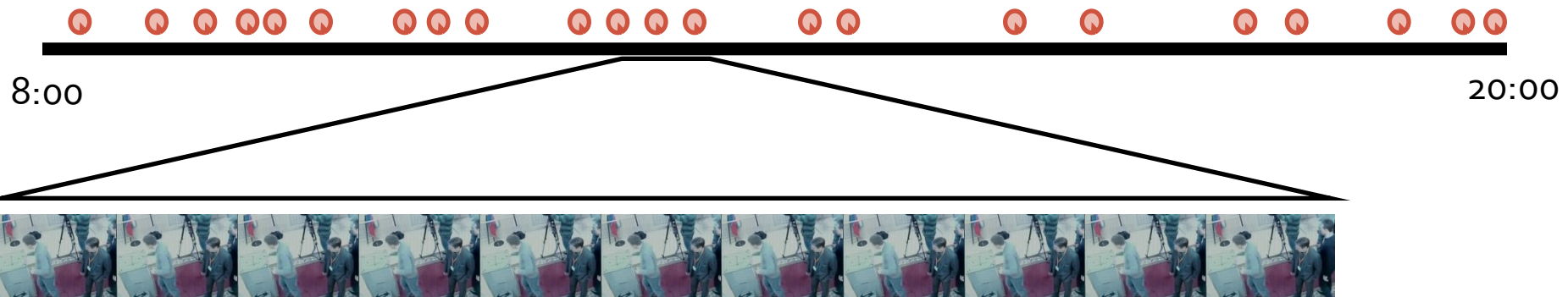
Визуализация результатов

Цель проекта: привлечение внимания к участкам с массовыми нарушениями, в том числе на УИК, как правило, не "покрытых" наблюдателями (сельские районы, национальные республики, и др.)

Ключевая задача: наглядная визуализация результатов

Идея 1: показ сверхускоренного видео (15-30 минут на участок). Каждый кадр снабжается количеством проголосовавших на данный момент. Моменты опускания выделяются тем или иным способом.

Идея 2: показ "ленты событий" и ленты кадров с возможностью прокрутки:



Дальнейшие идеи по визуализации приветствуются!

Требуемые ресурсы

- *"Лобби"*. Необходимо добиться получения видео с избирательных участков. Без этого проект невозможен.
- *Вычислительные ресурсы*. Для обсчета одного видео нужно минимум часов 12 компьютерного времени. Итого: на обсчет 1000 участков нужно 12000 машинных часов.
- *Емкости для хранения видео*. Для хранения одного видео нужно минимум 5 Гб места. Итого: на хранение и обработку видео с 1000 участков нужно 5000 Гб = 5 Тб.
- *Дополнительный человеческий ресурс*. Для запуска системы необходима разметка ограниченного количества неускоренного видео (несколько десятков часов). Данная разметка будет использована для тренировки системы компьютерного зрения.
- *Сетевые ресурсы*. Для ручной проверки результатов 10 людьми одновременно нужно иметь интернет-канал с пропускной способностью $10 \text{ мбит/сек} \times (5 \text{ Гб} / 24) / 1800 \text{ сек} \times 8$ (мегабит в мегабайте). Для одновременной проверке тысячи видео нужен 1 Гбит/сек. Если организовать работу распределенно, можно снизить требования на канал.

Текущая команда

Александр Гаген-Торн (powercrowd.ru, Россия)

Вадим Канторов (ENS Cachan, Франция)

Александр Кривошеков (powercrowd.ru, Россия)

Иван Лаптев (INRIA/ENS Paris, Франция)

Виктор Лемпицкий (Яндекс, Россия)

Вадим Лихолетов (powercrowd.ru, Яндекс, Россия)

**Проекту срочно нужны frontend/web-
программисты и верстальщики!**

schitaytesami@gmail.com

Примечание: все участники проекта участвуют в нём как частные лица. Их участие не подразумевает какой-либо поддержки проекта со стороны указанных организаций.