

# Метод реконструкции и анимации модели головы с использованием RGBD-изображения

Ю.Х. Козлова

Самарский национальный исследовательский университет им. академика С.П. Королева

Самара, Россия

jganeeva99@gmail.com

**Аннотация**—В статье представлен метод реконструкции и анимации цифровой модели головы человека по одному RGBD-изображению. Предложен подход к оптимизации параметрической модели FLAME с использованием облака точек лица, соответствующего одному RGBD-изображению. Предложенный подход оптимизации позволяет получить модель головы с более выраженными чертами исходного лица чем подходы оптимизации, использующие RGB-изображения или те же подходы, обобщенные на RGBD-изображения.

**Ключевые слова**—3D-реконструкция, 3D анимация, виртуальная реальность, дополненная реальность, FLAME, RGBD, облако точек, оптимизация

## 1. ВВЕДЕНИЕ

3D-реконструкция цифровой модели головы человека является довольно сложной и актуальной задачей в области компьютерного зрения. Актуальность данной задачи в последние годы сильно возросла, что связано с внедрением в повседневную жизнь человека технологий дополненной и виртуальной реальности. На практике только лишь реконструированная цифровая модель головы представляет меньший интерес, чем, например, ее последующая анимация. Такое применение технологии позволяет решать проблему виртуального присутствия человека.

Методы реконструкции головы можно классифицировать на параметрические [1], [2] непараметрические [3], [4] и смешанные [5]. Для реконструкции головы в рамках экспериментальных исследований за основу выбрана параметрическая модель FLAME [2]. К достоинствам параметрических методов реконструкции головы можно отнести высокую точность экстраполяции поверхности даже в тех местах, данные для которых отсутствуют. К недостаткам же можно отнести отсутствие возможности реконструкции волос, маленьких деталей на лице и относящихся к поверхности головы и лица аксессуаров. Анимация модели головы в случае использования параметрических моделей реализуется путем введения этапа оптимизации параметров, отвечающих за положение головы и выражения лица отслеживаемого в кадре субъекта.

## 2. ОПИСАНИЕ МЕТОДА РЕКОНСТРУКЦИИ И АНИМАЦИИ МОДЕЛИ ГОЛОВЫ

На «Рис. 1» представлена схема метода реконструкции и анимации головы по RGBD-

изображению. Метод состоит из трех основных этапов. На первом этапе реализуется обработка и подготовка входных данных для реконструкции. На втором этапе реализуется процедура реконструкции и текстурирования модели головы. Основным интересом на данном этапе представляет оптимизация модели FLAME.

Процедура оптимизации параметров, описывающих модель головы, является итеративной и включает в себя три оптимизационных блока. Первый блок отвечает за оптимизацию параметров положения камеры и положения головы относительно входного RGBD-кадра на основе 3D-координат ключевых точек лица. Второй блок отвечает за оптимизацию параметров формы головы (лица) и выражения лица на основе 3D-координат ключевых точек лица и отфильтрованного облака точек. Наконец, третий блок необходим для уточнения формы головы (лица) с помощью оптимизации соответствующих параметров на основе отфильтрованного облака точек. Количественная мера, показывающая качество реконструкции и участвующая в обновлении параметров модели с помощью механизма обратного распространения ошибки определяется как усредненное евклидово расстояние, которое вычисляется по формуле (1). При одновременном использовании данных различного характера (3D-координаты ключевых точек и облако точек) итоговая мера вычисляется как сумма значений соответствующих мер.

$$loss_{L_2} = \sqrt{\frac{1}{K} \sum_{k=0}^{K-1} (kpt_{gt_k} - kpt_{predict_k})^2}, \quad (1)$$

где  $K$  – количество ключевых точек или количество точек в целевом облаке точек,

$kpt_{gt_k}$  – целевые координаты  $k$ -ой ключевой точки или целевые координаты  $k$ -ой точки из облака точек,

$kpt_{predict_k}$  – предсказанные координаты  $k$ -ой ключевой точки или координаты  $k$ -ой ближайшей точки полигона предсказанного меша.

Заключительным этапом процедуры реконструкции является получение текстурной развертки. На текущем этапе производится оптимизация параметров освещения

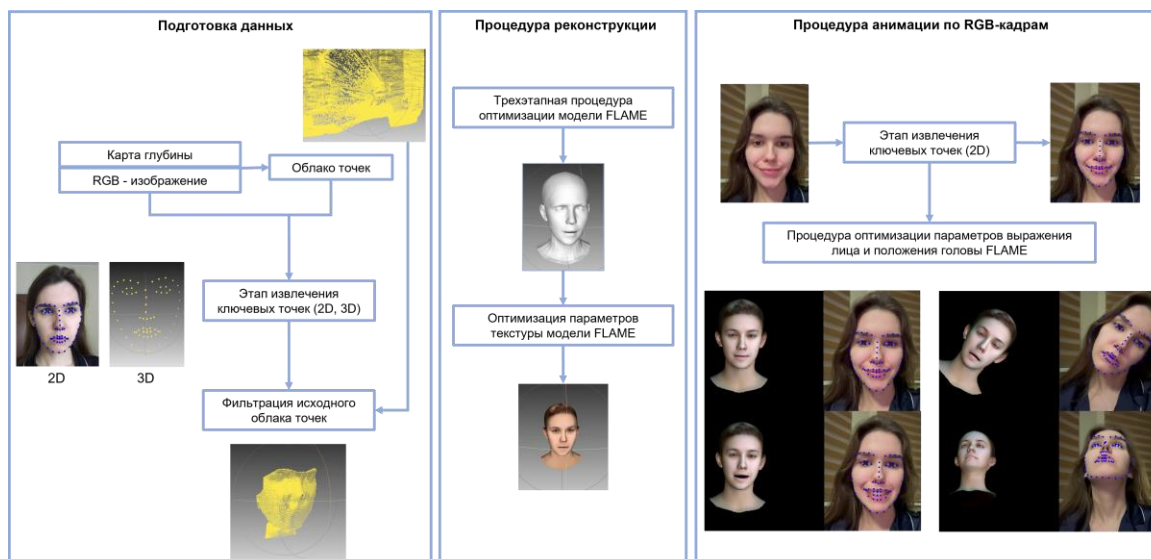


Рис. 1. Схема метода реконструкции и анимации модели головы с использованием RGBD-изображения

и текстуры. В роли количественной меры для оптимизации, которая вычисляется по формуле (2), выступает среднее значение абсолютной попиксельной разницы между исходным RGB-изображением и отрендеренной моделью с наложенной текстурой, оптимизированной позой и положением камеры, таким образом, чтобы отрендеренная голова занимала положение аналогично положению головы в кадре:

$$loss_{img} = \frac{1}{N \cdot M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} |img_{target_{i,j}} - img_{predict_{i,j}}|, \quad (2)$$

где  $N$  – длина изображения,  $M$  – ширина изображения,  $img_{target}$  – целевое изображение,  $img_{predict}$  – отрендеренная модель с наложенной текстурой, оптимизированной позой и положением камеры.

Этап процедуры анимации принимает на вход реконструированную модель головы, полученную на предыдущем шаге метода, и RGB-видеопоследовательность. Для каждого кадра видеопоследовательности извлекаются двумерные ключевые точки лица с использованием метода, представленного в [6], которые затем используются для оптимизации параметров положения головы в кадре и выражения лица. В роли количественной меры для оптимизации выступает усредненное евклидово расстояние по 2D-координатам, которое вычисляется по формуле (1).

### 3. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ

Для проведения экспериментальных исследований с использованием стереокамеры ZED2 был собран набор данных для нескольких людей. Каждому человеку соответствует RGB-изображение, карта глубины и облако точек.

На рисунке 1 представлены результаты работы представленного метода реконструкции и анимации для каждого этапа. Итоговые значения количественной меры для оптимизации для представленных данных на каждом из этапов оптимизации при выполнении процедуры реконструкции следующие: 0,0476; 0,0177; 0,0013;

0,0353. Важно отметить, что прирост значения количественной меры для оптимизации на заключительном этапе оптимизации обусловлен другим характером данных, поступающих в формулу (1). При оптимизации на этапе анимации максимальное значение количественной меры для оптимизации для кадра составило 0,04.

### ЗАКЛЮЧЕНИЕ

По результатам проведенных экспериментальных исследований было выявлено, что предложенный подход к оптимизации на этапе реконструкции модели головы позволяет получить модель с достаточно выраженными мелкими чертами лица. В дальнейшем планируется перейти от параметрического метода реконструкции модели головы к смешанному и на основе полученной в текущей работе модели головы реализовать метод уточнения таких деталей, как волосы, маленькие детали на лице и относящиеся к поверхности головы и лица аксессуары.

### ЛИТЕРАТУРА

- [1] Feng, Y. Learning an Animatable Detailed 3D Face Model from In-The-Wild Images / Y. Feng, H. Feng, M.J. Black, T. Bolkart // ACM Transactions on Graphics. — 2021. — Vol. 40(4). — P. 1-13.
- [2] Li, T. Learning a model of facial shape and expression from 4D scans / T. Li, T. Bolkart, M.J. Black, H. Li, J. Romero // ACM Transactions on Graphics. — 2017. — Vol. 36(6). — P. 1–17. — DOI: 10.1145/3130800.3130813.
- [3] Dou, P. End-to-End 3D Face Reconstruction with Deep Neural Networks / P. Dou, S.K. Shah, I.A. Kakadiaris // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2017. — P. 1503–1512. — DOI: 10.1109/CVPR.2017.164.
- [4] Jackson, A. Large Pose 3D Face Reconstruction from a Single Image via Direct Volumetric CNN Regression / A. Jackson, A. Bulat, V. Argyriou, G. Tzimiropoulos // 2017 IEEE International Conference on Computer Vision (ICCV). — 2017. — P. 1031-1039
- [5] Grassal, P.-W. Neural Head Avatars from Monocular RGB Videos / P.-W. Grassal, M. Prinzel, T. Leistner, C. Rother, M. Nießner, J. Thies // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). — 2022. — P. 18632-18643
- [6] Kazemi, V. One millisecond face alignment with an ensemble of regression trees / V. Kazemi, J. Sullivan // 2014 IEEE Conference on Computer Vision and Pattern Recognition. — 2014. — P. 1867–1874. — DOI: 10.1109/CVPR.2014.241.