

НЕКООПЕРАТИВНАЯ БИОМЕТРИЧЕСКАЯ ИДЕНТИФИКАЦИЯ ПО 3D-МОДЕЛЯМ ЛИЦА С ИСПОЛЬЗОВАНИЕМ ВИДЕОКАМЕР ВЫСОКОГО РАЗРЕШЕНИЯ

А.И. Манолов, А.Ю. Соколов, О.В. Степаненко, А.С. Тумачек, А.В.Тягт, А. К. Цискаридзе,
Д.Н. Заварикин, А.А. Кадейшвили,
Компания Vocord

Аннотация

Получены результаты по распознаванию лиц, основанные на 3D реконструкции без использования какой-либо структурированной подсветки. 3D реконструкция основана на использовании камер высокого разрешения. Вероятность распознавания составляет 92-98%.

Ключевые слова: 3D реконструкция, 3D распознавание лиц.

1. ВВЕДЕНИЕ

Системам распознавания лиц, основанным на двумерных изображениях, присущи определенные недостатки. Такие системы чувствительны к изменениям яркости. Свет, собранный с лица, является функцией геометрии лица, отражательной способности лица, свойствами источника света и свойствами камеры. С учетом этого, сложно создать модель лица, которая учитывает одновременно все вариации этих факторов. В трехмерном распознавании изменения в освещенности влияют лишь на текстуру на лице, в то время как реконструкция поверхности сохраняет свои свойства [1].

Другим отличительным фактором между 2D и 3D распознаванием является эффект изменения позы лица (ракурса камеры). Чтобы компенсировать этот эффект, в двумерном распознавании использовалась трансформация изображения к каноническому положению [2]. Тем не менее, эффективность данного подхода зависит от точности расположения антропометрических точек на лице, и плохо работает при затенении части лица. Более того, данная задача в двумерном виде не может быть решена теоретически строго из-за свойства перспективной проекции двумерных изображений. Чтобы обойти эти проблемы, было предложено сохранять в базе данных различные виды лица под разным ракурсом [3]. Такой подход, тем не менее, требует сохранения большого количества двумерных изображений и снижает производительность системы распознавания. Были предложены статистические модели, учитывающие изменения позы лица [4, 6], а также обобщенные модели лица [5]. С использованием 3D изображений, интерполяция положения может быть сведена к повороту восстановленной 3D модели лица к новой позе. Было предложено использовать 3D модель, реконструированную из двумерных нефронтальных изображений, и воссоздать фронтальный вид лица для 2D распознавания [7].

Другой проблемой 2D распознавания, связанной с вариацией положения лица, является то, что физические размеры лица нельзя полностью восстановить из 2D изображений. Размер лица на двумерном изображении существенно зависит от расстояния между лицом и камерой. В 3D распознавании физические размеры лица известны и могут использоваться явно для распознавания.

Традиционные системы 2D распознавания используют участки изображения с высоким контрастом, таких как глаза, рот, нос, границы лица и плохо используют информацию в областях низкого контраста – на щеках, лбу, подбородке и др. В противоположность 2D, 3D распознавание включает для анализа и информацию на данных участках. Более того, информация о форме поверхности, сосредоточенная на лобных костях и других слабо деформируемых участках, мало подвержена изменениям при различных выражениях лица (как например, улыбка и др.).

3D распознавание, тем не менее, не является полностью идеальным. Например, освещение не является проблемой на этапе 3D распознавания, но оно может отрицательно повлиять на результат 3D реконструкции формы лица. В зависимости от алгоритма реконструкции, некоторые части лица (например, засвеченные участки, либо области с очень низким контрастом) могут внести пробелы в реконструкцию, или артефакты (выбросы на поверхности). Другим недостатком 3D распознавания также является высокая стоимость оборудования. Кроме этого, 3D распознавание требует больших вычислительных ресурсов, чем 2D.

Наконец, до настоящего времени, по-видимому, недостаточное внедрение 3D систем распознавания было связано и с отсутствием на рынке видеосенсоров высокого разрешения. Разработанные компанией Vocord доступные по стоимости видеокамеры NetCam и проводимые исследования в области 3D-распознавания с использованием этих камер могут существенно повлиять на внедрение 3D распознавания в России.

Среди различных подходов 3D распознавания можно выделить три основных направления: анализ формы 3D поверхности лица, статистические подходы, и использование параметрической модели лица.

Методы, основанные на анализе формы, используют непосредственно геометрию поверхности, которая описывает лицо. Эти подходы можно классифицировать на те, которые используют либо локальные, либо глобальные свойства поверхности (например, кривизну), на подходы, основанные на профили линий, либо на методы, которые используют метрику расстояний между двумя поверхностями.

В работе [8] было предложено использовать кривизну поверхности для сегментации поверхности лица на признаки, которые могут быть использованы для сравнения поверхностей. Другой подход основан на 3D дескрипторах поверхности лица в терминах средней и гауссовой кривизны [9], либо в терминах расстояний и отношении углов между характерными точками поверхностей [10].

Другим локально-ориентированным методом является подход, использующий точки-сигнатуры [11]. Идея метода заключается в формировании представления-описания точки по соседним точкам вокруг заданной точки поверхности. Эти сигнатуры точек используются для сравнения поверхностей.

Было также предложено удалять из рассмотрения те части поверхности, которые подвержены изменениям в результате изменения мимики на лице. Только жесткие части лица являются информацией для распознавания. Такой подход применен в работе [12]. Кроме того, в этой же работе дополнительно с 3D информацией 2D использовалась текстурная информация на участках лица.

Существуют также и гибридные методы, основанные на объединении локальной информации о поверхности в виде локальных моментов, с глобальной трехмерной сеткой, описывающей все лицо [13].

Глобальными методами являются такие методы, которые задействуют все 3D лицо как информацию на входе системы распознавания. В одной из ранних систем лицо выравнивалось на основе его зеркальной симметрии [14], после чего выделялись и сравнивались профили лица вдоль плоскости выравнивания. В работе [15] было предложено сравнение лиц на основе максимального и минимального значений и направления кривизны профилей.

Другой подход основан на методе сравнения расстояний между поверхностями для распознавания. В работах [16] было предложено использовать метрику наименьших расстояний для вычисления расстояния между поверхностями. Было также предложено измерять расстояние не только между поверхностями, но и текстурой на поверхности [17]. Существенным ограничением данных методов является то, что лицо не может деформироваться и его поверхность является жесткой.

Другим подходом является извлечение и анализ трехмерных профилей и контуров, выделенных на лице [18].

Статистические методы, например, метод главных компонент (principal component analysis - PCA) широко использовались ранее в 2D распознавании. Недавно метод PCA был реализован для 3D распознавания [19, 20, 21]. PCA был расширен одновременно на комбинацию карты глубины и цвета [22]. Альтернативным для PCA является метод линейного дискриминантного анализа [23]. В отличие от PCA, в линейном дискриминантном анализе один объект (заданный человек) задается не одним лицом, а набором моделей (3D лиц).

До сих пор все описанные статистические методы не учитывали эффект изменения формы поверхности лица из-за его различной мимики. Для минимизации этого эффекта были разработаны подходы, основанные на инвариантных изоморфных преобразованиях. Такими преобразованиями являются те, которые не изменяют расстояния между двумя заданными точками на лице под действием мимических изменений формы лица. В работе [24, 25] было предложено для этого использовать преобразование формы лица к каноническому виду. Подобные методы использовали PCA на заключительной стадии распознавания, который применялся к канонической форме лица.

Существуют также способы распознавания, основанные на параметрических моделях лица. Ключевая идея распознавания по моделям основана на так называемых параметрических 3D моделях. В данном подходе форма лица контролируется набором параметров модели. Эти коэффициенты описывают 3D форму лица и также могут задавать цвет (текстуру) на ее поверхности. Подобная модель затем проецируется на двумерные изображения, из которых определяются параметры коэффициентов модели для данного изображения [26]. Недостатком метода является высокая вычислительная сложность и чувствительность к начальной инициализации параметров модели. Для преодоления этих сложностей были придуманы модели, состоящие из независимых участков [27].

Было также предложено использовать трехмерную поверхность среднего лица, которая с использованием анатомических антропометрических точек на лице деформируется до заданной трехмерной поверхности [28]. Параметры деформации использовались в качестве отличительных признаков для данного лица.

К сожалению, как отмечается в работе [17] прямое сравнение различных 3D алгоритмов распознавания по их эффективности крайне затруднено. Связано это с отсутствием общей стандартизированной базы данных трехмерных лиц и различием точности моделей 3D изображений. Тип представления и пространственное разрешение 3D данных существенно различается среди публикаций. По-видимому, стандарты на 3D изображения и стандартизированные базы изображений появятся в самое ближайшее время.

Среди наиболее современных методов, согласно обзору в [29], выделяются глобальные методы (вероятность распознавания 90-96%) и статистические методы (93-100%). Параметрические модели лица характеризуются вероятностью порядка 88-96%.

Среди рассматриваемых моделей, которые предложены до настоящего времени для компенсации мимики на лице, изоморфические модели (приведение к каноническим формам лица) являются, видимо, наиболее перспективными [24,25].

Исходя из анализа современного состояния проблемы компанией Vocord было принято решение использовать и разработать систему, основанную именно на 3D реконструкции лица с последующим его распознаванием, обеспечивающей достаточно высокие показатели, на базе выпускаемых камер Netcam высокого разрешения.

2. 3D РЕКОНСТРУКЦИЯ ПОВЕРХНОСТИ ЛИЦА

Реконструкция выполняется с использованием камер высокого разрешения собственной разработки Vocord NetCam. Данные камеры обеспечивают разрешение до 2048 x 1576 пикселей, частота кадров до 12 кадров/сек.



с.1 Камеры Vocord NetCam

Ри

Особенностью алгоритма является отсутствие какого либо лазерного или другого дополнительного света, направленного на лицо человека (структурированной подсветки). Реконструкция проводится в условиях минимальной освещенности (400-600 люкс) на расстоянии до 2.5 м от камер. Использование камер разрешением 5 Мегапикселей, выпуск которых намечен в скором плане компании, позволит осуществить реконструкцию с теми же параметрами на расстоянии 3-4 м до объективов.

Особенностью алгоритма является также возможность использовать реконструкцию по многим камерам. Например, оптимальным является использование трех

синхронизированных во времени камер, разнесенных друг относительно друга на расстояние порядка 20 см между камерами. Такой режим улучшает среднестатистические показатели реконструированной поверхности.

Кроме того, помимо высокого исходного разрешения камер, разработанный алгоритм поиска точек сопряжения изображений на стереопарах использует субпиксельную точность до $\frac{1}{4}$ пиксела. В результате, суммарная эффективная погрешность, обусловленная дискретностью сенсоров камер, имеет порядок 0.3-0.5 мм, что достаточно для дальнейшего высококачественного распознавания.

3D реконструкция модели лица осуществляется поэтапно:

- построение поля 3D точек;
- предварительная фильтрация точек;
- построение и окончательной фильтрация триангулированной модели лица (рис 2, 3);
- совмещение нескольких поверхностей лица человека, полученных в разные моменты времени (видеоряд), в одну общую объединенную 3D модель (рис.4, 5).



(a)



(б)

Рис. 2 Пример 3D реконструкции лица. (a) – лицо с текстурой, (б) – трангулированная модель.

В настоящее время, 3D реконструкция с использованием 4х ядерного процессора 2.9 ГГц занимает от 2 до 10 сек, в зависимости от настроек параметров системы. Компания продолжает работы по оптимизации алгоритма 3D реконструкции, в частности, с использованием графических плат акселераторов на платформе NVIDIA CUDA.

3. 3D РАСПОЗНАВАНИЕ

Распознавание выполняется в два этапа:

- (1) быстрое сравнение по базе данных с использованием вектора признаков лица небольшой длины;
- (2) уточнение результата по отобранным нескольким наиболее похожим изображениям из базы данных.

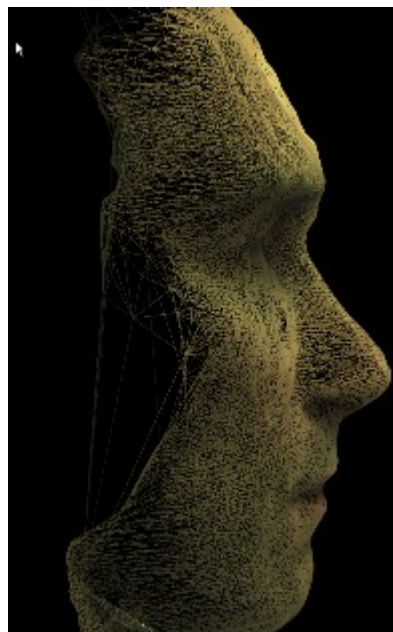
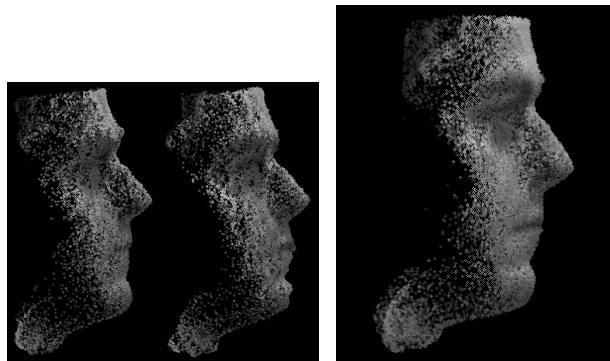


Рис.3 Пример 3D реконструкции лица (вид в профиль поперек к направлению на камеры).



(a)

(б)

Рис. 4. Обработка 3D видео. (a) – 3D модели лица в разные моменты времени, (б) – после совмещения 3D моделей в общую модель.

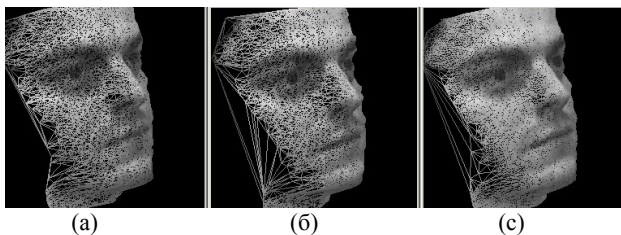


Рис. 5. Обработка 3D видео. (а), (б) – 3D модели в разные моменты времени, (с) – после совмещения 3D моделей в общую модель.

Система 3D распознавания, разработанная в Vocord, также может использовать приведение 3D поверхности к форме, более устойчивой при мимических искажениях (Рис.6) [24, 25].

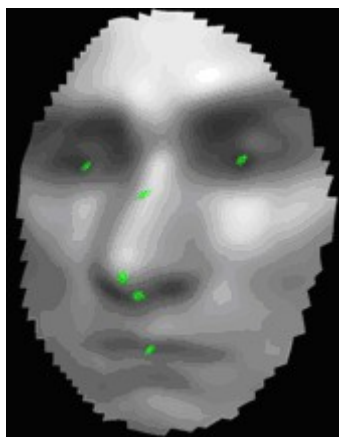


Рис. 6. Использование геодезического преобразования к маске лица

4. РЕЗУЛЬТАТЫ

Тестирование проводилось на базе данных 100 лиц, из которых 50 чел были внесены в базу данных, 50 чел были незнакомыми.

Для неподвижно стоящих людей вероятность распознавания составила $P=98\%$ при $FAR=FRR=2\%$. Для людей, движущихся навстречу камеры со скоростью примерно 3 км/ч, были продемонстрированы результаты $FAR=FRR=7\%$, вероятность распознавания $P=93\%$.

Разработанный алгоритм также устойчив к поворотам лица в пределах $+50/-30$ градусов (вверх-вниз, и примерно ± 30 градусов влево –вправо, см. Рис.7, 8, таблицу 1. На тестовых изображениях диапазон максимальных углов 3D распознавания выше диапазона 2D распознавания.

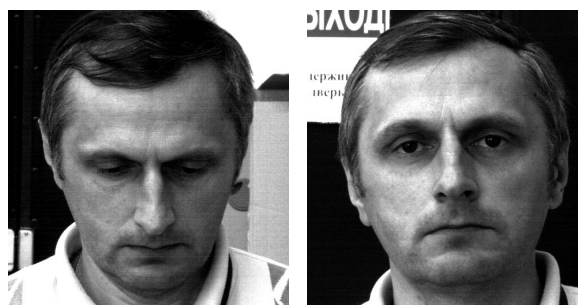
Таблица 1. Сравнение Vocord-3D и Cognitec-2D [30]. Человек считается распознанным, если мера сходства более 50%.

Угол наклона головы, град	Vocord 3D %	Cognitec 2D (Ver. 8.1, algorithm B.4) %
Вниз		
30	94	43.3
40	32	8
Вверх		
45	96	99

50	91.4	39.6
----	------	------



Рис. 7 Мера сходства лица с 3D моделью того же человека, хранимой в базе данных, в зависимости от положения головы.



(а) Взгляд вниз Vocord-3D: 94% Cognitec-2D: 43%



(б) Взгляд вверх Vocord-3D: 91.4%, Cognitec-2D: 39.6%

Рис. 8 Сравнение эффективности с двумерным алгоритмом распознавания Cognitec (Version 8.1, algorithm B4) [30]. Слева – тестируемое изображение, справа – фото из базы

данных. Человек считается распознанным, если мера сходства более 50%.

5. ЗАКЛЮЧЕНИЕ

Компанией Vocord разработан прототип системы 3D распознавания с использованием камер высокого разрешения, без использования какой-либо дополнительной подсветки. Данные преимущества позволяют использовать разработанную систему для некооперативной идентификации людей на основе 3D моделей. Показаны высокие характеристики распознавания с вероятностью 92-98%. На некоторых изображениях (большой угол поворота головы, см. Табл.1) 3D алгоритм имеет преимущество по сравнению с 2D распознаванием.

6. СПИСОК ЛИТЕРАТУРЫ:

[1] Heshner, C., Srivastava, A., and Erlebacher, G. (2003). A novel technique for face recognition using range imaging. *International Symposium on Signal Processing and Its Applications*, стр 201–204.

[2] Kim, T. and Kittler, J. (2005). Locally linear discriminant analysis for multimodally distributed classes for face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(3):318–327.

[3] Li, Y., Gong, S., and Lidell, H. (2000). Support vector regression and classification based multiview face detection and recognition. *International Conference on Face and Gesture Recognition*, стр 300–305.

[4] Lanitis, A., Taylor, C., and Cootes, T. (1995). Automatic face identification system using flexible appearance models. *Image and Vision Computing*, 13(5):393–401.

[5] Prince, S. and Elder, J. (2006). Tied factor analysis for face recognition across large pose changes. *British Machine Vision Conference*.

[6] Gross, R., Matthews, I., and Baker, S. (2002). Eigen light-fields and face recognition across pose. *International Conference on Automatic Face and Gesture Recognition*.

[7] Blanz, V., Grother, P., Phillips, J., and Vetter, T. (2005). Face recognition based on frontal views generated from non-frontal images. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, стр 454–461.

[8] Gordon, G. (1992). Face recognition based on depth and curvature features. *IEEE Computer Vision and Pattern Recognition (CVPR)*, стр 808–810.

[9] Moreno, A., Sanchez, A., Velez, J., and Diaz, F. (2003). Face recognition using 3D surface extracted descriptors. In *Irish Machine Vision and Image Processing Conference*.

[10] Lee, Y., Song, H., Yang, U., Shin, H., and Sohn, K. (2005). Local feature based 3D face recognition. *International Conference on Audio- and Video-based Biometric Person Authentication*, стр 909–918.

[11] Chua, C. and Jarvis, R. (1997). Point signatures - a new representation for 3D object recognition. *International Journal of Computer Vision*, 25(1):63–85.

[12] Wang, Y., Chua, C., and Ho, Y. (2002). Facial feature detection and face recognition from 2D and 3D images. *Pattern Recognition Letters*, 23:1191–1202.

[13] Xu, C., Wang, Y., Tan, T., and Quan, L. (2004). Automatic 3D face recognition combining global geometric features with local shape variation information. In *International Conference on Automated Face and Gesture Recognition*, pages 308–313.

[14] Cartoux, J., LaPreste, J., and Richetin, M. (1989). Face authentication or recognition by profile extraction from range

images. *Workshop on Interpretation of 3D Scenes*, стр 194–199.

[15] Tanaka, H., Ikeda, M., and Chiaki, H. (1998). Curvature-based face surface recognition using spherical correlation principal directions for curved object recognition. *International Conference on Automated Face and Gesture Recognition*, pages 372–377.

[16] Ackermann, B. and Bunke, H., et al. (2000). *15th International Conference on Pattern Recognition*, стр 809–813.

[17] Papatheodorou, T. and Rueckert, D. (2004). Evaluation of automatic 3D face recognition using surface and texture registration. *International Conference on Automated Face and Gesture Recognition*, стр 321–326.

[18] Nagamine, T., Uemura, T., and Masuda, I. (1992). 3D facial image analysis for human identification. *International Conference on Pattern Recognition*, pages стр 324–327.

[19] Mavridis, N., Tsalakanidou, F., Pantazis, D., Malasiotis, S., and Strintzis, M. (2001). The hiscore face recognition application: Affordable desktop face recognition based on a novel 3D camera. *International Conference on Augmented Virtual Environments and 3D Images*, стр 157–160.

[20] Chang, K., Bowyer, K., and Flynn, P. (2003). Face recognition using 2D and 3D facial data. *Multimodal User Authentication Workshop*, стр 25–32.

[21] Papatheodorou, T. and Rueckert, D. (2005). Evaluation of 3D face recognition using registration and PCA. *Audio- and Video-based Biometric Person Authentication*, стр 997–1009.

[22] Tsalakanidou, F., Tzocaras, D., and Strintzis, M. (2003). Use of depth and colour eigenfaces for face recognition. *Pattern Recognition Letters*, 24:1427–1435.

[23] Gokberk, B., Salah, A., and Akarun, L. (2005). Rank-based decision fusion for 3D shapebased face recognition. In *International Conference on Audio- and Video-based Biometric Person Authentication*, стр 1019–1028.

[24] Bronstein, A., Bronstein, M., and Kimmel, R. (2003). Expression-invariant 3D face recognition. *International Conference on Audio- and Video-Based Person Authentication*, стр 62–70.

[25] Bronstein, A., Bronstein, M., and Kimmel, R. (2005). Three-dimensional face recognition. *International Journal of Computer Vision*, 64:5–30.

[26] Blanz, V., Romdhani, S., and Vetter, T. (2002). Face identification across different poses and illuminations with a 3D morphable model. In *International Conference on Automatic Face and Gesture Recognition*, стр 202–207.

[27] Huang, J., Heisele, B., and Blanz, V. (2003). Component-based face recognition with 3D morphable models. In *International Conference on Audio- and Video-Based Person Authentication*.

[28] Passalis, G., Kakadiaris, I., Theoharis, T., Toderici, G., and Murtuza, N. (2005). Evaluation of 3D face recognition in the presence of facial expressions: an annotated deformable model approach. стр. 1022 – 1029.

[29] Papatheodorou, T. and Rueckert, D. (2007). 3D face recognition, книга Face recognition, Vienna, Austria, 2007, стр 417–423.

[30] <http://www.cognitec-systems.de/FaceVACS-SDK.19.0.html>