



Распознавание эмоций, вызываемых музыкой, на основе анализа оптического потока изображений лица человека

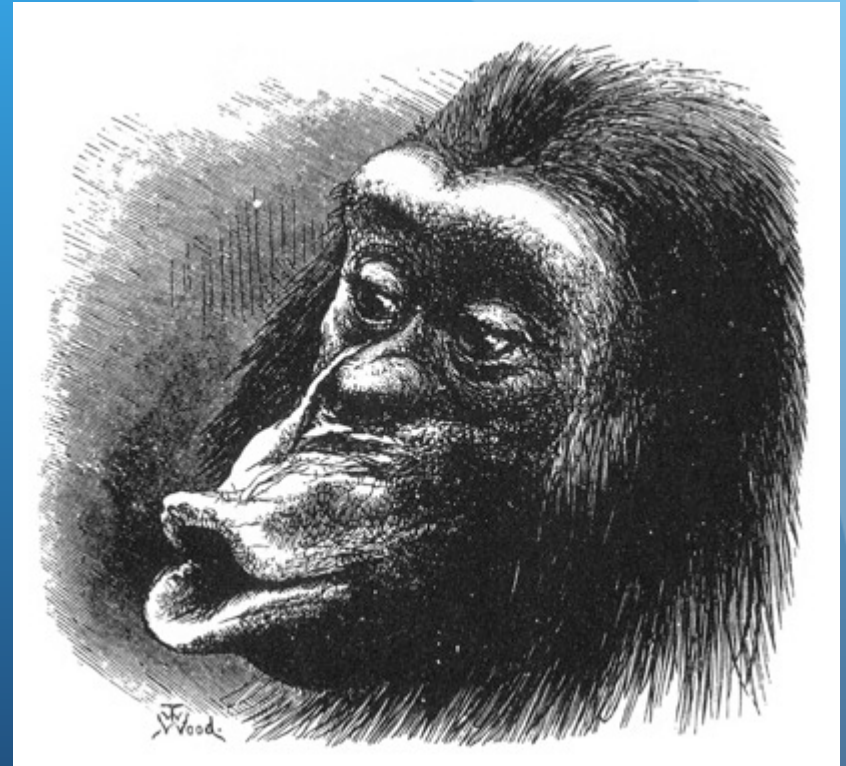
Князь В.В.

ГосНИИ Авиационных Систем (ГосНИИАС), Москва, Россия

Задача распознавания эмоций

Эмоциональный интеллект—
“способность человека
распознавать эмоции,
понимать намерения,
мотивацию и желания других
людей и свои собственные, а
также способность управлять
своими эмоциями и эмоциями
других людей в целях
решения практических
задач.”

‘Imagination, Cognition, and
Personality’, Саловей, П.,
1990



Мрачный и расстроенный шимпанзе
иллюстрация из книги «Выражение
эмоций у людей и животных»,
Чарльз Дарвин, 1872 г.

Выражение эмоций у Гибралтарских макак



Расслаблен

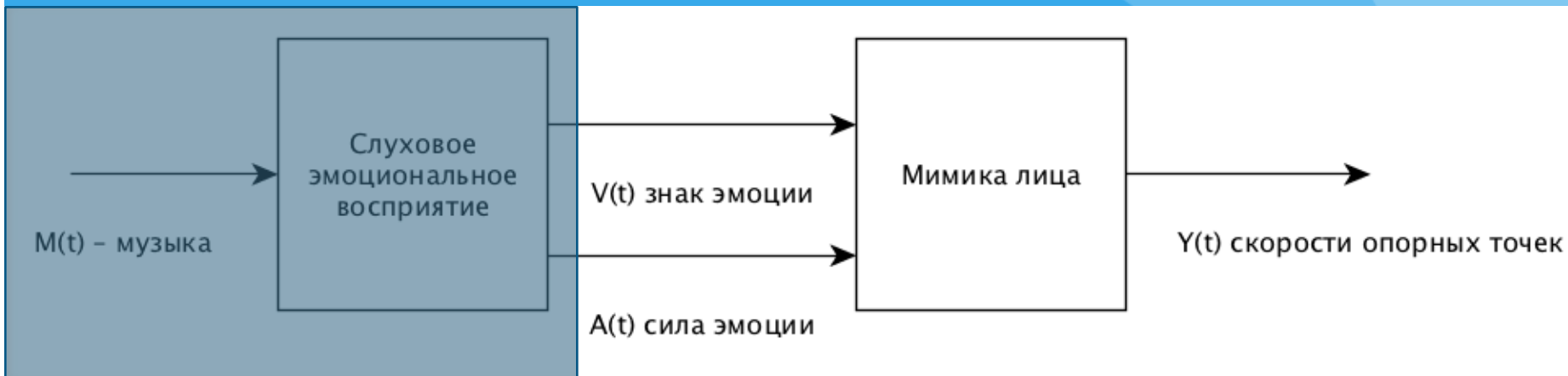
Разъярённый

Испуганный

Отрицающий

Фото: Зоопарк Коркеасаари, Хельсинки, Финляндия

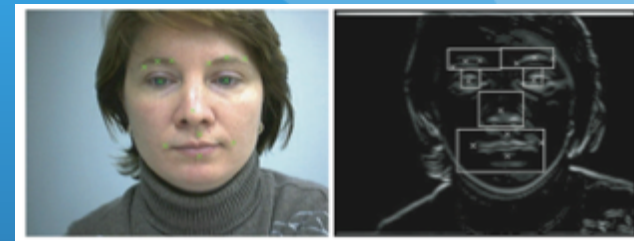
Постановка задачи



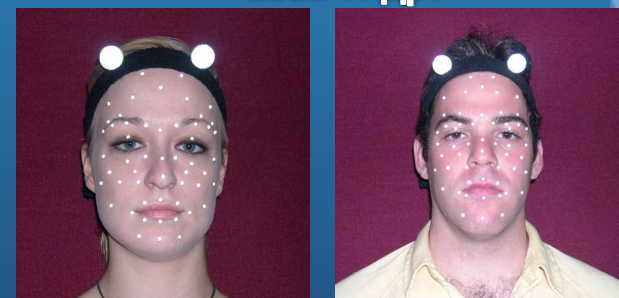
- Имеется последовательность изображений лица человека, слушающего музыкальное произведение
- Необходимо распознать эмоции, используя математическую модель, устанавливающую соответствие между эмоциями и мимикой человека
- Мимика распознаётся методами технического зрения
- Динамическая модель мимики определяется методами идентификации систем

Работы в данной области

- Kita S., Mita A., 2015. Распознавание эмоция по цвету лица человека / Emotion Identification Method using RGB information of Human Face
- Esau, N., Wetzel, et. al. 2007. Распознавание эмоций в реальном времени на основе нечёткой логики / Real-Time Facial Expression Recognition Using a Fuzzy Emotion Model
- Busso C., Bulut M., et.al, 2008. IEMOCAP: База данных захвата эмоциональных движений, вызванных диалогом / IEMOCAP: Interactive emotional dyadic motion capture database



Модель эмоций на основе нечёткой логики, Esau и др.



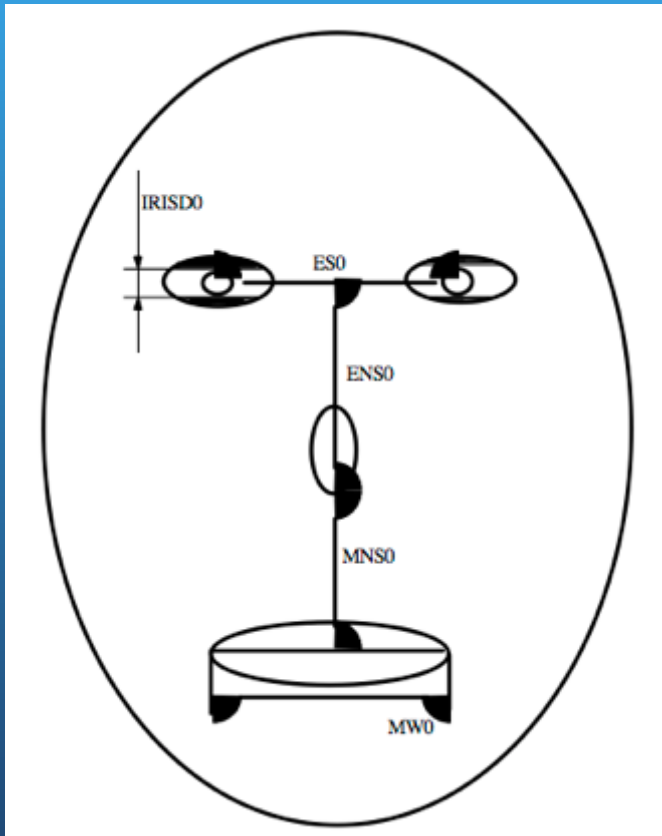
Запись эмоций с использованием системы захвата движения, Busso C., Bulut M., и др.

Пространство эмоций «знак-сила»

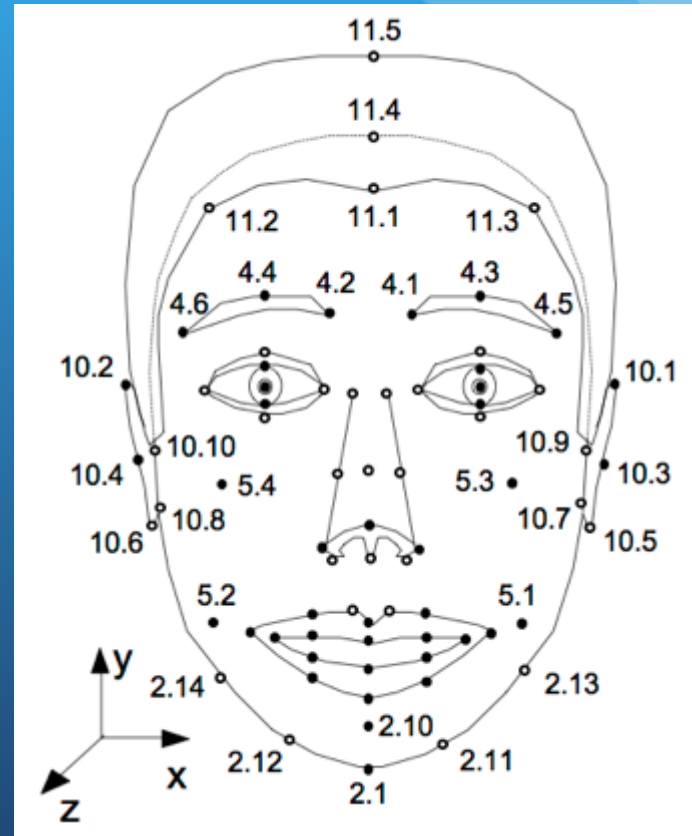
- «Круговая модель» или пространство эмоций «знак-сила» (З-С) была предложено Джеймсом Расселом в 1980г.
- Все эмоции представляются в двумерном пространстве
- Абсцисса - знак эмоции - положительное или отрицательное отношение человека к чему-либо
- Ордината - сила эмоции



Международный стандарт анимации лица и тела MPEG-4



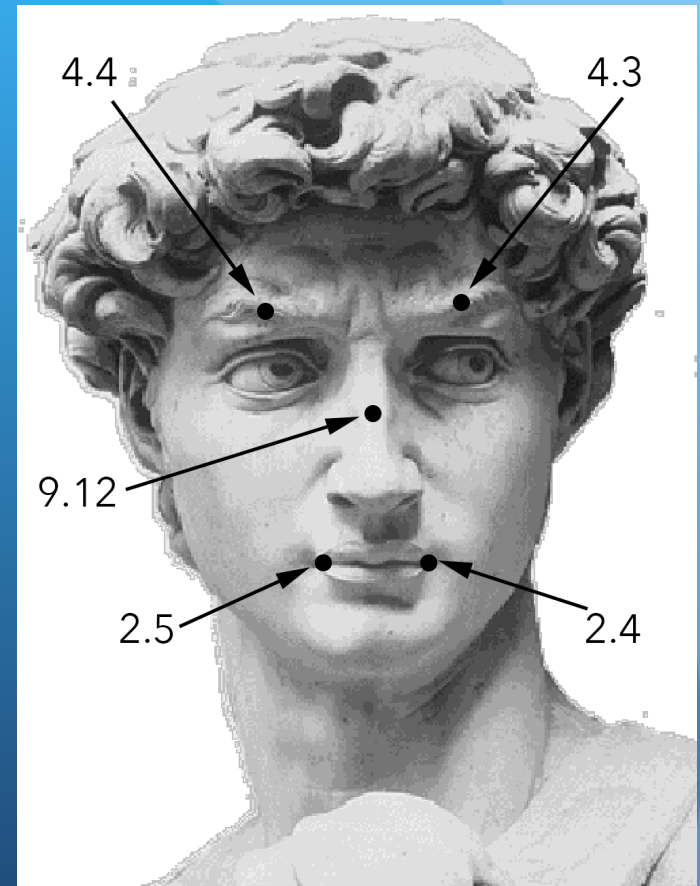
Единицы измерения анимации лица



Опорные точки анимации лица

Использованные параметры анимации лица человека

FAР №	Название параметра анимации	№	Название опорной точки
49	Поворот головы	9.12	Середина нижнего края носовой кости
50	Наклон головы	9.12	Середина нижнего края носовой кости
6	Горизонтальное смещение левого угла губ	2.4	Левый угол внутреннего контура губ
7	Горизонтальное смещение правого угла губ	2.5	Правый угол внутреннего контура губ



Вектор состояния анимации лица

- Для представления мимики лица используется двумерный вектор:

$$Y(t) = \begin{bmatrix} v_h \\ v_l \end{bmatrix}$$

$$v_h = |\varphi| + |\psi|$$

$$v_l = |v_{24}| + |v_{25}|$$

- v_h - модуль суммы угловых скоростей поворота ψ и наклона φ ГОЛОВЫ
- v_l - модуль суммы скоростей углов губ v_{24} , v_{25}
- В данной работе рассматриваются углы поворота и наклона головы в диапазоне $\pm 30^\circ$

Динамическая модель эмоциональной мимики

- Для записи динамики мимики лица используется классическая форма в пространстве состояний:

$$\dot{\mathbf{X}}(t) = \mathbf{A}\mathbf{X}(t) + \mathbf{B}\mathbf{U}(t)$$

$$\mathbf{Y}(t) = \mathbf{C}\mathbf{X}(t) + \mathbf{D}\mathbf{U}(t)$$

$$\mathbf{U}(t) = \begin{bmatrix} a \\ v \end{bmatrix} \quad \mathbf{Y}(t) = \begin{bmatrix} v_h \\ v_l \end{bmatrix}$$

- Вектор управления $\mathbf{U}(t)$ составлен из значений знака эмоции - v и силы эмоции - a

Методика сбора данных

- Для записи наборов данных для идентификации и проверки моделей были приглашены десять добровольцев, которым предлагалось прослушать музыкальные произведения

iPad с программой
'EmotionGrabber'



Доброволец



Музыкальное произведение

Мимика лица, удовольствие,
возбуждение

ЭВМ



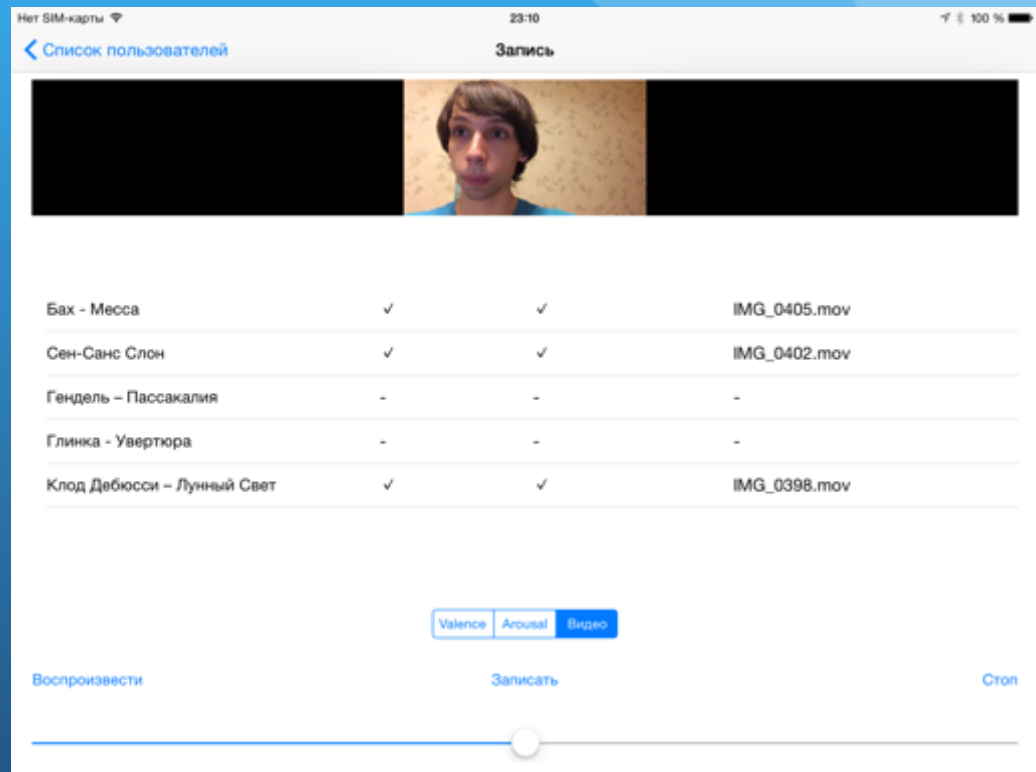
Данные

Идентификация
модели

$$\begin{aligned} \dot{X}(t) &= AX(t) + BU(t) \\ Y(t) &= CX(t) + DU(t) \end{aligned}$$

Функции программы записи эмоций и мимики «EmotionGrabber»

- Запись видеопоследовательности с частотой кадров 30 Герц, синхронизованная с воспроизведением музыки
- Синхронизованная запись значений знака и силы эмоции, реализованная с помощью подвижного ползунка
- Ведение базы данных добровольцев



Вид разработанной программы
«EmotionGrabber»

Отобранные музыкальные произведения:

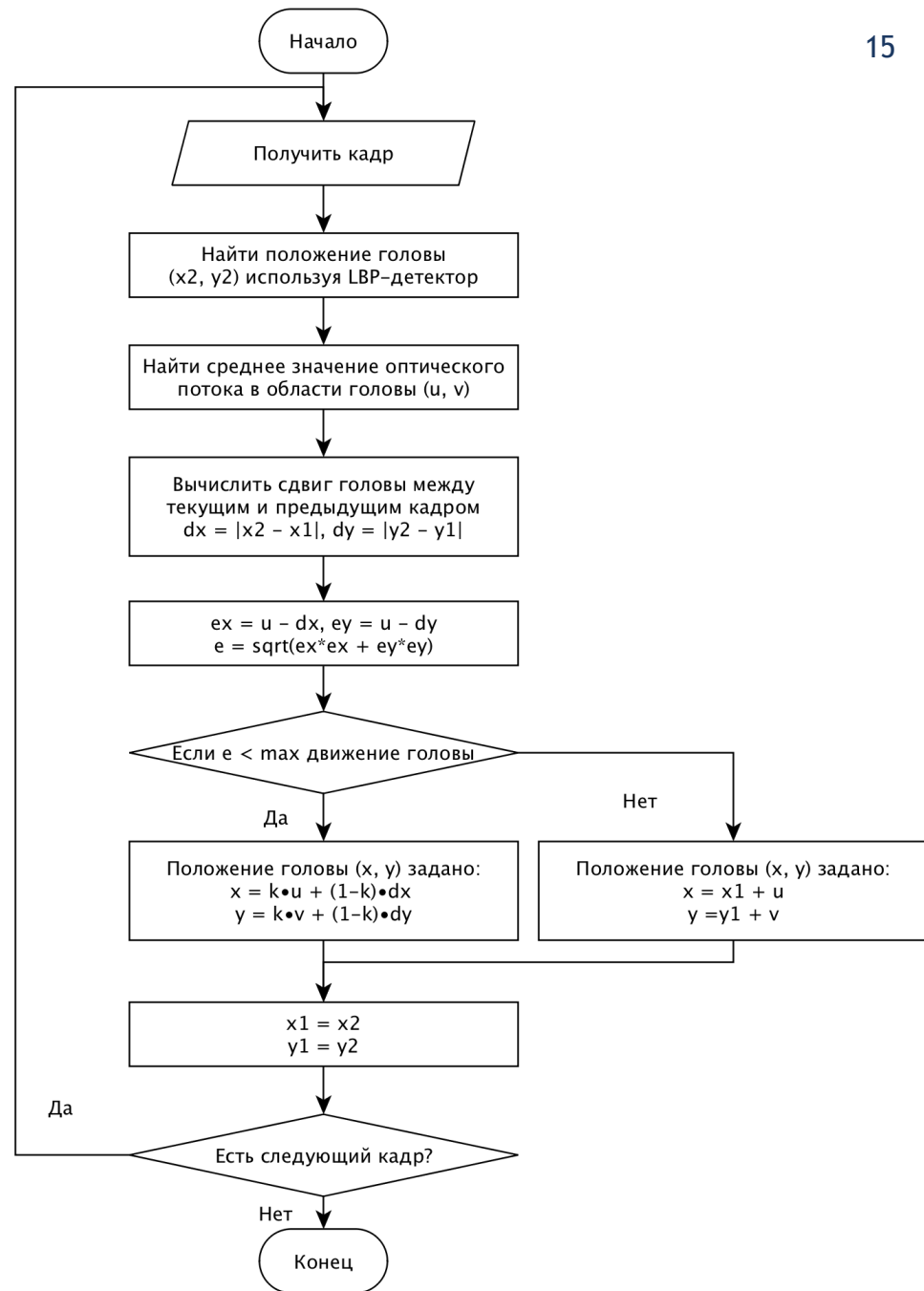
№	Музыкальное произведение	Композитор	Средняя эмоция	V знак	A сила
1	Пассакалья (№6), Сюита соль минор, BWV 432	Гендель Г.Ф.	Грустный	-0.8	-0.3
2	Месса си минор, BWV 232	Бах И.С.	Несчастный	-0.9	-0.1
3	«Lacrimosa» (Реквием)	Моцарт В.А.	Грустный	-0.8	-0.2
4	Мефисто вальс, №1, S.514	Лист Ф.	Встревоженный	-0.1	0.8
5	Увертюра к опере «Руслан и Людмила»	Глинка М.И.	Возбуждённый	0.6	0.6
6	«Поэма экстаза», Op. 54	Скрябин А.Н.	Исступление	0.1	0.9
7	«Остров мёртвых», Op. 29	Рахманинов С.В.	Напряжённый	-0.1	0.8
8	Карнавал животных, Часть V «Слон»	Сен-Санс Ш.К.	Счастливый	0.8	0.2
9	«Лунный свет»	Дебюсси К.	Спокойный	0.3	-0.5
10	Великая священная пляска из балета «Весна священная»	Стравинский И.Ф.	Испуганный	-0.3	0.7

Обработка данных

- Для обработки полученных наборов данных использовался модифицированный алгоритм прослеживания лица с помощью каскада классификаторов на основе локальных бинарных шаблонов (Cascade boosting LBP)
- Алгоритм был реализован в виде прикладного программного обеспечения для обработки собранных данных

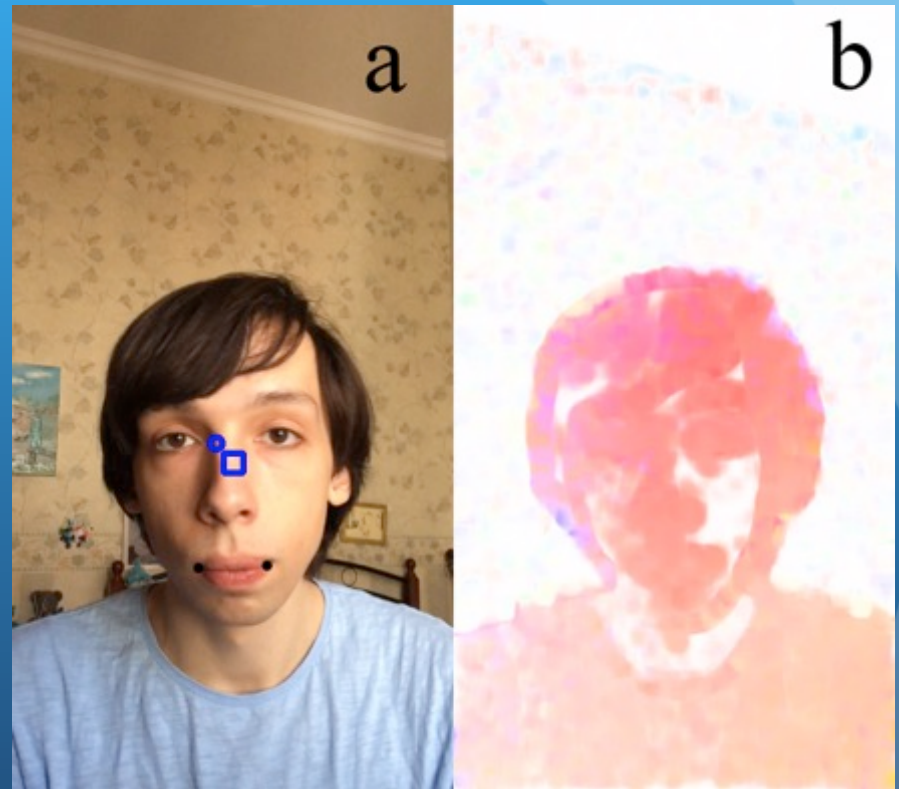
Модифицированный алгоритм прослеживания лица

- Положение лица на первом кадре определяется с использованием LBP детектора
- Для каждого последующего кадра для области головы рассчитывается оптический поток - плотная карта векторов скоростей пикселей последовательных кадров
- Оценка перемещения по оптическому потоку сравнивается с оценкой положения головы и использованием LBP детектора
- Если полученная разность меньше, чем заданное максимальное движение головы между кадрами, оценки положения головы по оптическому потоку и LBP детектору усредняются



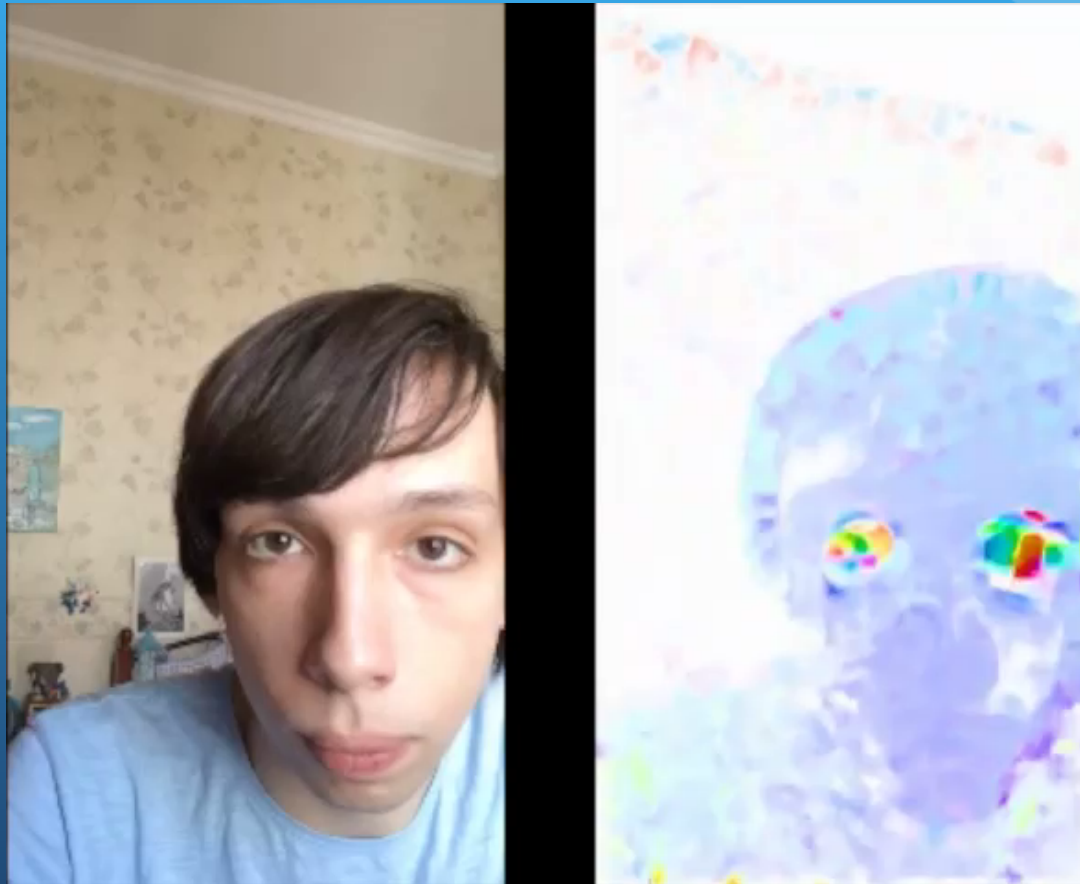
Прослеживание опорных точек на лице

- Расположение опорных точек на лице взято из стандарта MPEG-4
- Для уточнения положений точек на лице используется LBP-классификаторы, обученные на наборах изображений данных точек



Определение мимики лица с использованием LBP и оптического потока. (а) Найденные опорные точки. (б) Оптический поток в условных цветах (кодирование Мидлберри)

Демонстрация работы алгоритма поиска опорных точек



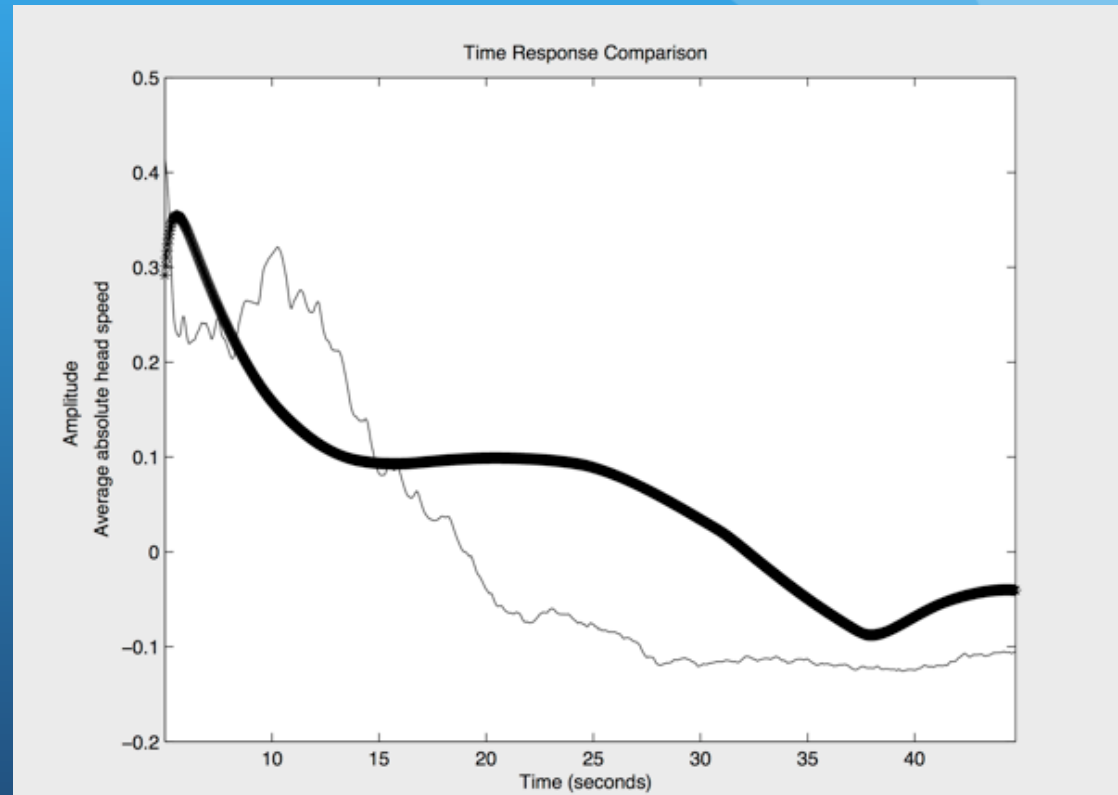
Музыка: «Карнавал животных», Часть V «Слон»,
Камиль Сен-Санс

Идентификация модели

- Два вида моделей использовались для представления связи между вектором управления (знак-сила эмоции) и мимикой лица:
 - линейные стационарные модели
 - нелинейные авто регрессионные модели wavenet (NARX)
- Для идентификации систем по полученным наборам данных использовался набор инструментов для идентификации систем пакета Matlab

Линейная стационарная модель

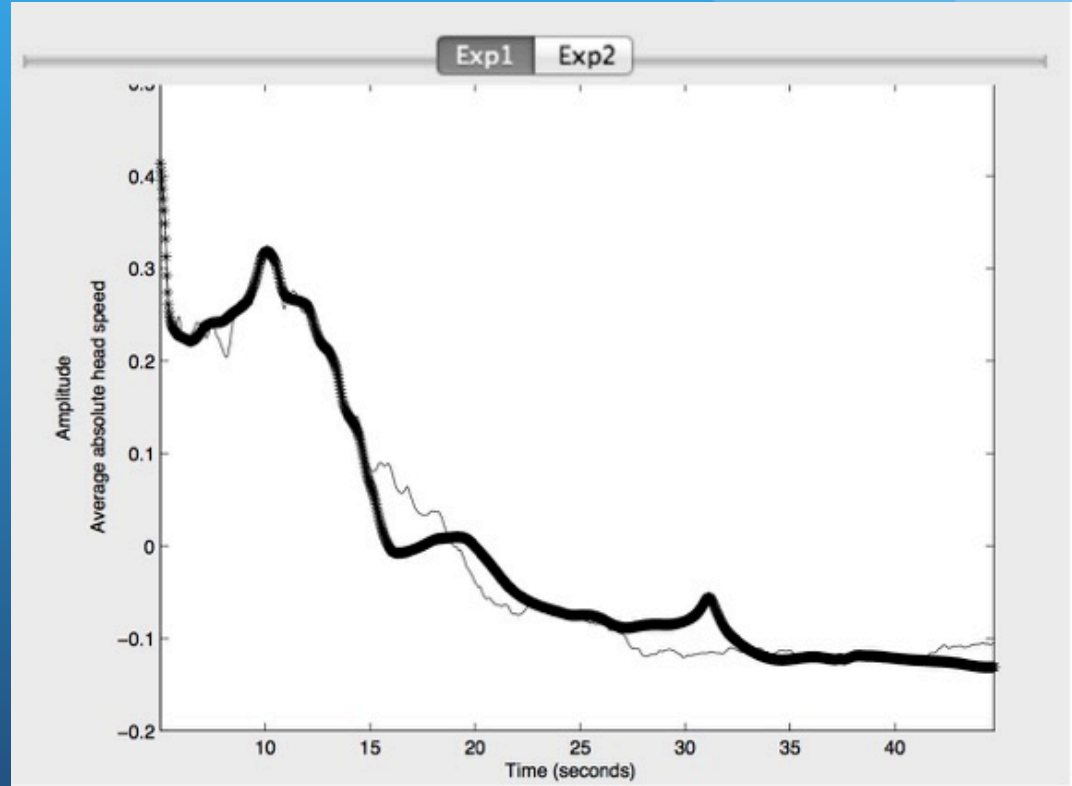
- Данная модель не смогла обеспечить высокой точности
- Нормированная среднеквадратическая ошибка составила 75%



Записанная средняя скорость движения головы и выход полученной модели

Нелинейная модель

- Для нелинейной авто регрессионной модели получена более высокая точность
- Нормированная среднеквадратическая ошибка составила 17%



Записанная средняя скорость движения головы и выход полученной модели

Выводы:

- Предложен метод идентификации динамической модели эмоций, вызванных музыкой
- Полученная модель может быть использована для восстановления мимики лица по записанным значениями вектора входа (знак-сила эмоции), например, для анимации 3D-персонажа
- Предложенный метод был опробован с использованием предварительно отобранного списка музыкальных произведений с разнообразной эмоциональной окраской
- Идентифицированная линейная стационарная модель оказалась недостаточно точной для описания изменений в мимике лица
- Нелинейная авто регрессионная модель обеспечила высокую сходимость с проверочными данными со среднеквадратической ошибкой 17%

Дальнейшая работа

- Оценка эмоции по наблюдаемой мимике лица с использованием расширенного фильтра Калмана
- Расширение базы данных записей добровольцев с использованием новых добровольцев и музыкальных произведений
- Оценка точности полученных моделей для музыкальных произведений, записанных в другом жанре
- Создание и оценка качества моделей с дополнительными параметрами анимации лица
- Исследование других пространств описания эмоций, таких как модель эмоций Плутчика и куб эмоций Лёвхейма



Спасибо за внимание!