но и поможет избежать незапланированных затрат при ее дальнейшем масштабировании.

1. Бережной А.В. Вертикальная дифференциация ландшафтов среднерусской лесостепи: монография / А.В. Бережной, А.С. Горбунов, Т.В. Бережная; под ред. А.В. Бережного. – Воронеж: Научная книга, 2007. – 274 с.
2. Вишневская Ирина Александровна. Компьютерная технология оценки зон затопления при наводнениях : дис. канд. техн. наук : 05.11.13 Москва, 2006 145 с. РГБ ОД, 61:07-5/1515
3. Грачев, А. В. Информационные технологии в экологии и природопользовании [Электронный ресурс] : учеб. пособие / В. Ю. Орлов, Яросл. гос. ун-т им. П. Г. Демидова, А. В. Грачев .— Ярославль : ЯрГУ, 2013 .— 108 с. — ISBN 978-5-8397-0960-7 .— Режим доступа: <https://rucont.ru/efd/272162>
4. Шашкова, И.Г. Информационные технологии на транспорте [Электронный ресурс] / Н.В. Бышов, Е.В. Лунин, В.С. Конкина, Е.И. Ягодкина, И.Г. Шашкова .— 2014 .— 300 с. : ил. — ISBN 978-5-98660-208-0 .— Режим доступа: <https://rucont.ru/efd/275481>
5. <http://www.radixtools.ru/publish-gis-tech>
6. http://egov.ifmo.ru/news_egov/news_17_09_25-0
7. <https://www.osp.ru/cw/1996/06/10265>

Жиленков И.В.

Анализ алгоритмов и методов распознавания человеческого лица в цифровом видеосигнале

*Севастопольский государственный университет
(Россия, Севастополь)*

doi:10.18411/sciencepublic-12-08-2020-02

idsp: sciencepublic-12-08-2020-02

Аннотация

В данной работе производится анализ актуальных подходов к распознаванию человеческого лица в цифровом видеосигнале, выявление их недостатков.

Ключевые слова: цифровой видеосигнал, распознавание образов.

В научной теории и практике существует большое количество различных подходов к выделению и распознаванию человеческого лица в цифровом видеосигнале, однако данная область является достаточно новой и до сих пор не изучена в полном объеме. Существующие подходы для распознавания лиц и жестов обладают недостаточной точностью, надежностью и скоростью в сложной реальной обстановке, характеризующейся присутствием шумов на изображениях, что обуславливает необходимость в совершенствовании существующих методов.

Для усовершенствования существующих подходов к распознаванию человеческого лица в цифровом видеосигнале необходимо проанализировать их и выделить недостатки, требующие доработки. В данной работе для анализа взяты наиболее часто используемые из них:

1. Метод гибкого сравнения на графах (Elastic graph matching);
2. Метод главных компонент (PCA);
3. Метод Виолы-Джонса;
4. Метод сопоставления шаблонов (Template Matching);
5. Метод опорных векторов (SVM);
6. Активные модели внешнего вида (AAM);
7. Активные модели формы (ASM);
8. Алгоритм CAMShift (Continuously AdaptiveMeanShift);
9. Алгоритм Гистрограммы направленных градиентов (HOG).

Рассмотрим метод гибкого сравнения на графах (Elastic graph matching). Метод основывается на эластичном сопоставлении графов, описывающих изображения лиц, со взвешенными вершинами и ребрами [1]. На этапе распознавания один из графов (эталонный) остается неизменным, в то время как другой деформируется с целью наилучшей подгонки к первому. Ребра графа взвешиваются расстояниями между смежными вершинами. Различие между двумя графами вычисляется при помощи ценовой функции деформации. Недостатками данного метода являются высокая вычислительная сложность процедуры распознавания, низкая технологичность при запоминании новых эталонов и линейная зависимость времени работы от размера базы данных лиц.

Метод главных компонент (PCA) преобразует обучающий набор лиц в общую матрицу данных, где каждая строка представляет собой один экземпляр изображения лица, разложенного в строку. Затем производится нормировка данных и приведение строк к 0-му среднему и 1-й дисперсии, вычисляется матрица ковариации, для которой решается задача определения собственных значений и соответствующих им собственных векторов (собственные лица) [2]. Данный метод обладает следующими недостатками: способен находить только линейные подпространства исходного пространства, однако на практике «поверхность» может быть значительно нелинейной; является инвариантным относительно поворота координат в пространстве латентных переменных, что делает восстановление их значений неоднозначным.

Метод Виолы-Джонса основан на трех главных идеях: интегральном представлении изображения, методе построения классификатора на основе алгоритма адаптивного бустинга (AdaBoost) и методе комбинирования классификаторов в каскадную структуру, что позволяет построить детектор, способный работать в режиме реального времени [3]. Существенным является то, что при угле наклона больше 30° вероятность обнаружения лица резко падает, что не позволяет в стандартной реализации детектировать лицо человека, повернутое под произвольным углом, и является недостатком данного метода.

Метод сопоставления шаблонов (Template Matching) применяется для распознавания изображений, аналогичных некоему заданному шаблону [4]. Цель работы данного алгоритма – найти на изображении область, которая в наибольшей степени совпадет с шаблоном объекта. Из недостатков метода мы выделили следующие: требует много ресурсов как для хранения участков, так и для их сравнения; обладает низкой скоростью работы; изображения для сравнения должны быть сняты в строго установленных условиях – не допускается заметных изменений ракурса, освещения, физических помех.

Рассмотрим метод опорных векторов (Support Vector Machines, SVM). Основной целью метода является перевод исходных векторов в пространство более высокой размерности и поиск разделяющей гиперплоскости с максимальным зазором в этом пространстве [5]. Метод основан на предположении о том, что чем больше разница или расстояние между этими параллельными гиперплоскостями, тем меньше будет средняя ошибка классификатора. Метод обладает следующими недостатками: он чувствителен к шумам и стандартизации данных; не имеет отбора признаков; необходимо подбирать оптимизируемый коэффициент C при помощи кросс-валидации; не существует общего подхода к автоматическому выбору ядра (и построению спрямляющего подпространства в целом) в случае линейной неразделимости классов.

Далее рассмотрим активные модели внешнего вида (AAM). Они содержат два типа параметров: параметры формы, и параметры, связанные со статистической моделью пикселей изображения или текстурой (параметры внешнего вида) [6]. Перед использованием модель должна быть обучена на множестве заранее размеченных изображений, которая производится вручную. Каждая метка имеет свой номер и определяет характерную точку, которую должна будет находить модель во время

адаптации к новому изображению. Главным недостатком данного метода является необходимость большого объема вычислений для его использования.

Активные модели формы (ASM) также часто используются в данной области. Суть метода ASM заключается в учете статистических связей между расположением антропометрических точек на имеющейся выборке изображений лиц, снятых в анфас [5]. На каждом изображении антропометрические точки пронумерованы в одинаковом порядке. Однако главной целью ASM является не распознавание лиц, а точная локализация лица и антропометрических точек на изображении для дальнейшей обработки. Недостатками данного метода являются: частичное использование изображения, необходимость хорошей разметки, рассмотрение локальных моделей как независимых.

Алгоритм CAMShift (Continuously Adaptive Mean Shift) используется при отслеживании объектов на основе информации о цвете движущейся в кадре цели [7]. Также он комбинирует алгоритм отслеживания объекта Mean Shift, основанный на карте вероятности цвета кожи, с адаптивным шагом изменения размера области отслеживания. Однако это обуславливает и недостаток алгоритма Camshift, состоящий в чувствительности к информации о цвете: он не может отследить желаемую цель, когда фон (или объект рядом) имеет тот же цвет.

Алгоритм гистограммы направленных градиентов (HOG) основан на подсчете количества направлений градиента в локальных областях изображения [8]. Данный метод производит вычисления на плотной сетке равномерно распределенных ячеек и использует нормализацию перекрывающегося локального контраста для увеличения точности. Данный алгоритм обладает следующими недостатками: для работы метода необходимо наличие большого набора изображений объекта и еще больший набор изображений, на которых объекта нет; все изображения необходимо масштабировать до фиксированного размера.

Таким образом, видно, что на сегодняшний день существует большой нерешенный спектр задач, связанный с распознаванием видеосигналов на фоне общего потока видеoinформации, таких как частота распознавания человеческих лиц, распознавание лиц в условиях плохой освещенности, повышение точности распознавания лица по его части, высокая вычислительная трудоемкость. Выделенные в данной работе недостатки существующих методов, моделей и алгоритмов показывают направления дальнейшей разработки.

1. Wiskott, L Face Recognition by Elastic Bunch Graph Matching /L. Wiskott, J. Fellous, N. Krüger, C. von der Malsburg// InIntelligent Biometric Techniques in Fingerprint and Face Recognition. – 1999. – Ch. 11. – P. 355-396.
2. Спицын В.Г. Распознавание лиц на основе метода главных компонент с применением вейвлет-дескрипторов Хаара и Добеши/ В.Г. Спицын, Ю.А. Болотова, Н.В. Шабалдина, Б.Т.Т. Чанг, Ф.Н. Хоанг// Научная визуализация. – 2016. – Т. 8. – №5. – С. 103-112.
3. Viola, P. Rapid object detection using a boosted cascade of simple features/P. Viola, M.J. Jones//IEEE Conf. on Computer Vision and Pattern Recognition. Kauai, Hawaii, USA. – 2001. – V. 1. – P. 511-518.
4. Половинка, Д.В. Применение технологий mobilevision в разработке мобильных приложений/Д.В. Половинка// Теоретическая и прикладная информатика. – 2015. – С. 21-28.
5. Le Hoang Thai, Vo Nhat Truong Face Alignment Using Active Shape Model And Support Vector Machine/ Le Hoang Thai, Vo Nhat// International Journal of Biometrics and Bioinformatics. – 2015. – P. 224 -234.
6. Paul, I. Active Appearance Modelsfor Face Recognition/ Ivan Paul. – Vrije Universiteit Amsterdam, 2004. – 39 P.
7. Буй, Т.Т.Ч. Распознавание жестов на видеопоследовательности в режиме реального времени на основе применения метода Виолы Джонса, алгоритма CAMShift, вейвлет-преобразования и метода главных компонент / Т.Т.Ч. Буй, Н.Х. Фан, В.Г. Спицын // Вестник Томского гос. у-та. – 2013. – № 2(23). – С. 102-111.
8. Dalal, N. Histograms of oriented gradients for human detection/ N. Dalal, B. Triggs// Computer Vision and Pattern Recognition: IEEE Comput. Soc. Conf. – 2005. – V. 1. – P. 886-893.