

УДК 004.932

DOI: 10.26102/2310-6018/2019.26.3.005

К.В. Климов
**РЕГИСТРАЦИЯ ТРЕХМЕРНОЙ ПОЛИГОНАЛЬНОЙ МОДЕЛИ
ЧЕЛОВЕЧЕСКОГО ЛИЦА С ИСПОЛЬЗОВАНИЕМ РЕЗУЛЬТАТОВ
ТРЕХМЕРНОГО СКАНИРОВАНИЯ И ДЕТЕКТОРА ЛИЦЕВЫХ
ГРАНИЦ**

*Воронежский государственный университет,
Воронеж, Россия*

Актуальность данного исследования обусловлена необходимостью создания высокоточного алгоритма регистрации трехмерных результатов сканирования человеческого лица. Несмотря на то, что за последние десятилетия появилось большое количество различных алгоритмов регистрации трехмерных моделей, включая модели человеческого лица, подавляющее большинство из них не способны проводить точную регистрацию лица в области губ и глаз. Данная работа призвана частично исправить данную ситуацию, путем использования детектора лицевых границ на основе глубокого обучения, а именно с использованием сверточных нейронных сетей. Само по себе, использование детекторов ключевых точек лица (как и на основе нейронных сетей, так и на основе других методов) не является новинкой в области неригидной регистрации. Однако такой подход не дает достаточной точности – сама по себе задача поиска ключевых точек является некорректно поставленной задачей, т.к. не все ключевые точки можно разметить согласованно между разными лицами или между выражениями одного и того же лица. Например, очень сложно поставить точки в середине глаза так, чтоб на разных кадрах они были анатомически на одном месте. В ходе исследования был предложен альтернативный способ – использование лицевых границ, разметка которых является однозначной задачей. Полученные результаты показывают преимущество предложенного подхода.

Ключевые слова: неригидная регистрация, трехмерное сканирование, метод ближайших точек, обнаружение лицевых границ, глубокое обучение, сверточная нейронная сеть.

Введение

Одной из динамично развивающихся областей, в игровой и киноиндустриях, является лицевая анимация на основе трехмерного сканирования лиц реальных актеров. Для получения сканов лиц актеров используют один из методов трехмерного сканирования, чаще всего фотограмметрию, в которой трехмерная полигональная модель актера восстанавливается на основе одновременно снятых фотографий его головы с разных ракурсов.

Под трехмерной полигональной моделью подразумевается упорядоченное множество вершин в трехмерном пространстве, и множеством полигонов, которые строятся на данных вершинах. Топологией называют связность вершин между собой.

К сожалению, методы сканирования общего назначения не имеют никакой информации о типе сканируемого объекта, поэтому, каждый скан имеет свою собственную нерегулярную топологию. При создании анимации, требуется, чтобы все модели лица находились в одной эталонной топологии (базовой топологии). Перевод трехмерной модели из одной топологии в другую может быть осуществлен при помощи алгоритмов регистрации трехмерных моделей.

Материалы и методы

1. Регистрация трехмерных полигональных моделей

Задача регистрации трехмерных моделей заключается в поиске соответствий между вершинами исходной (базовой) модели, и вершинами целевой. Одним из самых популярных способов решить данную задачу является использование метода ближайших точек (ICP - Iterativeclosestpoints) [1].

Рассмотрим классический алгоритм ICP. В общем случае алгоритм состоит из цикла, в котором последовательно повторяются следующие шаги:

1. Для каждой точки деформированной исходной модели ищется ближайшая точка на целевой модели.
2. С учетом найденных соответствий ищутся параметры модели деформации, которые приближают вершины базовой модели к целевым.

Шаги 1 и 2 повторяются до сходимости (когда параметры модели не изменились после очередного шага), или же до достижения максимального числа итераций. Шаг 1 тривиален в реализации, шаг 2 заключается в минимизации функции следующего вида:

$$E(p) = E_{icp}(p) + E_{fixed}(p) + E_{reg}(p) \quad (1)$$

$$E_{icp}(p) = \sum_i \|v(p)_i - v'_i\| \quad (2)$$

$$E_{fixed}(p) = \sum_{c \in C} \|v(p)_c - v_c^*\| \quad (3)$$

Где $\|x\| = \sum_i x_i^2$, p – параметры модели деформации, $v(p)_i$ – i -ая вершина базовой модели после деформации, конкретный вид зависит от модели деформации. v'_i – i -ая вершина целевой модели, найденная в ходе шага 1 алгоритма ICP, $E_{reg}(p)$ – ошибка регуляризации, которая показывает, насколько текущие параметры деформации нарушают гладкость модели, конкретный вид зависит от модели деформации. C – множество индексов фиксированных точек, v_c^* – целевые положения фиксированных точек, как правило, задается пользователем, или внешним алгоритмом.

Строго говоря, в продвинутых вариантах ИСР первое и второе слагаемое может использовать более сложные способы оценки, чем просто расстояние между точками в трехмерном пространстве. Например, расстояние от точки до плоскости, или расстояние между точкой в двумерном пространстве и проекцией трехмерной точки.

Модель деформации может отличаться, в зависимости от того, насколько точная требуется регистрация, и сколько информации мы можем предоставить алгоритму.

Самой простой моделью деформации является ригидная модель

$$v(r, t)_i = R(r, v_i) + t \quad (4)$$

где v_i – изначальная вершина базовой модели, r и t параметры модели – параметризованный поворот (конкретный вид параметризации зависит от задачи) и вектор переноса. $R(r, v)$ – функция выполняющая поворот входной вершины.

Такая модель деформации очень устойчива к шумам в входных данных, и может работать без дополнительной информации о фиксированных точках. Однако такая модель обладает низкой выразительностью – все, что она может это повернуть и переместить исходную модель.

Для задачи регистрации человеческого лица требуется модель с большей выразительностью, способная захватить мелкие детали целевого скана. Например, модель на основе блендшейпов, или As-rigid-as-possible модель [2].

К сожалению, чем более выразительной становится модель, тем больше шанс, что при регистрации ИСР остановится в плохом минимуме. Данную проблему можно избежать, если указать алгоритму заранее известные соответствия, например, между глазами базовой и целевой моделей. Это можно сделать вручную, но для массовой обработки трехмерных сканов этот способ слишком дорог в применении. Поэтому особый интерес представляют методы, позволяющие находить и использовать соответствия автоматически.

2. Обнаружение ключевых точек лица

Классический подход [3] заключается в использовании детектора ключевых точек лица (faciallandmarksdetector), которому на вход подаются рендеры базовой модели и целевого скана. На выходе детектор дает набор координат точек, фиксированного размера. В простейшем случае обнаруженные точки на изображении могут быть спроецированы обратно на модели, для получения точек в трехмерном пространстве. В сложном варианте, может быть использована функция ошибки на основе проекции:

$$E'_{fixed}(p) = \sum_{i \in C} \|P(v_i(p)) - u_i\| \quad (5)$$

где $P(v)$ – функция, проецирующая трехмерную вершину в пространство рендера, u_i – двумерные координаты точки обнаруженной детектором.



Рисунок1. Пример разметки ключевых точек в базе HELEN [4]

Существует большое количество разнообразных алгоритмов для обнаружения ключевых точек лица. В частности, не так давно были распространены алгоритмы на основе решающих деревьев [5]. Однако, в последнее время алгоритмы на основе глубокого обучения (сверточных нейронных сетей) показали свое превосходство в задаче поиска ключевых точек лица [6], [7].

К сожалению, в задаче разметки ключевых точек лица нет однозначного решения [8]. Можно однозначно отметить точки соответствующие, например, уголкам рта, или уголкам глаз. Но очень сложно согласованно отметить точки на краю губ у разных людей так, чтобы анатомически они соответствовали друг другу. На практике это приводит к тому, что разметка ключевых точек содержит в себе шум, что затрудняет обучение качественного детектора. Кроме этого, шум в разметке мешает численной оценке качества детектора, т.к. на каком-то моменте, расхождение с *groundtruth* перестает означать ошибку детектора. Так же, разметка ключевыми точками плохо учитывает поворот головы в кадре.

Для того чтобы преодолеть эти проблемы в ходе данного исследования было предложено использовать вместо детектора ключевых точек, детектор лицевых границ. Под лицевой границей подразумевается любая часть человеческого лица, которая может быть размечена при помощи кривой (ломанной, сплайном), например внешний контур губ, внешний контур глаз, контур бровей, и т.д.



Рисунок 2. Пример разметки глаза при помощи сплайнов.

Такой подход позволяет учитывать неоднозначность разметки лица, а так же повышает точность разметки, т.к. теперь у нас вместо дискретных точек непрерывные кривые.

К сожалению, на сегодняшний день, среди открытых баз данных [4], [9] не существует базы, с достаточно точной разметкой. Кроме этого, существующие базы содержат изображения слишком низкого качества, чтобы их можно было использовать в задаче регистрации трехмерных сканов. Поэтому, в ходе данной работы была собрана новая база человеческих лиц, с разметкой лицевых границ. В базе содержится около 10 000 фотографий приблизительно 500 человек.

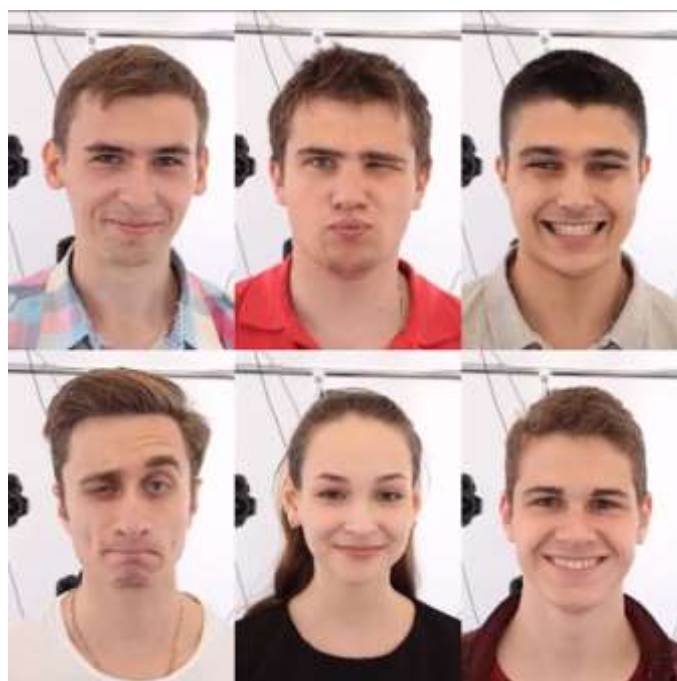


Рисунок 3. Примеры собранных фотографий лиц.

Имея собранные данные можно приступить к обучению детектора. Существует много способов организовать регрессию кривой, например использование сверток на графах [10], однако в данной работе используется

самый простой – размеченные кривые разбиваются на конечное число равноудаленных друг от друга точек, после чего данные точки предсказываются при помощи простой сверточной сети (CNN) с линейным выходным слоем. Для реализации сети был использован фреймворк PyTorch [11].



Рисунок 4. Пример работы детектора.

3. Регистрация трехмерной модели с использованием лицевых кривых.

Имея задетектированные лицевые границы можно приступить к регистрации модели. К сожалению, мы не можем использовать простую функцию (5) для лицевых границ, т.к. у нас больше нет однозначного соответствия между точками базовой модели, и обнаруженными границами. Для решения этой проблемы предлагается следующий подход.

Пусть $k \in R^{N \times 2}$ - множество точек задающих ломаную в двумерном пространстве. Пусть $s(k) : R^{N \times 2} \rightarrow R^{N \times 2}$ функция которая переразбивает ломаную таким образом, чтобы полученная ломаная была по форме как можно ближе к входной ломаной, и при этом, расстояния между ее точками были одинаковы. Реализация такой функции тривиальна, поэтому опустим ее для простоты.

Теперь мы можем задать расстояние между ломаными как

$$L(k_1, k_2) = \sum_i ||s(k_1)_i - s(k_2)_i|| \quad (6)$$

Теперь, мы можем задать положение требуемых кривых в трехмерном пространстве на базовой модели $\mu_i(p) : R^{N_{parameters}} \rightarrow R^{N \times 3}$, взять кривые обнаруженные на рендере при помощи детектора k_i , и составить следующую функцию ошибки по аналогии с функцией (5).

$$E''_{fixed}(p) = \sum_i L(P'(\mu_i), k_i) \quad (7)$$

где $P'(\mu)$ – функция, проецирующая точки ломаной из трехмерного пространства в пространство рендера.

Результаты

К сожалению, в общем случае, не существует однозначного способа численно оценить качество регистрации моделей. Поэтому продемонстрируем качество регистрации визуально.



Рисунок 5. Базовая модель, и скан с результатами обнаружения лицевых границ.

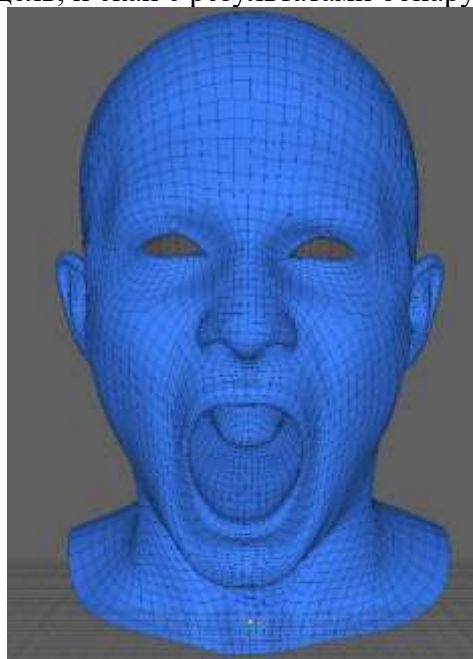


Рисунок 6. Результат регистрации.

На Рисунке 5 можно увидеть входные модели для регистрации, на Рисунке 6 можно увидеть результат работы алгоритма регистрации с предложенным способом обнаружения лицевых границ.

Заключение

В рамках данной работы был предложен алгоритм регистрации трехмерных полигональных моделей, который способен эффективно использовать информацию о лицевых границах. К недостаткам метода можно отнести необходимость в специфической разметке, и усложненную функцию ошибки.

ЛИТЕРАТУРА

1. Rusinkiewicz S. Efficient Variants of the ICP Algorithm / Rusinkiewicz S, Levoy M. // 3D Digital Imaging and Modeling — Stanford University: 2001. — URL: <https://graphics.stanford.edu/papers/fasticp/>
2. Sorkine O. As-rigid-as-possible surface modeling / Sorkine O., Alexa M. // SGP '07 Proceedings of the fifth Eurographics symposium on Geometry processing – 2007 – page 109-116.
3. Fyffe G. Multi-View Stereo on Consistent Face Topology / Fyffe G., Nagano K. // EUROGRAPHICS 2017 – 2017 – URL: <http://vgl.ict.usc.edu/Research/XimeaRiver/>.
4. Le V. Interactive Facial Feature Localization / Le V., Brandt J., Lin Z. // ECCV2012 – 2012 – URL: http://www.ifp.illinois.edu/~vuongle2/helen/eccv2012_helen_final.pdf.
5. Kazemi V. One Millisecond Face Alignment with an Ensemble of Regression Trees / Kazemi V., Sullivan J. // Conference: Computer Vision and Pattern Recognition – 2014 – page 1867-1874.
6. Николенко С. Глубокое обучение. Погружение в мир нейронных сетей / Николенко С, Кадури А., Архангельская Е. — Санкт-Петербург: Питер, 2018. — 480 с.
7. Wu W. Look at Boundary: A Boundary-Aware Face Alignment Algorithm / Wu W., Qian C., Yang S. // CVPR 2018 – 2018 - URL: <https://wywu.github.io/projects/LAB/LAB.html>
8. Dong X. Supervision-by-Registration: An Unsupervised Approach to Improve the Precision of Facial Landmark Detectors / Dong X., Yu S., Weng X. // CVPR 2018 – 2018 – URL: <https://arxiv.org/abs/1807.00966>
9. Sagonas C. 300 Faces In-The-Wild Challenge: database and results / Sagonas C., Antonakos E., Tzimiropoulos G., Zafeiriou S., Pantic M. // Image and Vision Computing – 2016 – pages 3-18.
10. Wu Z. A Comprehensive Survey on Graph Neural Networks / Wu Z., Pan S., Chen F. // arxiv – 2018 – URL: <https://arxiv.org/pdf/1901.00596.pdf>

11. Paszke A. Automatic differentiation in PyTorch / Paszke A., Gross S., Chintala S. // NIPS 2017 – 2017 – URL: <https://openreview.net/pdf?id=BJJsrnfCZ>

K.V. Klimov

3D FACIAL SCAN REGISTRATION WITH FACIAL EDGES DETECTION

*Voronezh state university,
Voronezh, Russia*

The relevance of this work is caused by need of high-fidelity non-rigid registration algorithm for facial scans. Despite the fact that over the past decade, a large number of various algorithms for 3d model registration were proposed, including models of a human face, the vast majority of them are unable to provide accurate face registration for lips and eyes. This work is intended to improve this situation by using a detector of facial boundaries based on deep learning, namely using convolutional neural networks. Usage of facial landmarks detectors (deep learning based or other algorithms) is not novel in the field of non-rigid registration. However such approach is not good enough for precise registration. Facial landmark annotation is ill posed problem in general case, due to ambiguity of landmarks position. For example, it is very difficult to put landmarks in the middle of the eye so that they are anatomically in the same place in different frames. In this work we propose an alternative method - facial edges detection for registration. Obtained results show the advantage of the proposed approach.

Keywords: non-rigid registration, 3d scanning, iterative closest point, facial edges detection, deep learning, convolutional neural networks.

REFERENCES

1. Rusinkiewicz S. Efficient Variants of the ICP Algorithm / Rusinkiewicz S, Levoy M. // 3D Digital Imaging and Modeling — Stanford University: 2001. — URL: <https://graphics.stanford.edu/papers/fasticp/>
2. Sorkine O. As-rigid-as-possible surface modeling / Sorkine O., Alexa M. // SGP '07 Proceedings of the fifth Eurographics symposium on Geometry processing – 2007 – page 109-116.
3. Fyffe G. Multi-View Stereo on Consistent Face Topology / Fyffe G., Nagano K. // EUROGRAPHICS 2017 – 2017 – URL: <http://vgl.ict.usc.edu/Research/XimeaRiver/>.
4. Le V. Interactive Facial Feature Localization / Le V., Brandt J., Lin Z. // ECCV2012 – 2012 – URL: http://www.ifp.illinois.edu/~vuongle2/helen/eccv2012_helen_final.pdf.
5. Kazemi V. One Millisecond Face Alignment with an Ensemble of Regression Trees / Kazemi V., Sullivan J. // Conference: Computer Vision and Pattern Recognition – 2014 – page 1867-1874.
6. Nikolenko S. Deep Learning. / Nikolenko S., Kadurin A., Arkhangelskaya E. — Saint Petersburg: Piter, 2018. — 480 p.

7. Wu W. Look at Boundary: A Boundary-Aware Face Alignment Algorithm / Wu W., Qian C., Yang S. // CVPR 2018 – 2018 - URL: <https://wywu.github.io/projects/LAB/LAB.html>
8. Dong X. Supervision-by-Registration: An Unsupervised Approach to Improve the Precision of Facial Landmark Detectors / Dong X., Yu S., Weng X. // CVPR 2018 – 2018 – URL: <https://arxiv.org/abs/1807.00966>
9. Sagonas C. 300 Faces In-The-Wild Challenge: database and results / Sagonas C., Antonakos E., Tzimiropoulos G., Zafeiriou S., Pantic M. // Image and Vision Computing – 2016 – pages 3-18.
10. Wu Z. A Comprehensive Survey on Graph Neural Networks / Wu Z., Pan S., Chen F. // arxiv – 2018 – URL: <https://arxiv.org/pdf/1901.00596.pdf>
11. Paszke A. Automatic differentiation in PyTorch / Paszke A., Gross S., Chintala S. // NIPS 2017 – 2017 – URL: <https://openreview.net/pdf?id=BJJsrnfCZ>