

УДК 534.87

Д.С. Безыкорнов, А.И. Тюрин, М.А. Степаненко, Э.С. Соколова

## АНАЛИЗ ЭФФЕКТИВНОСТИ АЛГОРИТМОВ ПОИСКА КЛЮЧЕВЫХ ТОЧЕК НА ЛИЦЕ В УСЛОВИЯХ НЕДОСТАТОЧНОЙ ГРАФИЧЕСКОЙ ИНФОРМАЦИИ

Нижегородский государственный технический университет им. Р.Е. Алексеева

**Цель работы:** Анализ эффективности двух различных алгоритмов поиска ключевых точек: активных моделей внешнего вида (ААМ) и ограниченных локальных моделей (CLM). Определение наиболее эффективных условий применения каждого алгоритма.

**Методология:** Тестирование проводилось на созданном наборе изображений, приближенном к реальным условиям распознавания. Данные изображения отбирались с максимально возможными различиями. Качество распознавания проверялось на пяти категориях: технические характеристики изображения, частичное перекрытие лица, поворот лица, индивидуальные особенности (борода, очки и т.д.), активное проявление эмоций.

**Результаты:** Определены условия применения каждого алгоритма: в хороших условиях (без перекрытия, человек смотрит в камеру и др.) лучший выбор алгоритм CLM. Если системе не могут быть предоставлены хорошие условия или необходимо применение в системах реального времени, следует использовать алгоритм ААМ.

**Последующая работа:** Опираясь на полученные выводы, дальнейшие исследования могут быть проведены для построения более стабильной системы, основанной на CLM или ААМ модели.

**Оригинальность:** Данная работа позволила выбрать алгоритм ААМ для дальнейшего улучшения и использования в построении реальной системы распознавания эмоций.

*Ключевые слова:* ключевые лицевые точки, эффективность алгоритмов, комплексное тестирование, активные модели внешнего вида, ограниченные локальные модели.

### Введение

Базовой операцией систем автоматического распознавания эмоций является поиск лиц на изображении и дальнейшее распознавание ключевых точек в заданной области. Количество ошибок на данном этапе критически сказывается на качестве распознавания всей системы в целом, что недопустимо в области обеспечения национальной безопасности, как самой распространенной сфере применения этих систем.

Большинство существующих систем строятся для определенных «идеальных условий» и не подходят для практического использования. Например, системы аутентификации по изображению лица требуют взгляда непосредственно в камеру и хорошего освещения, наличие которых не всегда возможно.

Целью данной работы является оценка эффективности работы алгоритмов распознавания ключевых лицевых точек в условиях, приближенным к реальным. Для систем обработки реального времени будем рассматривать следующие параметры:

- вариативность вида и формы лица у различных людей;
- ориентация лица в пространстве;
- присутствие индивидуальных особенностей (борода, очки, веснушки и т.д.);
- эмоции на лице;
- перекрытие части лица различными предметами;
- технические особенности съемки (шумы, некачественное освещение, цветовой баланс и др.).

Данный набор охватывает большинство ситуаций на изображениях и позволяет использовать результаты анализа для гибкого выбора алгоритмов в различных системах.

Предварительное исследование существующих алгоритмов позволило исключить

заведомо бесперспективные для решения задачи распознавания эмоций и выделить два различных подхода к наиболее успешному поиску ключевых точек на лице. Это алгоритм активных моделей внешнего вида и алгоритм ограниченных локальных моделей.

### Активные модели внешнего вида

В основе алгоритма активных моделей внешнего вида (ААМ - active appearance models) используются статистические модели изображений, которые путем особых преобразований, называемых деформацией, накладываются на реальное изображение или видеопоток [1]. Деформация представлена в виде композиции переноса, поворота и масштабирования конкретной модели.

Алгоритм распространения созданной модели на анализируемое изображение выглядит так:

1. Вычислить параметры модели и аффинные преобразования формы исходного изображения;
2. Вычислить вектор ошибки  $\delta t$ . Извлечение текстуры из анализируемого изображения происходит с помощью его кусочной деформации [2];
3. Вычислить вектор возмущений  $\delta p = R\delta t$ ;
4. Обновить вектор комбинированных параметров и аффинных преобразований суммированием текущих их значений с соответствующими компонентами вектора возмущений [3];
5. Обновить формы и текстуры;
6. Выполнить последующие итерации, начиная с п. 2, пока не будет достигнута необходимая сходимость модели с реальным изображением.

Анализ математического обеспечения данного алгоритма позволяет предположить что при частичном перекрытии части лица и небольшом повороте статистическая модель сможет с достаточной точностью совпадать с входным изображением.

### Ограниченные локальные модели

В отличие от метода активных моделей внешнего вида алгоритм ограниченных локальных моделей (CLM - constrained local models) строит базовую модель на основе выборки ключевых точек по интенсивности пикселей соседней области с использованием различных фильтров Хаара [4, 5]. Основная идея алгоритма заключается в следующем:

- 1) выбрать некоторую область на изображении, расположенную предположительно около целевой точки с учетом проекции на изображение-пример;
- 2) для каждой точки сгенерировать изображение-отклик, записывая в него стоимость каждого пикселя (стоимость вычисляется исходя из совпадения по определенному набору параметров) [6];
- 3) найти такую комбинацию точек, которая является оптимальной по общей стоимости, изменяя при этом параметры модели.

Алгоритм поиска точек выполняется следующим образом:

1. Получить на вход начальный набор ключевых точек;
2. В цикле:
  - a) найти модель для текущего набора ключевых точек, чтобы создать набор шаблонов;
  - b) используя метод поиска с учетом ограниченной формы модели, уточнять набор точек до тех пор, пока точки модели и тестируемого изображения не сойдутся.

При отслеживании начальные точки выбираются с предыдущего кадра. В случае неудачи можно использовать глобальный поиск с нуля [7].

Основным недостатком алгоритма для работы в сложных условиях является исполь-

зование фильтров Хаара: определенные шаблоны яркости будут хуже работать в условиях перекрытия части лица, что снижает эффективность применения данных алгоритмов на всем многообразии входных данных[8].

### Экспериментальная часть

Для исследования возможностей алгоритмов необходим набор изображений, отвечающий определенным требованиям. Стандартные наборы изображений (FERET, M2VTS) зачастую предоставляют качественные, заранее выверенные примеры, в которых не проявляется всего многообразия признаков. Большинство исследованных систем очень плохо реагируют на такие особенности, как темная кожа или веснушки на лице. В примерах они не указываются.

Для решения задачи комплексного анализа в условиях, приближенных к реальным, был составлен собственный набор данных, состоящих из 150 изображений. 50 из них были размечены ключевыми точками в полуавтоматическом режиме: с помощью алгоритма ААМ находилось приблизительное положение точки и после проводилась ее коррекция вручную. 100 изображений тестовой выборки помечались, как принадлежащие к одной из пяти категорий:

1) ориентация лица в пространстве. Лицо считалось повернутым при повороте более чем на 15 град;

2) выражение лица. Проявление эмоций сильно влияет на сходимость построенных моделей и реального изображения. В данной категории учитывался открытый при улыбке рот и закрытые глаза;

3) индивидуальные особенности. Наличие бороды у мужчин, темный цвет кожи, наличие веснушек или прозрачных очков также затрудняет распознавание. Если дужки очков не закрывают глаза, алгоритм все еще может распознать ключевые точки, поэтому очки с прозрачными стеклами отнесены в эту категорию;

4) частичное перекрытие лица. Самое частое – пышные прически у женщин, когда волосы закрывают глаза, а также темные солнцезащитные очки;

5) плохие технические характеристики. К ним относятся низкое качество изображения (шум, недостаточная резкость), большая разница в яркости частей лица (часто встречается, что половина лица затенена).

Характеристики тестового набора:

- общее количество изображений – 100;
- минимальное разрешение – 278 x 541;
- максимальное разрешение – 1280 x 853;
- цветные изображения – да.

Количество изображений в каждой категории приведено в табл. 1.

**Таблица 1**

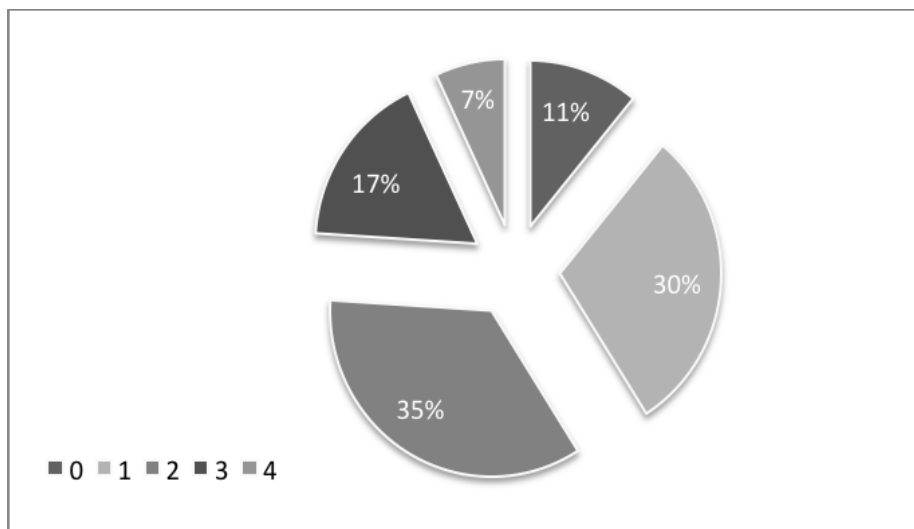
**Количество изображений в категориях**

Поворот лица	Эмоции	Индивидуальные особенности	Перекрытие	Плохие технические характеристики
54	48	21	23	37

На диаграмме (рис. 1) визуальное представлено распределение изображений по количеству категорий.

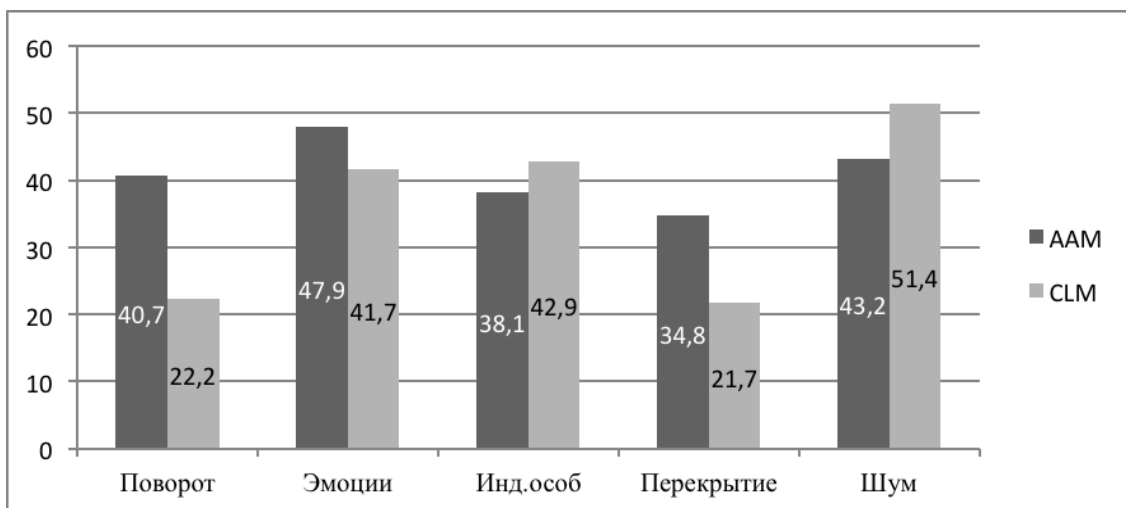
Как видно, большую часть выборки составляют изображения с первым или вторым сложными для распознавания признаками – 30 и 35% соответственно. Такие условия проверки алгоритмов позволяют провести комплексный стресс-тест.

Для алгоритма активных моделей внешнего вида было произведено обучение на обучающей выборке по 81 точке. Алгоритм ограниченных локальных моделей прошел обучение по 65 точкам. Такое различие связано с использованием для теста стандартных реализаций и особенностей алгоритмов. Основные различия в количестве точек, обозначающих глаз и очерчивании положения носа - не влияют на общие наборы для распознавания эмоций. В связи с этим был введен порог ошибочного определения положения лиц в 10% за счет меньшей важности локальных точек, очерчивающих овал лица, и некоторых других.



**Рис. 1. Диаграмма количественного распределения категорий изображений**

Результаты работы алгоритмов представлены на рис. 2.



**Рис. 2. Процентное соотношение распознанных лицевых изображений (порог ошибок 10%) по категориям**

Как видно из приведенного графика, алгоритм ААМ показывает эффективность в два раза выше на одной из самых распространенных категорий – повороте лица. Также отмечается сильное падение качества распознавания, если часть лица закрыта. На рис. 3 представлены качественные сравнения алгоритмов.

Анализ результатов нахождения ключевых лицевых точек позволяет наглядно удостовериться в статистических данных. Главными точками в задачах распознавания эмоций

являются края губ, глаз, бровей, точки на носу. В первую очередь для сравнения качества используются именно эти точки.

Рисунки 3 (1), 3 (5) подтверждают непригодность алгоритма CLM к условиям изменения ориентации лица в пространстве, данные наборы точек не позволяют использовать даже небольшую их часть для эффективного применения в последующих алгоритмах. Рисунок 3 (4) показывает уязвимость алгоритма ААМ к изображениям лиц с очками, что исходит из их теоретических основ, а конкретно – разных вариантов формирования базовой модели во время обучения. Фильтры Хаара, при условии прозрачности очков позволяют более точно выделить участки глаз, в отличие от полностью адаптивной модели ААМ.

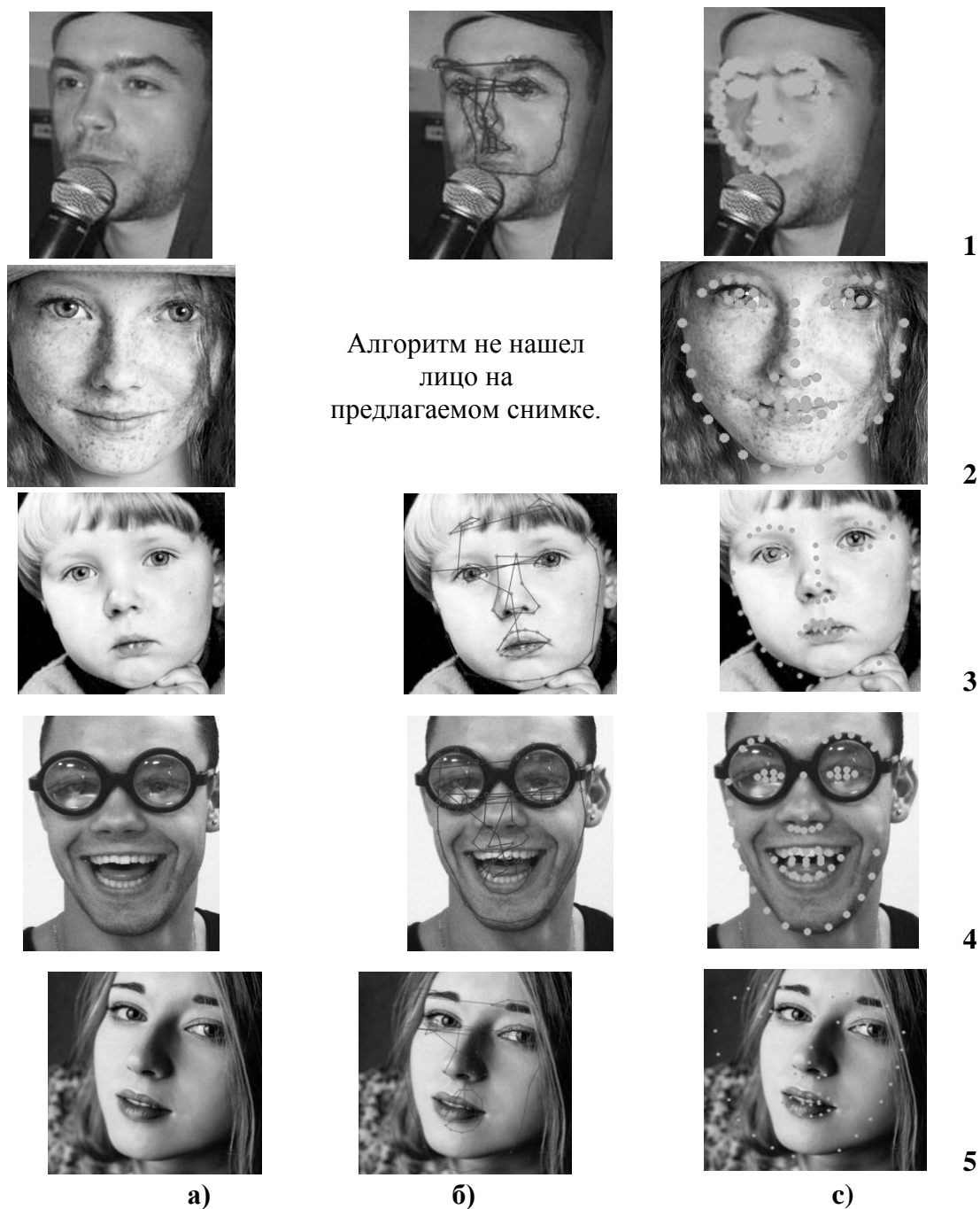


Рис. 3. Качественное сравнение работы алгоритмов:  
а – оригинал; б – ААМ; в – CLM

Перекрытие части лица на рис. 3 (1) и рис. 3 (5) представляет сложность для алгоритма CLM, однако не мешает алгоритму ААМ эффективно распознать верхнюю часть лица и оставить несколько главных точек в наборе с высоким качеством распознавания.

На основе полученных экспериментальных данных можно сделать предположения об оптимизации алгоритма ААМ как наиболее перспективного в большинстве случаев:

1. Для повышения качества распознавания сильно зашумленных изображений при тренировке базовой модели следует производить графическое выравнивание параметров изображений, исходя из среднего по обучающей выборке. Данный метод позволит свести к минимуму влияние освещения и различий в разрешении изображений

2. Увеличение скорости обработки за счет исключения из сравнения модели текстур компонентов точек базовой формы при достаточном количестве итераций должно быть достаточно для качественного распознавания. Это позволит увеличить скорость обработки, а также снизить влияние условий частичного перекрытия лица и наличия индивидуальных особенностей.

3. Добавление обновления базовой модели с учетом поступающих данных будет эффективно в условиях систем реального времени, что позволит повысить качество распознавания без увеличения нагрузки на вычислительную систему.

### Заключение

Проведенное тестирование алгоритмов поиска ключевых точек в условиях недостаточной графической информации позволяет определить области применения каждого из алгоритмов. Для систем распознавания эмоций в режиме реального времени эффективнее использование алгоритма ААМ. Средняя ошибка распознавания данного алгоритма на 5% ниже, чем показывает алгоритм CLM. Для систем с возможностью задавать условия съемки (прямой взгляд в камеру, открытое лицо без перекрытия) наиболее целесообразным является использование алгоритмов CLM. Данные наработки будут использованы в модификации алгоритма ААМ при построении реальной системы распознавания эмоций.

### Библиографический список

1. **Cootes, T.F.** Active appearance models / T.F. Cootes, G.J. Edwards, C.J. Taylor // Proc. European Conference on Computer Vision 1998. V. 2. P. 484-498
2. **Li, P.** Context-based additive logistic model for facial keypoint localization DOI: 10.5244/C.24.28 / P.Li [et al.] // British Machine Vision Conference, BMVC 2010. – Aberystwyth (UK), 2010.
3. **Cristinacce, D.** A comparison of shape constrained facial feature detectors / D. Cristinacce, T.F. Cootes // Proc. Int. Conf. on Face and Gesture Recognition, 2004. P. 375–380.
4. **Zhang, Z.** Facial Landmark Detection by Deep Multi-task Learning / Z. Zhang [et al.] // Proceedings of European Conference on Computer Vision (ECCV), 2014.
5. **Zhu, X.** Face detection, pose estimation and landmark localization in the wild / X. Zhu, D. Ramanan // Computer Vision and Pattern Recognition (CVPR) Providence, Rhode Island. 2012.
6. **Hasan, K.** Localizing Facial Keypoints with Global Descriptor Search, Neighbour Alignment and Locally Linear Models, DOI: 10.1109/ICCVW.2013.55 / K. Hasan, C. Pal, S. Moalem // Proceedings of the 2013 IEEE Int. Conf. on Computer Vision Workshops, 2013.
7. **Bolme, D.S.** Average of Synthetic Exact Filters / D.S. Bolme, B.A. Draper, J.R. Beveridge // Computer Vision and Pattern Recognition, 2009.
8. **Valstar, M. F.** Facial Point Detection using Boosted Regression and Graph Models / M. F. Valstar [et al.] // Proceedings of IEEE Int. l Conf. Computer Vision and Pattern Recognition (CVPR'10). – San Francisco (USA). 2010. P. 2729–2736.

Дата поступления  
в редакцию 02.07.2015

---

D.S. Bezykornov, A.I. Tyurin, M.A. Stepanenko, E.S. Sokolova

## KEY POINTS SEARCH ALGORITHMS EFFICIENCY ANALYSIS IN A LACK OF GRAPHICAL INFORMATION

Nizhny Novgorod state technical university n.a. R.E. Alexeev

**Purpose:** Efficiency analysis of two different face models in computer vision: AAM (Active-Appearance Model), CLM (Constrain Local Model) in real cases and defining areas of use of these models.

**Design/methodology/approach:** We created image set with different pictures with real conditions. Recognition quality was checked in 5 categories: noise, partial occlusion, rotation, individual face landmarks (beard, glasses, etc.), emotions on the tested face.

**Findings:** We found different areas of using these face models: when system could be used in good conditions (no occlusion, humans look in camera, etc.) it better to use CLM model, in other case, when system cannot be set up in good conditions or if it needed to use system in real-time it better to use AAM.

**Research limitations/implications:** Further research could use our summary to create more stable system based on CLM or AAM model.

**Originality/value:** This research allow us to choose AAM model to further improvement and using in real system that can recognize emotions on a real face.

*Key words:* facial key points, algorithms efficiency, complex testing, active appearance models, constraint local models.