

Построение параметрической модели головы человека по набору изображений

Максим Федюков

Факультет вычислительной математики и кибернетики
Московский государственный университет им. М. В. Ломоносова
mfedyukov@graphics.cs.msu.ru

1 Введение

Задача моделирования головы человека является актуальной во многих областях: идентификации, отслеживании, видеомоделировании на базе трехмерных моделей, моделировании виртуального присутствия и других. В зависимости от области применения моделирование может производиться с помощью трехмерного лазерного сканера [1, 2], по стереопаре [3, 4], по одной фотографии [5, 6], либо по набору изображений: фотографий или кадрам видеопоследовательности [7, 8]. При постановке задачи реконструкции всей модели головы (но не только модели лица) с минимальными требованиями к входным данным оптимальным является последний подход. Также методы моделирования можно классифицировать по типу выходных данных. В некоторых работах ими является облако точек [9]. Во многих работах реконструированная форма представляется в виде полигональной модели [3, 5, 6, 8]. Отдельной задачей, ставшей актуальной в последние годы с развитием многопользовательских виртуальных сред является реконструкция модели, представляющей собой компактный набор параметров, описывающий морфологические характеристики головы человека.

2 Постановка задачи

Везде, где не оговорено обратное, будем понимать под *полигональной моделью* u вектор (u_1, u_2, \dots, u_N) , $\forall i = 1, \dots, N, u_i \in \mathbb{R}^3$, состоящий из радиус-векторов вершин. Множество граней фиксировано, и потому не рассматривается с целью упрощения модели. На отдельных же этапах, касающихся получения и работы с текстурой, будем рассматривать *текстурированную полигональную модель* $\tilde{u} := (\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_N)$, $\forall i = 1, \dots, N, \tilde{u}_i \in \mathbb{R}^5$, $\tilde{u}_i = (\tilde{u}_{i,x}, \tilde{u}_{i,y}, \tilde{u}_{i,z}, \tilde{u}_{i,u}, \tilde{u}_{i,v})$, $\tilde{u}_{i,x}, \tilde{u}_{i,y}, \tilde{u}_{i,z}$ — пространственные координаты вершины, $\tilde{u}_{i,u}, \tilde{u}_{i,v}$ — текстурные координаты вершины. Под *параметризованной полигональной моделью* a будем подразумевать $a := b + Sp$, где $b = (b_1, b_2, \dots, b_N)$ — базовая полигональная модель, задающая координаты вершин по умолчанию, $S := [s_{ij}]$ ($i = 1, 2, \dots, N; j = 1, 2, \dots, K; s_{ij} \in \mathbb{R}^3$) — матрица смещений, $p := (p_1, p_2, \dots, p_K)$ ($j = 1, 2, \dots, K; p_j \in \mathbb{R}, -1 \leq p_j \leq 1$) — вектор параметров модели, N — число вершин полигональной модели, K — число параметров. Далее такую модель a для краткости будем называть просто *параметрической моделью*. Базовая полигональная модель b и матрица смещений S считаются фиксированными. Решаемая задача заключается в разработке системы, определяющей значения всех параметров p и текстуру по набору из 1–4 фотографий (анфас, в профиль слева, в профиль справа и сзади).

3 Предлагаемый метод

Схема работы системы представлена на рис. 1. Базовая полигональная модель предварительно размечена. *Универсаль-*

ная разметка $T := (t_1, t_2, \dots, t_M)$ представляет собой набор векторов $t_i := (t_1^i, t_2^i, \dots, t_{L_i}^i)$ ($\forall i, j, t_j^i \in \mathbb{N}, t_j^i \leq N; i = 1, 2, \dots, M; j = 1, 2, \dots, L_i; M$ — число векторов, L_i — число элементов векторов, N — число вершин полигональной модели), элементами которых являются индексы t_j^i вершин базовой полигональной модели. Таким образом, вектор $q_i = (q_1^i, q_2^i, \dots, q_{L_i}^i)$, где $\forall i, j, q_j^i = b_{t_j^i}$ задает ломаную, узлами которой являются вершины базовой полигональной модели. Далее (q_1, q_2, \dots, q_M) будем обозначать как Q . Каждая ломаная описывает одну характерную черту лица человека, такую как контур брови, носа, или всей головы. Разметка (ломаные) заданы для каждой проекции — анфас, в профиль слева, в профиль справа и сзади: $T_f(Q_f)$, $T_l(Q_l)$, $T_r(Q_r)$ и $T_b(Q_b)$, соответственно. Однако, в силу того, что работа с разметкой и ломаными на большинстве этапов происходит одинаково для каждой проекции, для упрощения записи, где это возможно, будем писать T и Q , не уточняя индекс проекции.

3.1 Предобработка

На каждой фотографии распознаются характерные точки лица, такие как уголки глаз или кончик носа. На каждую фотографию проецируется соответствующий набор ломаных: $\Omega = PQ$, где $P: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ — оператор ортогонального проецирования на координатную плоскость. Узлы спроецированных ломаных, соответствующие найденным характерным точкам, перемещаются в соответствующие координаты. По ним вычисляется матрица аффинного преобразования, с помощью которой перемещаются все остальные узлы каждой ломаной. После этого дополнительно скорректировать положение узлов спроецированных ломаных Ω можно вручную. Далее фотографии, как текстурированные прямоугольники, располагаются в \mathbb{R}^3 , корректируются их положение, поворот и масштаб. Полагаем, что плоскость Oxy расположена горизонтально, ось Y направлена на наблюдателя, ось X — вправо, ось Z — вверх. Сначала определяется положение фотографии анфас. Она располагается в плоскости Oxz и поворачивается вдоль оси Y таким образом, чтобы ось симметрии ломаных Ω_f на ней совпала с осью Z . Затем фотография в профиль располагается в плоскости Oyz и ее положение, поворот, и масштаб определяются при помощи минимизации среднеквадратичного отклонения вершин, входящих одновременно как в набор ломаных для фотографии анфас Q_f , так и в набор ломаных для фотографии анфас Q_r .

3.2 Получение формы

На этапе построения параметрической модели положение узлов ломаных Ω считается фиксированным. Цель этапа заключается в подгонке проекций ломаных, задаваемых параметрической моделью, к ломаным, полученным на предыдущем этапе. Формально она может быть сформулирована как нахождение значений параметров p , при которых функционал $\mathcal{F} = \sum_{i=1}^M \sum_{j=1}^{L_i} (\omega_j^i - \omega_j'^i)^2$ принимает минимальное значе-

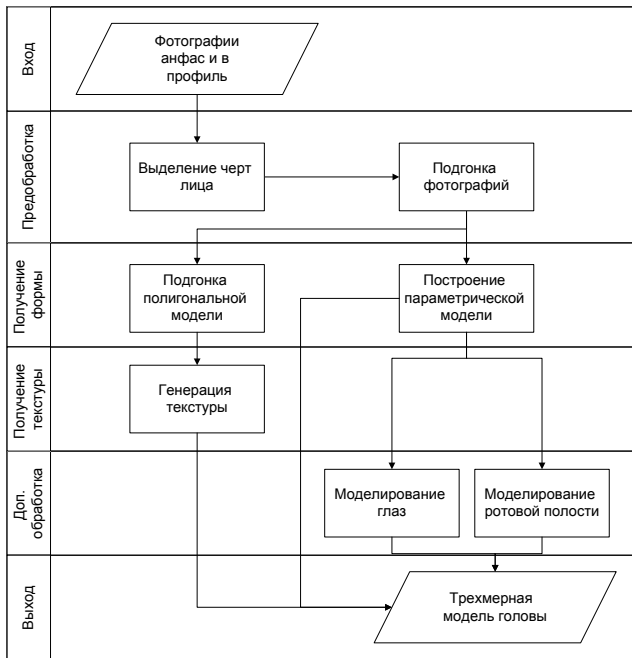


Рис. 1: Схема работы системы моделирования головы человека по набору фотографий.

ние. Здесь ω_j^i — узлы ломаных $\omega' = PQ$, а так как ломаные Q проходят по вершинам параметрической модели $a(p)$, $\forall i, j, \omega_j^i = \omega_j^i(p)$. Минимизация осуществляется методом градиентного спуска. В зависимости от заданной матрицы смещений S значения функционалов невязки G_m могут быть достаточно большими и получаемая полигональная модель a может быть недостаточно точной для генерации по ней текстуры. Для решения этой проблемы создается текстурированная полигональная модель \tilde{c} — копия базовой текстурированной модели \tilde{b} . Затем ломаные на фотографиях Ω проецируются оператором P^{-1} из \mathbb{R}^2 в \mathbb{R}^3 (получаемые ломаные будем обозначать \hat{Q}): $\forall i, j, \hat{q}_{j,x}^i = \omega_{j,x}^i, \hat{q}_{j,y}^i = \omega_{j,y}^i$, а значение z -компоненты восстанавливается из исходной модели: $\hat{q}_{j,z}^i = q_{j,z}^i$. Далее подгонка полигональной модели \tilde{c} осуществляется методом свободных деформаций Дирихле [8].

3.3 Получение текстуры

Для генерации текстуры создается текстурированная полигональная модель \tilde{d} . В случае фотографии анфас, значения элементов \tilde{d} задаются следующим образом: $\tilde{d}_x = 0, \tilde{d}_y = \tilde{c}_u, \tilde{d}_z = \tilde{c}_v, \tilde{d}_u = P\tilde{c}_y, \tilde{d}_v = P\tilde{c}_z$, где P — оператор ортогонального проецирования. Полученная плоская полигональная модель визуализируется в текстуру (в данном случае, фронтальную). Аналогичным образом генерируется текстура по фотографии в профиль. Затем по заранее заданной маске все текстуры смешиваются.

3.4 Дополнительная обработка

Если система, в которой будет использоваться полученная модель головы, предполагает анимацию (в частности, моргание глаз и открывание рта), то исходная базовая модель головы не включает себя модели глаз, зубов, языка и полости рта, и их моделирование является дополнительной задачей. Модель

каждого глаза, представляющая собой объединение двух сфер (глазное яблоко и зрачок) переносится и масштабируется в соответствии с положением ломаной \hat{Q} , описывающей контур данного глаза. Усредненный цвет радужки глаза определяется по фотографии анфас. Затем в пространстве HSL соответственно изменяется цвет радужки на предварительно подготовленной развертке фотографии глаза. Модели полости рта, а также языка и зубов, подгоняются аналогично моделям глаз. Текстура, полученная с фотографий, смешивается с предварительно полученной разверткой фотографий ротовой полости, включающей зубы и язык.

4 Заключение и предстоящая работа

В статье представлен разработанный метод построения параметрической модели головы человека по набору изображений. Данный подход применим к любым параметрическим моделям, обладающим описанными свойствами. В рамках дальнейшей работы планируется как улучшение предложенных алгоритмов, так и добавление новых этапов, таких как фильтрация текстуры и моделирование прически в соответствии с входными фотографиями.

Список литературы

- [1] C. Xu, L. Quan, Y. Wang, T. Tan, and M. Lhuillier. Adaptive multi-resolution fitting and its application to realistic head modeling. In *Proceedings of Geometric Modeling and Processing*, 2004.
- [2] Frank Hülken, Christian Eckes, Roland Kuck, Jörg Unterberg, and Sophie Jörg. Modeling and animating virtual humans for real-time applications. In *The International Journal of Virtual Reality*, 2007, 6(4):11-20, 2007.
- [3] G. Galicia and A. Zakhor. Depth based recovery of human facial features from video sequences. In *Image Processing, 1995. Proceedings., International Conference on*, volume 2, 1995.
- [4] R. Lengagne, P. Fua, and O. Monga. 3D stereo reconstruction of human faces driven by differential constraints. *Image and Vision Computing*, 18(4):337-343, 2000.
- [5] R. Dovgand and R. Basri. Statistical symmetric shape from shading for 3D structure recovery of faces. *Lecture Notes in Computer Science*, pages 99-113, 2004.
- [6] W.A.P. Smith and E.R. Hancock. A Model-Based Method for Face Shape Recovery. In *Pattern Recognition and Image Analysis: Second Iberian Conference, IbPRIA*, pages 268-276. Springer, 2005.
- [7] P. Fua. Regularized bundle-adjustment to model heads from image sequences without calibration data. *International Journal of Computer Vision*, 38(2):153-171, 2000.
- [8] VG Zhislina, DV Ivanov, VF Kuriakin, VS Lempitskii, EM Martinova, KV Rodyushkin, TV Firsova, AA Khropov, and AV Shokurov. Creating and Animating Personalized Head Models from Digital Photographs and Video. *Programming and Computer Software*, 30(5):242-257, 2004.
- [9] Y. Xu, C.S. Xu, Y.L. Tian, S.D. Ma, and M.L. Luo. 3D face image acquisition and reconstruction system. In *IEEE Instrumentation and Measurement Technology Conference, 1998. IMTC/98. Conference Proceedings*, volume 2, 1998.