

ВОССТАНОВЛЕНИЕ ТРЁХМЕРНОЙ СТРУКТУРЫ ЛИЦА С ПОМОЩЬЮ ОПОРНОЙ МОДЕЛИ¹

Струц С.В., Матвеев И.А.
Вычислительный Центр РАН, Москва

Поступила в редакцию 25.06.2011, после переработки 30.06.2011.

Для восстановления рельефа лица человека по стереопаре изображений предлагается использовать опорную 3D-модель, описывающую усреднённые свойства пространственной структуры лица. Представлен способ построения и использования такой модели. Работа метода проверена с помощью базы изображений, имеющейся в открытом доступе. Предложенный метод может быть применён в медицинских системах для измерения формы лица.

Supporting 3D model is proposed as a mean of reconstruction of human face 3D structure. The model contains averaged properties of spatial face geometry. Ways of creating and using the model are described. The proposed methods are verified with image database from public domain. The methods may be used for medical applications to measure face shape.

Ключевые слова: стереорекострукция, биометрия лица, 3D-модель.
Keywords: stereo-reconstruction, facial biometry, 3D-model.

1. Введение

Одно из важных направлений задач компьютерной биометрии - описание и распознавание трёхмерной структуры лица [1]. Во многих приложениях, в частности, в медицинских [2, 3], требуется определение формы лица с высокой точностью. Для этого может применяться трёхмерное сканирование с использованием специального оборудования [2, 3], однако, хотелось бы решить эту задачу при помощи более простых и доступных средств, а именно бытовых фотокамер и персонального компьютера. Наиболее удобным для такого применения является метод определения трёхмерной структуры по стереопаре изображений. Задача стереорекострукции плохо определена и вычислительно сложна [5] и не поддаётся удовлетворительному решению в общем случае. Но в некоторых приложениях (например, системы медицинской диагностики по форме лица) возможно воспользоваться двумя обстоятельствами для упрощения задачи. Во-первых, узкий класс объектов (это только человеческие лица), позволяет построить обобщённую, но при этом достаточно детальную модель, индивидуальные отличия объектов класса

¹Работа выполнена при финансовой поддержке РФФИ (проект № 11-01-00781).

от которой минимальны. Во-вторых, допустимо отказаться от полностью автоматической системы и построить автоматизированную, т.е. требующую нескольких простых операций, выполняемых вручную.

В статье последовательно даётся краткое описание используемого метода стереорекострукции, рассматривается построение опорной модели и её использование для восстановления трёхмерной структуры обрабатываемых лиц, приводятся результаты тестов на базе данных.

1. Восстановление трёхмерной структуры по стереопаре изображений

В описанных далее алгоритмах используются стереопары изображений. Предполагается, что эти изображения получены с помощью стереоскопической системы, подобной [4]. Схема такой стереоскопической системы, реализующей *нормальную стереосъёмку* [11], показана на Рис.1.

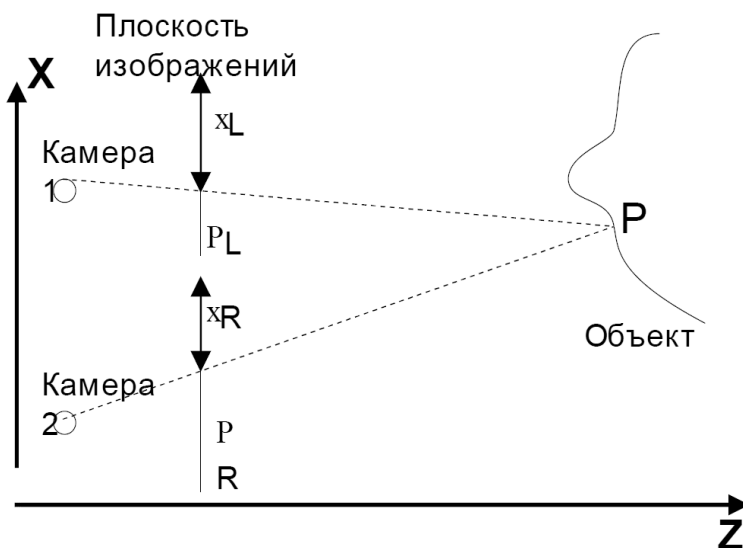


Рис. 1: Схема стереоскопической системы

Чтобы восстановить трёхмерную структуру объекта, необходимо знать трёхмерные координаты (X, Y, Z) для достаточного количества точек объекта. За пространственные координаты X и Y в случае нормальной стереосъёмки можно взять координаты x и y точек одного из изображений стереопары. Для лица человека это может быть фронтальное изображение. Z -координата любой точки $\vec{P} = (X, Y, Z)$ объекта, при нормальной стереосъёмке вычисляется по x -координатам изображений этой точки $\vec{p}_L = (x_L, y_L)$ и $\vec{p}_R = (x_R, y_R)$ (см. Рис.1):

$$Z = \frac{k}{x_L - x_R} \quad (1)$$

где k - коэффициент, зависящий от характеристик стереосистемы, x_L и x_R - x -координаты изображений точки. Здесь и далее индекс L используется для всего,

что связано с левым изображением стереопары, а индекс R - для всего, что связано с правым изображением («левое» и «правое» расположение как на Рис.1). Формула (1) показывает, что Z зависит только от разности x -координат. Эта разность

$$d = x_L - x_R \quad (2)$$

называется рассогласованием [4] или *диспаратностью*. Следует отметить, что в случае отсутствия или малости затенений [11] каждой точке на одном изображении соответствует какая-либо единственная точка на другом, а значит в каждой точке изображения задана диспаратность $d(x, y)$. Соответственно, задана также $Z(x, y)$ – трёхмерная координата точки объекта, изображение которой находится в (x, y) . Для восстановления 3D-структуры по стереопаре изображений необходимо найти соответствие между точками левого и правого изображений. Соответствием здесь называется пара точек \vec{p}_L и \vec{p}_R , являющихся изображениями одной и той же точки \vec{P} трёхмерного объекта. К настоящему моменту предложено большое количество методов поиска соответствий. Можно выделить два больших класса этих методов:

«**Плотные**» методы [5], использующие области изображения, действуют по следующей схеме: для каждой пары точек, где первая точка берётся на одном, а вторая - на другом изображении, (или для подмножества этих пар, ограниченного известной геометрией системы) вычисляется некоторая количественная характеристика соответствия, затем по этой характеристике находятся пары точек, наиболее «подходящих друг другу» и вычисляются диспаратности. Эти методы подходят для любых трёхмерных сцен, но вычисление количественных характеристик для каждой пары точек и нахождения среди них оптимальных пар - очень трудоёмкие операции.

«**Разреженные**» методы [6] используют характерные точки. Такими точками могут быть углы и края объектов, области с выраженной текстурой и т.п. Эти методы работают по следующей схеме: определяются характерные точки на изображениях, выясняется соответствие между характерными точками левого и правого изображений, по этим данным вычисляется геометрия сцены (а при необходимости и геометрия системы камер), и на основании этого находятся новые пары соответственных точек. Методы этого класса работают быстрее «плотных», но применимы только к тем трёхмерным сценам, для которых можно с достаточной точностью и надёжностью получить характерные точки, причём построить достаточно плотную их сеть по всему изображению.

Представленный метод решает проблему поиска соответствий несколько иным способом. Предлагается принять во внимание тот факт, что трёхмерная структура лица человека имеет некоторые характерные свойства, которые можно описать трёхмерной моделью лица. При этом, как метод построения такой модели, так и метод её использования для восстановления трёхмерной структуры совмещают некоторые признаки обоих классов методов стереорекострукции.

2. Построение опорной модели

Для построения 3D-модели была использована стереопара изображений, одно из которых было фронтальным (таким образом функция $Z(x, y)$ для него была однозначной). Далее без ограничения общности будем называть фронтальной

изображение левым, а второе изображение стереопары - правым. В методе предполагается возможность выполнения некоторых операций вручную. Наличие этих операций – безусловно, недостаток метода. Но их предполагается производить всего несколько раз, чтобы получить модель лица, достаточно хорошую для дальнейшего использования. Порядок построения модели следующий.

Шаг 1. Построение сетки характеристических точек. Выбирается область на фронтальном изображении и строится регулярная сетка точек в этой области. Можно вручную выбрать прямоугольную область, в которой будет нарисована сетка. Для лучшей работы алгоритма рекомендуется по возможности выбирать прямоугольник, целиком содержащийся в области лица. Пример прямоугольной области с нарисованной в ней регулярной сеткой показан на рис. 2 (изображение слева). Регулярная сетка размером $N \times N$ точек делит область на $(N - 1)^2$ прямоугольников. Использовалась решётка 10×10 точек.

Шаг 2. Задание соответствия на характеристических точках. Для каждой точки регулярной сетки на фронтальном изображении вручную отмечается соответственная точка на другом изображении. Кроме того, те точки регулярной сетки, которые находятся в районе глаз и носа, также могут быть перемещены на точки лица, для которых легче найти соответственные (например, зрачки, углы глаз, кончик носа). В итоге выбранная область на обоих изображениях делится на выпуклые четырёхугольники. Пример сеток, полученных в результате этого шага, показан на рис. 2 (изображение справа).



Рис. 2: Пример сеток с точками, отмеченными вручную

Шаг 3. Построение плотной карты диспаратностей. Для каждой точки фронтального изображения (в области, покрываемой сеткой точек) с использованием сети ищется соответственная точка, затем вычисляется диспаратность и Z -координата. Для каждой точки \vec{p}_L из покрытой решёткой области на фронтальном изображении соответственная точка находится следующим образом:

А. Определяется четырёхугольник решётки на фронтальном изображении, которому принадлежит точка \vec{p}_L , и положение точки в этом четырёхугольнике.

Б. Определяется точка \vec{p}_R , имеющая такое же положение в соответствующем четырёхугольнике сетки на другом изображении. Для этого используется алгоритм билинейного преобразования четырёхугольников, предложенный [7].

В. Соответственная точка для \vec{p}_L ищется в окрестности точки \vec{p}_R . За соответственную берётся та точка \vec{p}_R' , окрестность которой даёт максимум корреляции [4] функций яркости с окрестностью точки \vec{p}_L . Здесь можно брать достаточно малые окрестности (например 7×7 пикселей). Затем вычисляется диспаратность:

$$d(\vec{p}_L) = d(x_L, y_L) = x_{p_L} - x_{p_R'} \quad (3)$$

и Z -координата:

$$Z(x_L, y_L) = \frac{k}{d(x_L, y_L)} = \frac{k}{x_{p_L} - x_{p_R'}}. \quad (4)$$

Шаг 4. Привязка модели. На изображении отмечаются характерные точки лица (используются центры глаз и кончик носа). Координаты, которые имеют прообразы этих точек в трёхмерной модели, сохраняются для дальнейшего использования.

3D-модель, полученная в результате выполнения алгоритма, показана на рис. 3.

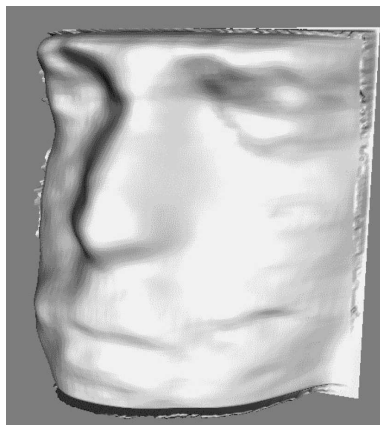


Рис. 3: 3D-модель лица

Описанный алгоритм, служащий для построения опорной модели, может рассматриваться и как метод построения собственно трёхмерной структуры. Однако, он требует слишком много операций, производимых вручную, чтобы быть использованным в этом качестве.

Модель, построенная по предложенному методу, имеет два характерных свойства, которые важны для дальнейшего её использования. Первое отражено в формуле (4): Z -координата каждой точки модели однозначно определяется ее x - и y -координатами. Второе свойство состоит в том, что самая высокая точка лица (кончик носа) имеет максимальную Z -координату.

Для достижения лучшего (в смысле «общности» итоговой модели лица) результата, следует взять несколько моделей, построенных по этому методу (из фотографий людей с разными типами лица) и построить «среднюю» модель, Z -координата

каждой точки которой является средним значением Z -координат точек индивидуальных моделей. Для построений опорной модели в данной работе было использовано 5 индивидуальных моделей.

3. Использование опорной модели

Основная идея заключается в том, что для каждой точки p_L исходного изображения опорная модель дает приблизительную величину диспаратности $d(x, y)$, т.е. область для поиска соответственной точки, причём достаточно малого размера.

Проблема здесь заключается в том, что обрабатываемые новые изображения лица могут иметь иной размер, а само лицо – иные пропорции, нежели опорная модель. Поэтому требуется приведение опорной модели к параметрам обрабатываемого изображения. Для этого устанавливается соответствие по нескольким точкам, а именно по кончику носа и центрам глаз, которые были зафиксированы для опорной модели при её построении. Эти точки надёжно ищутся автоматическими методами [10], впрочем, для автоматизированной системы допустимо отмечать их вручную.

Порядок восстановления 3D структуры лица по опорной модели следующий:

Шаг 1. Отмечаются (или автоматически определяются) кончик носа и центры глаз на обоих изображениях. Отметка нескольких точек позволяет вычислить их диспаратность и, используя (4), масштабирующий коэффициент k .

Шаг 2. Вычисляются диспаратности и Z -координаты. Соответственная точка для каждой точки p_L выделенной области вычисляется следующим образом:

А. Определяется Z -координата для точки с такой же относительной (x, y) позицией на модели.

Б. По известным k и Z вычисляется d и p_R на другом изображении.

В. Соответственная точка p_R' ищется в окрестности точки p_R так же, как и при построении модели, т.е. по максимуму корреляции.

Г. Вычисляется Z -координата по формуле (4).

Результаты численного эксперимента приведены в таблице 1. Восстановленная 3D-структура показана на рис. 4.

Полученная 3D-карта, возможно, не слишком детализированная, но достаточно точная для целей, приведённых в разделе 1, т.е. определение врождённой асимметрии лица.

4. Вычислительный эксперимент и выводы

Был проведён эксперимент для оценки предложенных методов. Для сравнения был выбран алгоритм, предложенный В.Колмогоровым [8]. Программа, реализующая этот алгоритм, имеется в свободном доступе в Интернет (<http://www.cs.ucl.ac.uk/staff/V.Kolmogorov/software/match-v3.3.src.tar.gz>). Этот алгоритм относится к классу «плотных» и использует минимизацию энергии для вычисления оптимальных пар соответственных точек. Для тестирования метода построения опорной 3D-модели был написан пользовательский интерфейс, позволяющий выполнять все требуемые операции, производимые вручную, т.е. выбор области и отметку соответственных точек для сетки. Также он даёт возможность

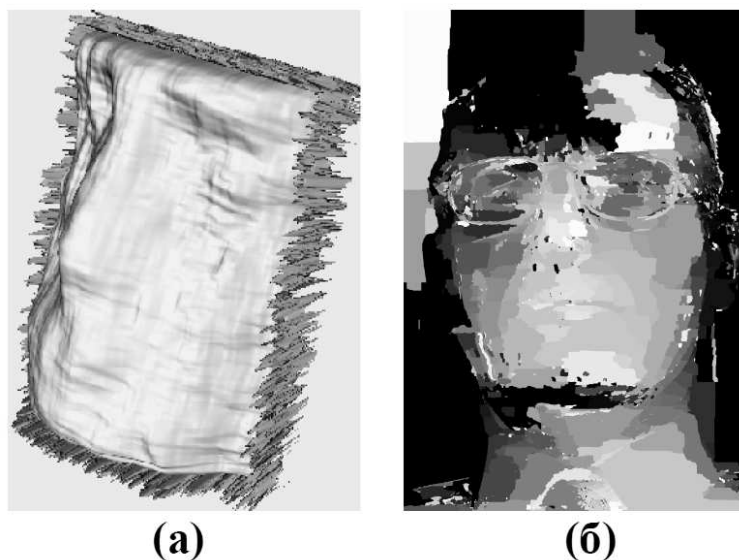


Рис. 4: Восстановленная трёхмерная структура лица

просматривать итоговую 3D модель с помощью программы DirectX Mesh Viewer. Модель сохраняется в файлах формата *.x [9].

Для тестирования метода, использующего 3D-модель также был написан пользовательский интерфейс, позволяющий выделять область на изображении и отмечать кончик носа, а также просматривать опорную 3D-модель и восстановленную трёхмерную структуру лица. Было проведено сравнение алгоритма Колмогорова с обоими методами (если считать построение опорной модели самостоятельным методом восстановления трёхмерной структуры лица), предложенными в этой статье, по двум параметрам. Первый параметр – время обработки изображений одного и того же размера. Было измерено время обработки всеми алгоритмами стереопары изображений размером 350×450 пикселей каждое. По результатам теста видно, что алгоритм Колмогорова (далее АК) обрабатывает изображения в четыре раза медленнее, чем алгоритм, использующий модель (далее АИМ). Результаты теста отображены в верхней части таблицы 1. Второй параметр измерений – время для достижения примерно одинакового результата. «Одинакового» в смысле видимого качества полученной 3D-структуры, т.к. для человеческого лица не существует «эталонной» стереопары изображений для численной оценки точности (как, например, стереопара Цукубы, использованная для оценки в [8]). Тестирование по второму параметру показало, что для достижения примерно такого результата, как на рис. 4 АК требуется изображения размером 1200×1700 пикселей, которые он обрабатывает около часа. АИМ получает такой результат за 5 минут. Результаты тестирования алгоритмов показаны в нижней части таблицы 1 и на рис. 4. Все тесты проводились на персональном компьютере с процессором Intel Celeron E3200 (частота 2,4 ГГц) и оперативной памятью 2 ГБ.

Таблица 1: Сравнение алгоритма Колмогорова (АК), алгоритма построения модели (АПМ) и алгоритма использования модели (АИМ).

Алгоритм	Размер изображения (пиксели)	Время обработки (с)
АК	350 × 450	120
АИМ	350 × 450	27
АК	1200 × 1700	3470
АИМ	500 × 800	300

Заключение

Предложенный в данной статье метод стереорекострукции, не применим для произвольной трёхмерной сцены, однако может быть успешно использован для обработки объектов, имеющих сходную форму, для которых возможно построение опорной 3D-модели. В приложении к лицу человека данный подход обеспечивает за малое время точность, достаточную для использования в некоторых медицинских задачах, например, при диагностике асимметрии лица [2].

Список литературы

- [1] Jain A.K., Ross A., Prabhakar S. An introduction to biometric recognition. IEEE Trans. Circuits and Systems for Video Technology. 2004. №14. V.1. P.4-20.
- [2] Knyaz V.A., Matveev I.A., Muryin A.B. Applying computer stereovision algorithms to study of correlation between face asymmetry and human vision pathology. Pattern Recognition and Image Analysis. 2009. №19. V.4. P.679-686.
- [3] Tang C.-K., Medione G., Duret F. Automatic, Accurate Surface Model Inference for Dental CAD/CAM. Proc. Int. Conf. Medical Image Comput. and Comput. Assist. Inter. Cambridge, MA, USA. 11-13 October 1998. P.732-742.
- [4] Шапиро Л., Стокман Д. Компьютерное зрение. М.:Бином. 2006. с.181, 221, 509.
- [5] Scharstein D., Szeliski R. A Taxonomy and Evaluation of Dence Two-Frame Stereo Correspondence Algorithms. Int. J. Computer Vision. 2002. №47. V.1. P.7-42.
- [6] Vincent E., Langanieri R. Matching Feature Points in Stereo Pairs: A Comparative Study of Some Matching Strategies. Machine Graphics and Vision. 2001. №10. V.3. P.237-259.
- [7] Melikov T., Nabiyeu V., Gunay A. Producing the Average Faces with Morphing. Applied and Computational Mathematics. 2011. in print.
- [8] Kolmogorov V., Zabih R. Multi-Camera scene Reconstruction via Graph Cuts. Proc. European Conf. Computer Vision. Copenhagen, Denmark. 27 May - 2 June 2002. Part III. P.8-40.
- [9] X File Reference: <http://msdn.microsoft.com/en-us/library/bb172982%28v=vs.85%29.aspx>

- [10] Bagherian E.R., Rahmita W.O.K. Facial feature extraction for face recognition: a review Proc. Int. Symp. Information Technology, 2008, V.2, P.1-9.
- [11] Мурынин А.Б., Матвеев И.А. Принципы построения системы стереоскопического зрения для управления движением робота Изв. РАН. Теория и системы управления. 2003. N.3. С.184-192.