

УДК 004.932.72'1

Исследование алгоритма определения ракурса лица по 2D изображению

***Петрук В.И.**, аспирант*

*Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедры «Биомедицинские технические системы»*

***Прояев П.О.**, студент*

*Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедры «Биомедицинские технические системы»*

*Научные руководители: Спиридонов И.Н., д.т.н., профессор
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана*

Самородов А.В., к.т.н., доцент

Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана

Для формирования шаблона при решении задач верификации и идентификации личности, анализа мимических выражений лица, оценки эмоционального состояния по изображению лица, как правило, необходимо производить оценку ракурса лица. Лучшие методы оценки ракурса опираются на использование информации о локализации контрольных точек (КТ) на изображении лица [1]. В настоящее время существуют различные алгоритмы оценки положения объекта относительно системы координат устройства регистрации изображения. Все они требуют явного задания модели объекта в виде координат его характерных точек в трехмерном пространстве, а также таких характеристик камеры, как фокусные расстояния в миллиметрах и пикселях, размер ПЗС-матрицы и разрешение выходного изображения [2, 3, 4]. Тогда, если на изображении удастся задать соответствие между характерными точками трехмерной модели и точками на полученном изображении, то представляется возможным вычислить матрицу поворота и вектор параллельного переноса, совмещающие систему координат объекта с системой координат устройства регистрации изображения.

В настоящей статье приводятся результаты исследований алгоритма оценки ракурса лица по его 2D изображению с использованием алгоритма POSIT и параметрической модели лица CANDIDE-3.

Комплексы КТ и модели лица. Исходя из необходимости задания модели лица в виде координат его характерных точек в трехмерном пространстве, был проведен сравнительный анализ наиболее часто используемых в настоящее время комплексов КТ и

моделей лица. Среди них: комплекс КТ MPEG-4, определенный международным стандартом ИСО/МЭК 14496-2-2001, комплекс КТ, определенный в ГОСТ Р ИСО/МЭК 19794-5-2006, параметрические модели лица CANDIDE, статистические модели распределения КТ.

Из всех перечисленных комплексов КТ и моделей лица указанному выше требованию удовлетворяет только модель CANDIDE-3 [5]. Статистические модели распределения точек не позволяют явным образом задать взаимное расположение КТ и зависимость их координат от углов вращения объекта вокруг осей системы координат. Кроме того, параметризация статистической модели неочевидна и трудна для восприятия человеком [6, 7]. Комплексы КТ MPEG-4 и ГОСТ Р ИСО/МЭК 19794-5-2006, в отличие от модели CANDIDE-3, изначально не содержат в себе базовую модель лица с нейтральным выражением [8, 9].

Параметрическая модель лица CANDIDE-3. Первая версия параметрической модели CANDIDE, содержащая 75 вершин и 100 полигонов, была создана в 1987 году Микаэлем Ридфолком и настоящее время практически не используется. Позднее Мартеном Штромбергом модель была несколько модифицирована и включила в себя 79 вершин, 108 полигонов и 11 двигательных единиц (ДЕ) лица. ДЕ представляют собой движения, совершаемые отдельными мышцами или группой мышц лица.

В 1991 году Билл Уэлш разработал модель CANDIDE-2, дающую описание всей головы в целом (включая волосы и зубы) и плечей. Эта версия модели содержит 160 вершин, 238 полигонов, но лишь 6 ДЕ лица.

В 2001 году Йоргеном Альбергом была существенно доработана модель CANDIDE-1. Модель CANDIDE-3, показанная на рисунке 1, содержит 113 вершин, 184 полигона, 65 анимационных единиц, часть которых соответствует параметрам движения лица (ПДЛ) стандарта MPEG-4, и 14 единиц формы (ЕФ) лица. В настоящее время это наиболее полная и реалистичная с точки зрения строения лица версия модели CANDIDE (рис. 2) [5].



Рис. 1. Параметрическая модель лица CANDIDE-3

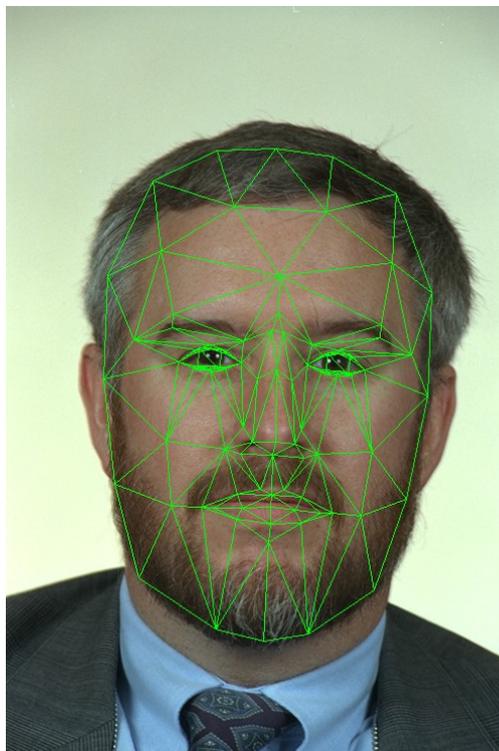


Рис. 2. Пример совмещения модели CANDIDE-3 с изображением лица

Параметры текстуры, формы и единиц действия мышц лица модели CANDIDE-3 введены так, чтобы достигнуть совместимости с ПДЛ и параметрами определения лица (ПОЛ) стандарта MPEG-4.

Простота моделей CANDIDE делает их хорошим инструментом для решения задач анализа изображений и анимации с низкой вычислительной сложностью. Модели CANDIDE имеют свой формат хранения данных, который поддерживается всеми ее версиями и содержит описание координат вершин (КТ), соответствующих средней форме

лица с нейтральным выражением, полигонов, ДЕ и ЕФ и прочих параметров в удобной для использования форме.

CANDIDE-3 представляет собой параметрическую модель лица. Начальная форма $\bar{\mathbf{g}}$ представляет собой $3N$ -мерный вектор координат КТ (вершин) модели, где N – количество КТ. Модель может изменять свою форму в соответствии с формулой

$$\mathbf{g}(\sigma, \alpha) = \bar{\mathbf{g}} + \mathbf{S}\sigma + \mathbf{A}\alpha,$$

где \mathbf{g} – обновленная форма модели, \mathbf{S} и \mathbf{A} – столбцы ЕФ и АЕ соответственно, σ и α – векторы параметров формы и анимации соответственно.

При учете преобразований вращения, масштабирования и переноса, формула принимает вид

$$\mathbf{g}(\sigma, \alpha) = \mathbf{R}_s(\bar{\mathbf{g}} + \mathbf{S}\sigma + \mathbf{A}\alpha) + \mathbf{t},$$

где $\mathbf{R} = R(r_x, r_y, r_z)$ – матрица вращения, s – коэффициент масштабирования, $\mathbf{t} = t(t_x, t_y, t_z)$ – вектор параллельного переноса.

Следовательно, модель параметризуется вектором

$$\mathbf{p} = [r_x, r_y, r_z, s, t_x, t_y, t_z, \sigma, \alpha].$$

Таким образом, параметрическая модель лица CANDIDE-3, во-первых, позволяет удобным образом использовать существующие алгоритмы оценки ракурса, а во-вторых, является максимально приближенной к требованиям стандартов ИСО/МЭК 14496-2-2001 и ГОСТ Р ИСО/МЭК 19794-5-2006.

Алгоритм POSIT для оценки положения объекта по изображению. Оценку ракурса лица по изображению можно получить с использованием алгоритма POSIT (Pose from Orthography and Scaling with Iterations), отличающегося высокой производительностью и быстрой сходимостью.

Алгоритм POSIT вычисляет матрицу поворота и вектор параллельного переноса так, чтобы ошибка между проекцией характерных точек трехмерной модели на плоскость изображения и соответствующими им размеченными точками на изображении была минимальна. Метод POSIT комбинирует два алгоритма. Первый, POS (Pose from Orthography and Scaling), аппроксимирует перспективную проекцию с помощью масштабированной ортографической проекции и вычисляет матрицу поворота и вектор параллельного переноса объекта посредством решения СЛАУ. Второй, POSIT (POS with Iterations), в цикле использует результаты работы POS для нахождения более точной масштабированной ортографической проекции характерных точек, и затем применяет POS к полученным точкам, а не к исходным. В итоге алгоритм POSIT сходится к точному решению за несколько итераций [2].

На вход алгоритма должно поступать не менее 4 точек объекта, не лежащих в одной плоскости. Увеличение числа используемых точек должно вести к повышению устойчивости к ошибкам разметки характерных точек объекта на изображении и точности определения матрицы поворота и вектора параллельного переноса.

Метод определения ракурса лица по изображению. Как упоминалось выше, параметрическая модель CANDIDE-3 содержит в себе координаты вершин (КТ), соответствующих средней форме лица с нейтральным выражением. Те вершины модели, которые соответствуют антропометрическим КТ лица, могут быть достаточно точно определены оператором и отмечены на изображении. Таким образом, выполняются все условия, необходимые для определения ракурса с помощью алгоритма POSIT.

Исследование точности определения ракурса. В настоящей работе исследуется точность работы алгоритма POSIT применительно к оценке ракурса лица по его изображению в зависимости от числа используемых КТ.

Для проведения исследований было разработано специальное программное обеспечение (СПО), предоставляющее оператору возможность разметки произвольного числа КТ лица в соответствии с параметрической моделью CANDIDE-3, и производящее определение ракурса лица с помощью алгоритма POSIT.

Как упоминалось ранее, алгоритм POSIT на выходе дает матрицу поворота и вектор параллельного переноса. Для интерпретации получаемых результатов, производится факторизация матрицы поворота в произведение трех матриц поворота вокруг осей x , y , z и вычисление соответствующих им трех углов Эйлера [10].

Оценка точности результатов работы алгоритма производилась с использованием верифицированной базы данных (БД) Pointing'04, содержащей 375 изображений 15 людей, снятых с различным ракурсом (рис. 3) [11]:

- вращение вокруг оси x (наклоны вперед/назад): 0° , -15° , $+15^\circ$, -30° , $+30^\circ$;
- вращение вокруг оси y (поворот влево, вправо): 0° , -15° , $+15^\circ$, -30° , $+30^\circ$.

Размер ПЗС-матрицы устройства регистрации изображений составляет 1/3 дюйма, а фокусное расстояние равно 50 мм. Таким образом, для данных изображений фокусное расстояние в пикселях составляет 2832 пикс.



Рис. 3. Примеры изображений БД Pointing'04

На изображениях из БД вручную были размечены КТ, являющиеся подмножеством вершин модели CANDIDE-3 (рис. 4).

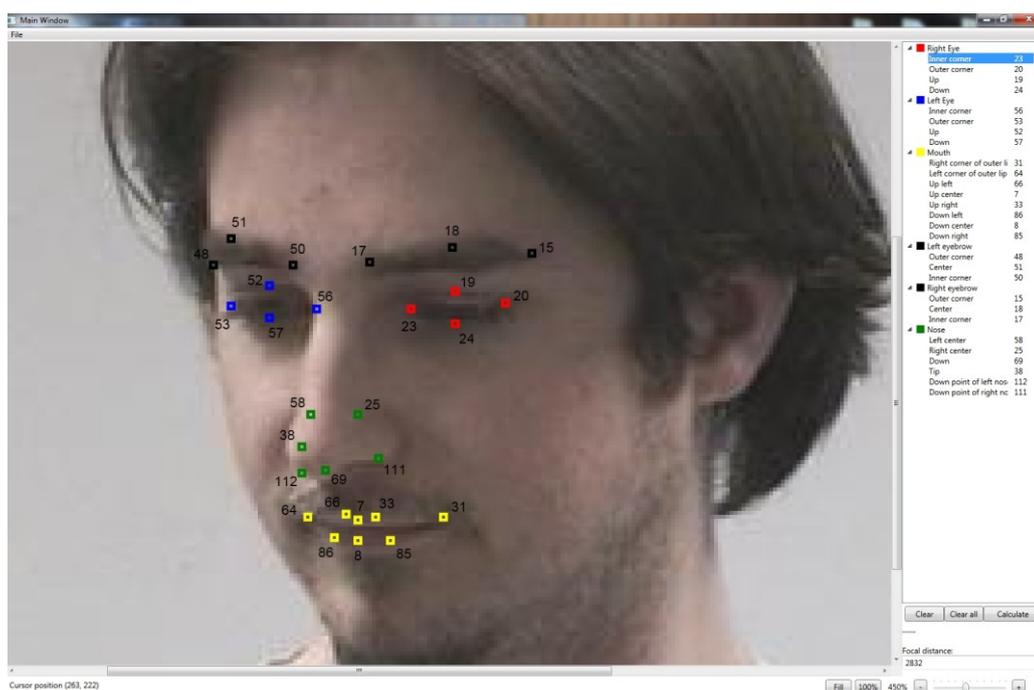


Рис. 4. Набор используемых КТ и пример их разметки на изображении лица в СПО

Нумерация КТ соответствует нумерации вершин модели CANDIDE-3 и приведена в таблице 1. Данный набор исследуемых КТ был выбран исходя из следующих соображений. Во-первых, перечисленные антропометрические КТ лица удобно размечать оператору. Во-вторых, при решении задач верификации и идентификации личности, анализа мимических выражений лица, оценки эмоционального состояния по изображению

лица, как правило, расстановка КТ происходит в автоматическом режиме. При этом приведенный набор КТ схож с наборами, используемыми в алгоритмах автоматической расстановки КТ.

Таблица 1

Идентификаторы и наименования используемых КТ

Идентификатор	Наименование КТ
23	Внутренний угол правого глаза
20	Внешний угол правого глаза
19	Середина верхней границы правого глаза
24	Середина нижней границы правого глаза
56	Внутренний угол левого глаза
53	Внешний угол левого глаза
52	Середина верхней границы левого глаза
57	Середина нижней границы левого глаза
31	Правый угол внешнего контура губ
64	Левый угол внешнего контура губ
66	Угол верхнего контура губы, левее середины
7	Середина верхнего контура губы
33	Угол верхнего контура губы, правее середины
86	Середина нижнего контура губы между её центром и левым углом
8	Середина нижнего контура губы
85	Середина нижнего контура губы между её центром и правым углом
48	Внешний угол левой брови
51	Середина левой брови
50	Внутренний угол левой брови
15	Внешний угол правой брови
18	Середина правой брови
17	Внутренний угол правой брови
58	Верхняя точка левого крыла носа
25	Верхняя точка правого крыла носа
69	Нижняя точка носа
38	Вершина носа
112	Нижний левый край левого крыла носа
111	Нижний правый край правого крыла носа

В таблице 2 приведены результаты исследования качества работы алгоритма POSIT в зависимости от количества КТ, а также от используемого набора КТ. В качестве критериев качества использовались следующие показатели:

1. время работы алгоритма, t (на процессоре Intel Core i5, 4.1 ГГц);
2. абсолютная ошибка определения угла вращения вокруг оси X, Δx ;
3. абсолютная ошибка определения угла вращения вокруг оси Y, Δy .

Таблица 2

Результаты тестирования алгоритма POSIT

№	Набор КТ	Число КТ	$\Delta x, ^\circ$	$\Delta y, ^\circ$	$t, \text{мс}$
1	2	3	4	5	6
1	53, 20, 64, 31	4	18,03	17,94	0,25
2	53, 20, 8, 38	4	12,18	8,18	0,25
3	53, 20, 7, 19	4	8,77	9,31	0,25
4	53, 20, 64, 38, 31	5	11,26	7,08	0,31
5	53, 20, 8, 38, 19	5	9,69	5,80	0,31
6	53, 20, 7, 38, 19	5	9,22	6,028	0,31
7	53, 20, 64, 56, 31, 23	6	15,78	12,15	0,35
8	53, 20, 25, 23, 58, 56	6	13,51	9,79	0,35
9	53, 20, 64, 23, 31, 56, 38	7	11,36	5,87	0,45
10	53, 20, 64, 23, 31, 56, 38, 51	8	11,32	5,87	0,48
11	19, 20, 24, 23, 52, 53, 57, 56 (точки глаз)	8	10,84	10,64	0,48
12	53, 20, 64, 23, 31, 56, 38, 51, 18	9	11,37	5,88	0,50
13	53, 20, 64, 23, 31, 56, 38, 50, 17	9	11,41	5,90	0,50
14	53, 20, 64, 23, 31, 56, 38, 57, 19	9	11,09	5,74	0,50
15	53, 20, 64, 23, 31, 56, 38, 57, 19, 7	10	10,85	5,65	0,52
16	53, 20, 64, 25, 31, 52, 38, 58, 19, 7	10	9,82	5,50	0,52
17	53, 20, 64, 25, 31, 52, 38, 58, 19, 7, 24	11	9,59	5,67	0,52
18	53, 20, 64, 25, 31, 52, 38, 58, 19, 7, 24, 57	12	9,60	5,63	0,52
19	53, 20, 64, 25, 31, 52, 38, 58, 19, 7, 24, 57, 8	13	9,69	5,55	0,52
20	53, 20, 64, 25, 31, 52, 38, 58, 19, 7, 24, 57, 8, 23	14	9,67	5,55	0,52
21	53, 20, 64, 25, 31, 52, 38, 58, 19, 7, 24, 57, 8, 23, 48	15	9,63	5,37	0,52

Продолжение табл. 2

1	2	3	4	5	6
22	19, 20, 24, 23, 52, 53, 57, 56, 7, 8, 31, 64, 33, 66, 85, 86 (точки глаз и рта)	16	9,85	6,71	0,52
23	53, 20, 64, 25, 31, 52, 38, 58, 19, 7, 24, 57, 8, 23, 48, 15, 69	17	8,83	5,64	0,53
24	53, 20, 64, 25, 31, 52, 38, 58, 19, 7, 24, 57, 8, 23, 48, 15, 69, 25, 58, 66	20	8,83	5,82	0,53
25	53, 20, 64, 25, 31, 52, 38, 58, 19, 7, 24, 57, 8, 23, 48, 15, 69, 25, 58, 66, 33, 85, 86, 8	24	9,23	5,83	0,54
26	Все 28 точек	28	11,55	5,78	0,54

Заключение. Анализируя результаты тестирования алгоритма POSIT, можно сделать следующие выводы:

- При увеличении числа КТ до определенного количества ошибка уменьшается и составляет $\Delta x = 8,8^\circ$, $\Delta y = 5,6^\circ$ при оптимальном числе КТ (17 точек).
- Визуальный анализ изображений БД Pointing'04 показал, что приведенная в ней оценка углов вращения головы носит примерный характер и существенно меняется от человека к человеку. Ошибка определения углов вращения вокруг осей алгоритмом POSIT примерно соответствует точности оценки поворотов головы в БД (так, например, ошибка Δy меньше Δx).
- Качество работы алгоритма зависит не только от количества, но и от набора КТ. Например, при использовании 4-х точек, минимальных ошибок можно добиться на основе трех точек глаз (точки 53, 20, 19) и середины верхней границы верхней губы (точка 7).
- Исходя из минимальных значений ошибок, рекомендуется использовать набор КТ № 23 из таблицы 2 в качестве входного параметра алгоритма POSIT при решении задачи оценки ракурса лица по его 2D изображению.

Список литературы

1. Murphy-Chutorian E., Trivedi M.M. Head pose estimation in computer vision: a survey //IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2009. – Vol.31, No.4. – P.607-626.
2. DeMenthon D.F., Davis L.S. Model-based object pose in 25 lines of code // International Journal of Computer Vision, Volume 15, Issue 1-2, June 1995. pp. 123-141

3. Gao X.S., Hou X.-R., Tang J., Chang H.-F. Complete Solution Classification for the Perspective-Three-Point Problem // IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 25, Issue 8, August, 2003. pp. 930 - 943.
4. Moreno-Noguer F., Lepetit V., Fua P. EPnP: Efficient Perspective-n-Point Camera Pose Estimation // Int. J. Comput. Vision, 81(2), 2009. pp. 155–166.
5. Ahlberg J. CANDIDE-3 – an updated parameterized face, Report No. LiTH-ISY-R-2326, Dept. of Electrical Engineering, Linköping University, Sweden, 2001. 16 p.
6. Cootes T.F., Edwards G.J. Active Appearance Models // Proc. European Conference on Computer Vision, 2, 1998. pp. 484-498
7. Edwards G.J., Cootes T.F., Taylor C.J. Face Recognition Using Active Appearance Models // Proc. European Conference on Computer Vision, 2, 1998. pp. 581-695
8. ISO/IEC 14496-2. Information technology – Coding of audio-visual objects – Part 2: Visual. 536 p.
9. ГОСТ Р ИСО/МЭК 19794-5-2006. Автоматическая идентификация. Идентификация биометрическая. Форматы обмена биометрическими данными. Часть 5. Данные изображения лица. М., Стандаринформ, 2006. 68 с.
10. Eberly D. Euler Angle Formulas // Geometric Tools, LLC. URL. <http://www.geometrictools.com/Documentation/EulerAngles.pdf> (дата обращения: 7.02.2013)
11. Gourier N., Hall D., Crowley J. L. Estimating Face Orientation from Robust Detection of Salient Facial Features // Proceedings of Pointing 2004, ICPR, International Workshop on Visual Observation of Deictic Gestures, Cambridge, UK, 22 August, 2004. 9 p.