

# Определение интенсивности эмоций по видео

Александра Кудряшова

Научный руководитель: д.ф.-м.н. Стрижов Вадим Викторович

Московский Физико-Технический Институт  
Факультет управления и прикладной математики  
Кафедра интеллектуальных систем

Москва,  
2015

## Цель проекта

Построить онлайн систему для определения интенсивности эмоций в режиме реального времени путем обработки изображения лица человека и определения пульса по данным веб-камеры

## Новизна

Совместное использование распознавания выражения лица и определение пульса для уточнения определения эмоций

## Требования к системе

- Работа в режиме реального времени
- Online
- Работа с изображениями размера 640\*480px
- Возможность внедрения в клиент-серверную архитектуру

Направление	Авторы, работа, год
роль эмоций в эффективности контента	Soleymani M., et.al.: Multimodal Emotion Recognition in Response to Videos, 2012
кодирование выражений лица	Saragih J., et.al.: Deformable Model Fitting by Regularized Landmark Mean-Shift, 2010
определение пульса	Liu, L., et.al.: Eulerian Video Magnification, 2014
методология определения выражения лица	Littlewort, G., et.al: Emotion detection using sub-image based features through human facial expressions, 2011
конкурирующее решение	Mariappan, M: FaceFetch: A User Emotion Driven Multimedia Content Recommendation System Based on Facial Expression Recognition, 2012

## Ввод

X - видео

## Вывод

Y - массив меток выражений лица

P - массив значений пульса

- Видео

$$X_{ijk} \in \overline{0, 255},$$

где  $k$  - номер кадра  
 $i$  - индекс по горизонтали  
 $j$  - индекс по вертикали

- Метки

$$Y_{mk} \in [0, 1],$$

где  $k$  - номер кадра  
 $m$  - индекс метки (0 - счастливый, 1 - грустный, 2 - сердитый, 3 - удивленный)

- Пульс

$$P_k \in \mathbb{Z},$$

где  $k$  - номер кадра

- Область лба

$$D_k = \{(i, j)\}$$

- Интенсивность зеленого канала

$$I_k \in \overline{0, 255}$$

$$I_k = \sum_{i, j \in D_k} (X_{ijk2}) / \| D_k \|^2$$

- Пульс

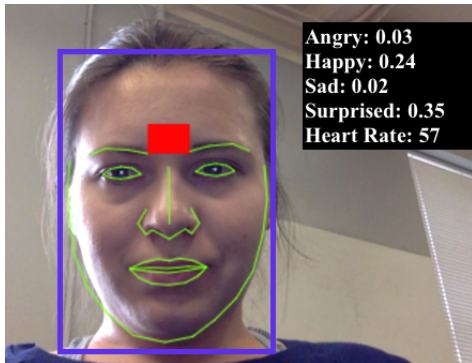
$$P_k = \operatorname{argmax}(FFT([I_{k-delay} \cdot I_k]))$$

- $Y_{mk} \in [0, 1]$ , - метки классов
- $b_1..b_t$  - базовые классификаторы
- $W^l = (w_1..w_l)$  - вектор весов классификатора
- $Q(b, W^l) = \sum_{i=0} w_i [y_i b(x_i) < 0]$  - стандартный функционал качества для  $b$
- Задача: оптимизировать  $\alpha_t$  по  $E(z) = \exp(-z)$

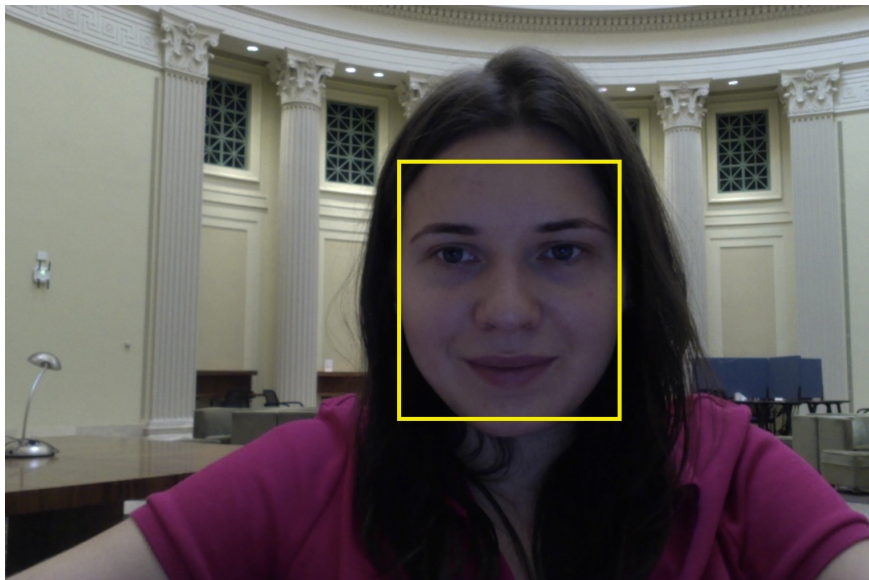
- Инициализация весов объектов:  $\phi_i = 1/l, i = 1, \dots, l$ ;
- Для всех  $t = 1, \dots, T$ , пока не выполнен критерий остановки:
  - Классификатор  $b_t : X \rightarrow \{-1, +1\}$ , минимизирует взвешенную ошибку  
 $b_t = \operatorname{argmin} Q_b(b, W^l)$ ;
  - Пересчитываем коэф. взвешенного голосования для  $b_t$ :  
$$\alpha_t = \frac{1}{2} \ln \frac{1 - Q(b, W^l)}{Q(b, W^l)}$$
  - Пересчет весов:  $w_i = w_i \exp(-\alpha_t y_i b_t(x_i)), i = 1, \dots, l$
  - Нормировка весов:  $w_0 = \sum_{j=1}^l w_j; w_i = w_i/w_0, i = 1, \dots, l$
- $\alpha(x) = \operatorname{sign}(\sum_{i=1}^T \alpha_i b_i(x))$



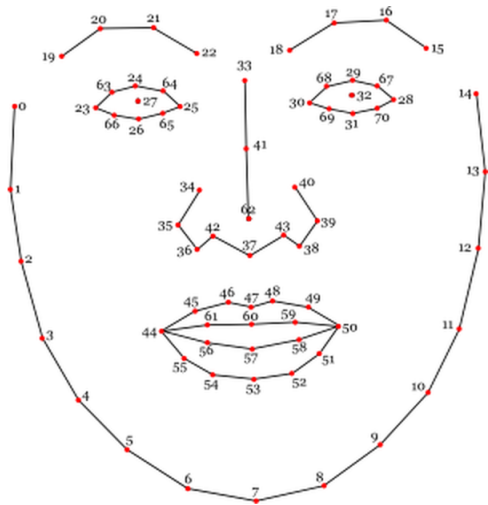
- 1 Обнаружить лицо
- 2 Кодировать выражение лица
- 3 Измерить пульс
- 4 Классифицировать эмоцию
- 5 Проверить корреляцию между пульсом и интенсивностью эмоций



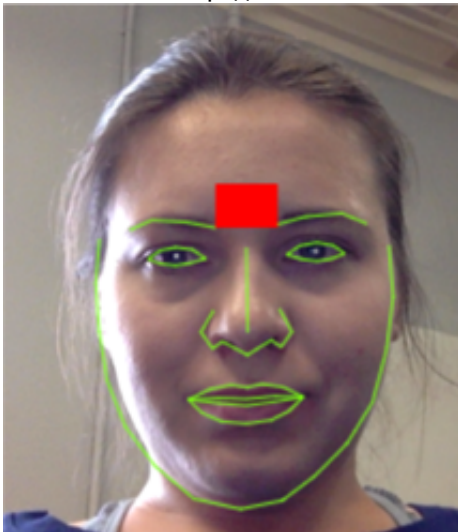
# Определение лица на изображении (Viola, Jones, 2001)



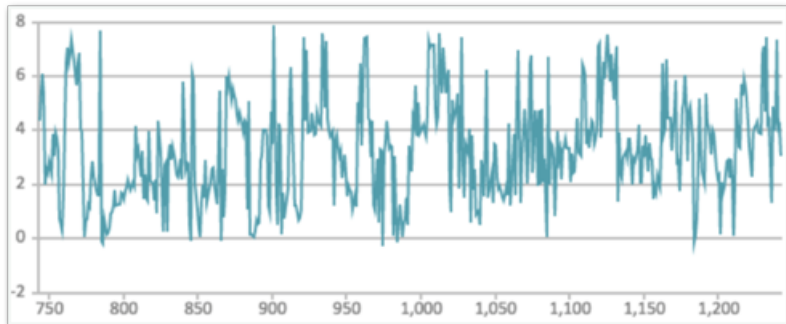
# Кодирование выражений лица (Oygar, 2011)



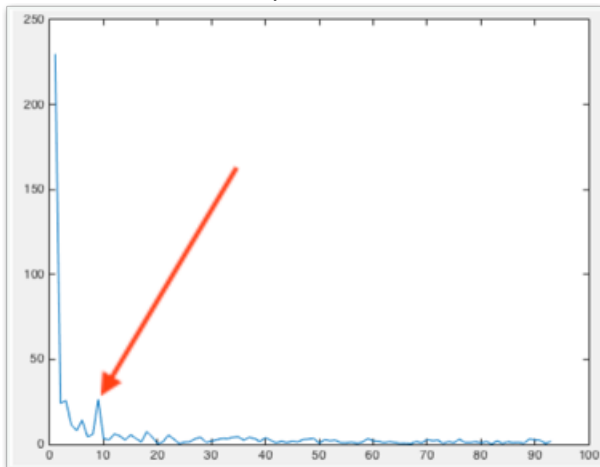
## Шаг 1: определение лба



## Шаг 2: Измерение интенсивности зеленого канала



## Step 3: FFT

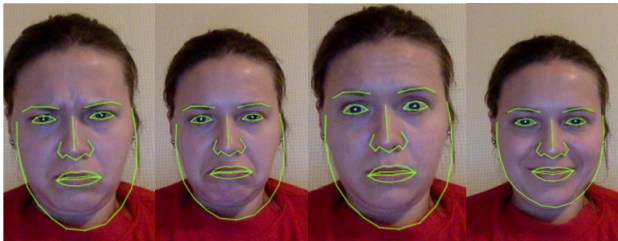


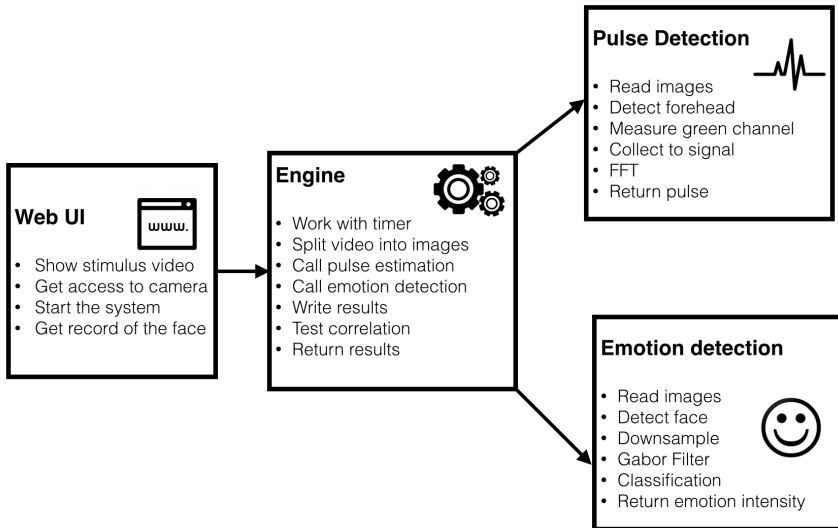
# Классификация эмоций (Saragih, 2012)

- типичные изображения для каждого из классов: сердитый, грустный, удивленный, счастливый



- различие маски







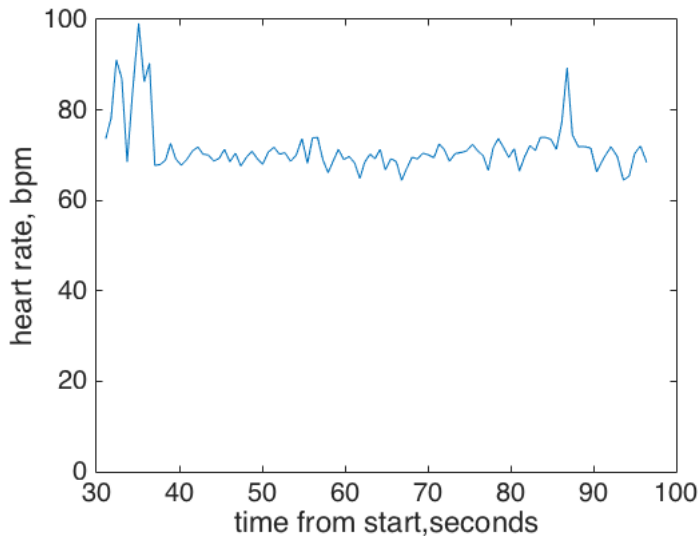
## Цель эксперимента

Измерить интенсивность эмоций, частоту сердцебиения и проверить корреляцию между ними.

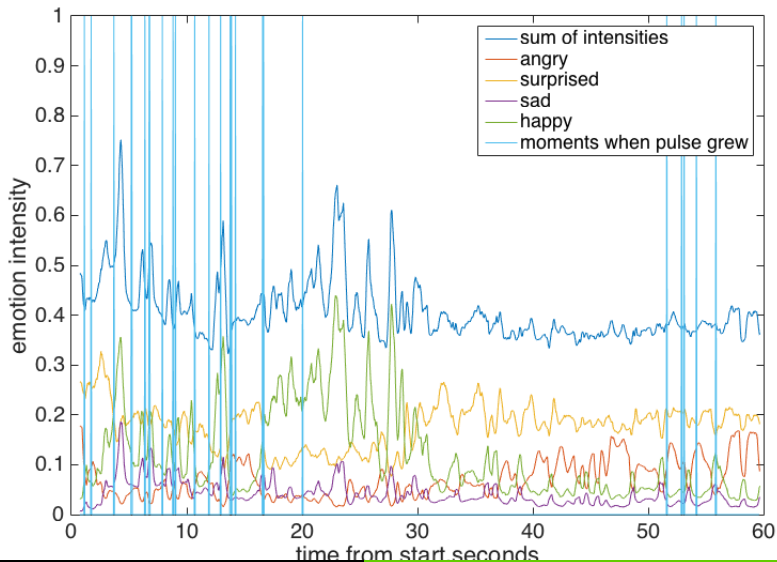
## Образец данных

Переменная	Пример	Единицы измерения
сердитый	$0.062 \pm 0.014$	[0..1]
грустный	$0.202 \pm 0.014$	[0..1]
удивленный	$0.034 \pm 0.014$	[0..1]
счастливый	$0.093 \pm 0.014$	[0..1]
время	100.92	секунд после запуска программы
пульс	$55 \pm 10$	ударов в минуту

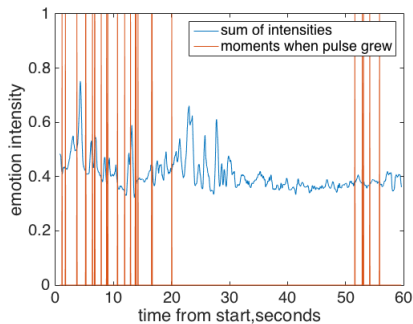
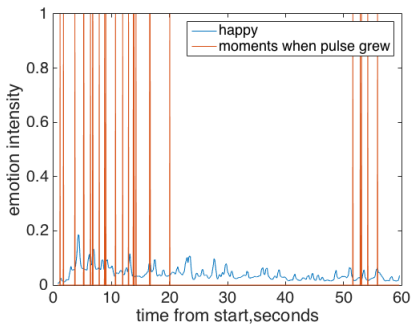
## Результаты: измерение пульса



# Результаты: распознанные эмоции и рост значения пульса



# Результаты: счастливый и сумма всех



$$r = \frac{\sum_m \sum_n (y_{mn} - \bar{y})(p_{mn} - \bar{p})}{\sqrt{(\sum_m \sum_n (y_{mn} - \bar{y})^2)(\sum_m \sum_n (p_{mn} - \bar{p})^2)}}$$

№	Сердитый	Грустный	Удивленный	Счастливый	Сумма
1	0.21	-0.31	-0.20	-0.21	0.36
2	0.20	-0.10	-0.15	-0.32	0.34
3	0.30	0.20	-0.20	0.18	0.29
4	0.00	-0.01	-0.00	-0.02	0.18
5	0.00	-0.02	0.005	0.03	0.14
6	0.00	-0.01	-0.00	0.02	0.11

- Прототип реализован на JavaScript и доступен по адресу <http://bit.ly/1Q98wt1>
- Система была протестирована на фокус-группе компании e-Contenta, которая занимается аналитикой медиа-контента
- Была получена договоренность о продолжении разработки и включении в их систему анализа эффективности контента



- Была спроектирована и построена система для определения и анализа интенсивности эмоций для проведения маркетинговых исследований
- Проверена возможность реализации, подобраны оптимальные параметры и протестирован инструмент для определения пульса по видео записи.
- Был реализован алгоритм определяющий интенсивность эмоций по атомарным выражениям лица.
- Проверена корреляция между пульсом и интенсивностью эмоций
- Система подготовлена к интеграции в существующую среду

## Launch results: some facts

c0	c1	c2	c3	c4	c5	c6	c7	c8	c9
1	67	1596	surp	0.10	sad	0.35	100	69	66
2	66	1577	happy	0.43	happy	0.43	135	52	45
3	59	1388	surp	0.21	happy	0.49	125	83	78
4	36	854	surp	0.31	happy	0.70	127	81	70
5	61	1475	surp	0.08	happy	0.95	115	60	45
6	55	1319	surp	0.35	happy	0.44	101	51	49
7	61	1505	surp	0.25	happy	0.56	112	75	75
8	62	1502	surp	0.09	happy	0.57	121	75	45
9	60	1431	surp	0.26	happy	0.82	125	87	99
10	59	1459	surp	0.09	happy	0.67	135	57	57

c0 - sample number, c1 - duration, c2 - number of records

c3 - most frequent emotion, c4 - its maximum value

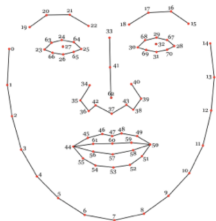
c5 - absolute maximum value, c6 - its emotion

c7 - maximum heart rate, c8 - mean heart rate, c9 - most frequent



# Feature extraction: encoding facial expressions

AU	Points of interest	Coordinate change	Class label
Inner brow raise	22, 18	[0; +2]	surprise
Outer brow raise	20, 16	[0; +2]	surprise
Brow corrugator	21, 17	[0; +2]	sad
Upper lid raise	47	[0; +2]	happy
Cheek raise	1, 13	[0; +2]	happy
Lower lid tight	53	[0; +2]	sad
Nose wrinkle	33	[0; +2]	angry



1 Initialize the weight  $w_{ij}$  ( $i = 1, \dots, m$  and  $j = 1, \dots, K$ ):

$$w_{ij} = \begin{cases} 0 & y_i = u_j \\ 1 & \text{otherwise} \end{cases}$$

2 For  $t = 1, \dots, T$ :

a. Normalize  $w_{ij}$ .

b. Train  $h_t(x)$  by minimizing loss function:

$$L = \sum_{i=1}^m \sum_{j=1}^K w_{ij} \exp((u_j - y_i) \cdot h_t(x_i)) \quad (1)$$

c. Update the weight matrix  $w_{ij}$ :

$$w_{ij} \leftarrow w_{ij} \exp((u_j - y_i) \cdot h_t(x_i)) \quad (2)$$

3 Final classifier:

$$H(x) = \sum_{t=1}^T h_t(x)$$