



УДК 004.932.2

DOI: 10.22184/NanoRus.2019.12.89.156.158

БЫСТРЫЙ АЛГОРИТМ ОБНАРУЖЕНИЯ И РАСПОЗНАВАНИЯ ЛИЦ ДЛЯ СЕМАНТИЧЕСКОГО АНАЛИЗА ИЗОБРАЖЕНИЙ НА DSP-ядрах ELcore

THE FAST FACE RECOGNITION ALGORITHM FOR SEMANTIC IMAGE ANALYSIS ON ELcore DSP-cores

КОСТУЛИН НИКОЛАЙ ВЛАДИМИРОВИЧ
vkostulin@elvees.com

KOSTULIN NIKOLAY V.
vkostulin@elvees.com

ЯНАКОВА ЕЛЕНА СЕРГЕЕВНА
helen@elvees.com

YANAKOVA ELENA S.
helen@elvees.com

АО Научно-производственный центр
«Электронные вычислительно-информационные
системы»
124498, г. Москва, г. Зеленоград, проезд №4922, 4, стр. 2

R&D Center "ELVEES", JSC
bld. 2, 4 Lane 4922, Zelenograd, Moscow, 124498, Russia

Рассмотрены этапы семантического анализа изображений с целью обнаружения и распознавания лиц, выделены и реализованы наиболее ресурсоемкие алгоритмы на DSP-ядрах ELcore. Общий тракт обработки изображений на DSP-ядрах ELcore с целью обнаружения и распознавания лиц составляет не более 32 мс, что соответствует требованиям по обработке видеок кадров в реальном времени.

Ключевые слова: семантический анализ; распознавание лиц; DSP-ядра.

The paper considers the stages of the image semantic analysis for faces detection and recognition. The most resource-intensive algorithms have been identified and implemented on ELcore DSP-cores. The common path of processing images for face recognition on ELcore DSP cores is no more than 32 ms, which corresponds to the requirements for real time video processing.

Keywords: semantic analysis; face recognition; DSP-cores.

ВВЕДЕНИЕ

В настоящее время алгоритмы семантического анализа изображений вызывают большой интерес в связи с бурным развитием компьютерной техники, интернет-технологий, робототехники, повышением требований к точности и качеству систем анализа видеоизображений, предназначенных, в частности, для охраны объектов различной важности [1]. Под анализом изображений понимается извлечение из них значимой информации. Семантический анализ является вершиной иерархической процедуры обработки изображений (рис. 1). Частными задачами семантического анализа являются распознавание лиц, позы, действий человека, распознавание объектов, сегментация, классификация и поиск изображений, обнаружение и распознавание текста.

Распознавание лица и определение его характеристик является одной из сложных и привлекательных задач семантического анализа изображений, алгоритмы которых решают общую

задачу верификации (проблема соответствия 1:1) и идентификации лица (проблема соответствия 1:N), учитывая наличие базы данных лиц [3]. В последние несколько лет наблюдается повышенный интерес к распознаванию аффективного психического состояния, когнитивных нарушений и интерпретации социальных сигналов на основе выражений лица и его поведения [4].

Лицо является важным каналом невербальной коммуникации, поэтому анализ лица необходим в различных приложениях в медицине для понимания и диагностирования состояния человека (депрессия, посттравматические стрессовые расстройства и другие), в системах помощи водителям (ADAS-системах), образовании, в охранных системах, системах компьютерного зрения, развлечениях и других системах. Такие системы для работы требуют встраиваемых решений с применением эффективных алгоритмов, экономично использующих соответствующие аппаратные вычислительные ресурсы в реальном времени.

Основной тракт обработки изображений в целях обнаружения и распознавания лиц является достаточно сформировавшимся процессом (рис. 2) и незначительно отличается в разных системах. Основные различия заключаются в наборе извлекаемых характеристик лица [3].

В настоящее время основная научная проблематика в области распознавания лиц и определения их характеристик заключается в разработке путей преодоления ограничений, обусловленных имеющимися ресурсами: возможностями элементной базы, допустимой величиной программно-аппаратных затрат; поэтому актуальной задачей является разработка быстрых алгоритмов и подходов к обнаружению и распознаванию лиц и их характеристик на встраиваемых вычислительных устройствах,

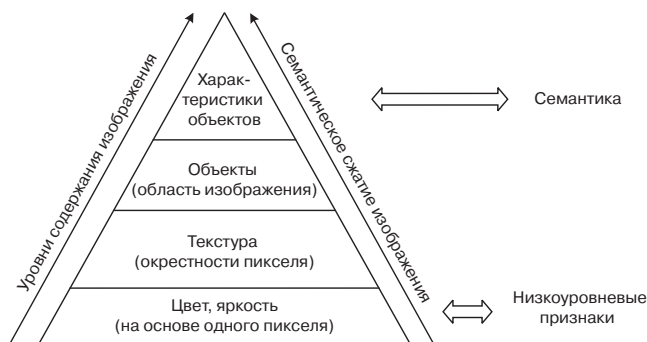


Рис. 1. Уровни семантического анализа изображений

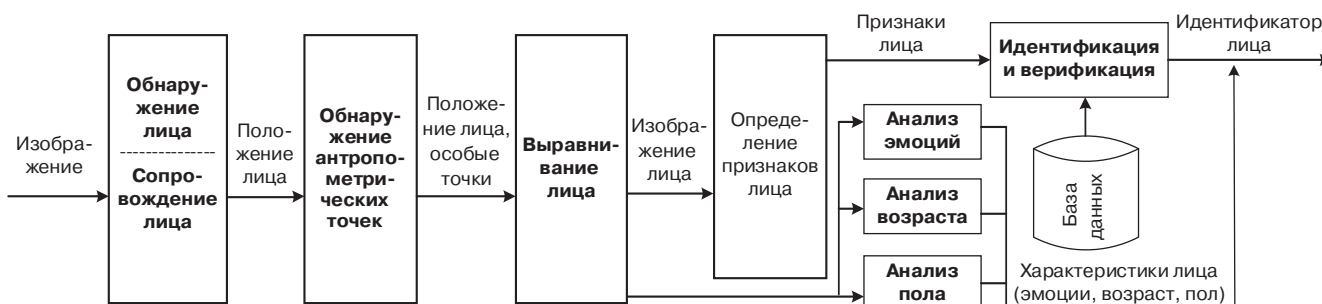


Рис. 2. Основной тракт обработки изображения в целях обнаружения и распознавания лиц

применяемых в составе информационно-измерительных и управляющих систем.

Устойчивое непрерывное развитие вычислительных процессоров и ядер компании НПЦ «ЭЛВИС» [5], постоянное повышение технико-экономической эффективности обработки сигналов на DSP-ядрах ELcore позволяют решать новые задачи компьютерного зрения, связанные с семантическим анализом сигналов и изображений путем синтеза достоинств программного и аппаратного подходов.

Целью работы является решение задачи обнаружения и распознавания лиц путем обеспечения высокого качества обработки с использованием аппаратных преимуществ отечественных DSP-ядер ELcore для применения во встраиваемых системах реального времени.

БЫСТРЫЙ АЛГОРИТМ ДЕТЕКТИРОВАНИЯ И РАСПОЗНАВАНИЯ ЛИЦ

В настоящее время используется несколько подходов к построению алгоритмов распознавания лиц [3]: с использованием алгоритмов машинного (Machine Learning) и глубокого обучения (Deep Learning). Синтез подходов обеспечивает наибольшую производительность с высокими характеристиками обнаружения. Такой комплексный подход, применяемый для решения задачи распознавания лиц, позволил получить российским компаниям, в частности 3DiVi [6], первые места в международном тестировании программного обеспечения для распознавания лиц NIST 2017 [7].

Наиболее ресурсоемким алгоритмом тракта (рис. 2) является алгоритм обнаружения и выравнивания лица с определением вектора признаков, по которому выполняются его верификация и идентификация. Основные принципы для построения эффективного детектора [8] заключаются в использовании множества слабых каскадных классификаторов с простыми функциями на первом этапе и нейросетевым алгоритмом на втором этапе для принятия решения о наличии лица на изображении с высокой вероятностью правильного обнаружения. Каскадный классификатор представляет собой дерево решений (каскадный детектор), а именно дерево регрессии с использованием бинарных NPD-признаков (Normalized Pixel Difference) [9]. Каскадный детектор описывается выражением (1):

$$f^N = \sum_{i=1}^N C^i(x), \tag{1}$$

где $C^i(x)$ — слабый классификатор, x — окно изображения. Окно изображения отбрасывается, если $f^n < \theta^n$, $n = 1, 2, \dots, N$, где θ^n — порог смещения. В качестве простой функции используются функции расчета NPD-признаков по следующей формуле (2):

$$f(i_1, i_2) = \frac{i_1 - i_2}{i_1 + i_2}, \tag{2}$$

где i_1, i_2 — интенсивности двух пикселей окна изображения, $f(i_1, i_2) = 0$, если $i_1 = i_2 = 0$.

Общая структура алгоритма детектора на основе каскадного классификатора достаточно проста. На вход алгоритма поступает изображение или кадры видеопотока не менее 30 пикселей по ширине и высоте (ограничение определяется минимальным размером обнаруживаемого лица). На первом этапе цветное изображение преобразуется в режим градации серого (grayscale), и все дальнейшие действия над изображением выполняются в этом режиме. Основной этап алгоритма состоит из функции скользящего окна с единичным шагом как по оси X, так и по оси Y изображения с разным масштабным коэффициентом, расчета NPD-функций для набора регрессионных деревьев, которые представляют собой модель детектора (табл. 1), и сравнении с пороговыми значениями для предварительного принятия решения о наличии лица в определенной области изображения. На последнем этапе алгоритма полученные области изображения с подозрением на наличие лица объединяются. Если число пересекающихся областей превысило порог, то считается, что на изображении обнаружено лицо.

Представленный алгоритм каскадного детектора на основе регрессионных деревьев эффективен для реализации

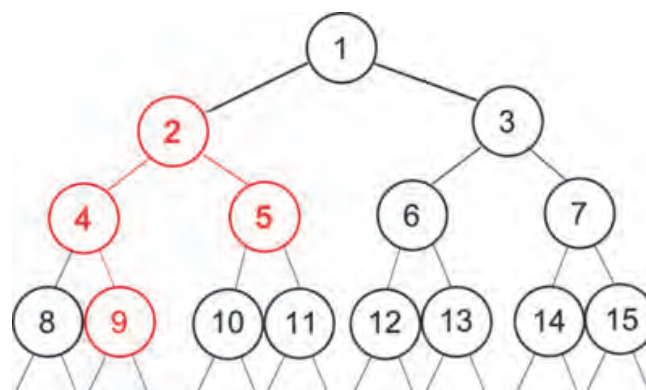


Рис. 3. Предполагаемое ускорение при использовании метода «обрезания» спектра

Таблица 1. Основные характеристики алгоритма каскадного детектора

№	Характеристика	Значение характеристики
1	Число деревьев	220
2	Глубина дерева	4
3	Объем памяти для коэффициентов модели детектора	150 КБ



Таблица 2. Результаты работы алгоритмов распознавания лиц на DSP-ядре ELcore50*

№	Алгоритм	Тип алгоритма	Число тактов, clock	Время выполнения на 1xDSP, 1 ГГц, мс	Время выполнения на 1xDSP, 1 ГГц, FPS
1	Обнаружение лица	Дерево решений	36e5	3,6	278
2	Подтверждение обнаруженного лица	Нейронная сеть	2,4e5	0,02	33210
3	Определение антропометрических точек	Дерево решений	90e5	9	111
4	Сопровождение лица	Комплексный алгоритм	130e5	13	76
5	Формирование характеристик для идентификации и верификации	Нейронная сеть	47,4e5	4,74	168
6	Анализ эмоций	Нейронная сеть	1,9e5	0,20	4232
7	Анализ возраста	Нейронная сеть	3,99e5	0,40	2001
8	Анализ пола	Нейронная сеть	1,98e5	0,20	4022
7	Итого:			31,16	32

* размер кадра составляет 256×384 пикселей.

на специальных вычислительных устройствах, которыми являются DSP-ядра, так как обладает рядом достоинств: отсутствие масштабирования самого изображения, операции выполняются над исходным изображением; использование grayscale-режима, который позволяет значительно сократить объем памяти (кеш или внутренней памяти) при выполнении вычислительных операций; использовании базовых арифметических операций для расчета NPD-признаков; применение однотипных вычислений для неглубоких регрессионных деревьев (рис. 3), которые легко векторизуются.

Для решения задач подтверждения наличия лица на изображении, выявления характеристик самого лица в целях его идентификации и верификации, а также для определения пола, возраста и эмоций на текущем изображении в настоящее время применяются алгоритмы глубокого обучения на основе сверточных нейронных сетей (convolutional neural network, CNN), которые реализованы на DSP-ядрах ELcore.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ И ВЫВОДЫ

Результаты работы алгоритмов обнаружения и распознавания лиц с определением их характеристик представлены в табл. 2, которая показывает высокую эффективность реализации алгоритмов машинного и глубокого обучения на DSP-ядрах ELcore50. Общий тракт обработки изображений в целях обнаружения лиц и определения характеристик одного лица составляет не более 32 мс, что соответствует требованиям по обработке видеок кадров в реальном времени.

Авторы считают, что в данной работе новыми являются следующие положения и результаты: модель обработки

изображения в целях обнаружения и распознавания лиц, реализация и результаты оптимизации модели на DSP-ядрах ELcore, численные результаты работы алгоритмов.

ЛИТЕРАТУРА

1. Shaogang Gong, Tao Xiang. *Visual Analysis of Behaviour: From Pixels to Semantics*. Publisher: Springer-Verlag London, 2011. 358 p.
2. Guo R. and Qi H. *Facial feature parsing and landmark detection via low-rank matrix decomposition*. In IEEE International Conference on Image Processing, 2015. P. 3773–3777.
3. AbdAlmageed W., Wua Y., Rawls S. *Face Recognition Using Deep Multi-Pose Representations*, 2016 IEEE Winter Conference on Applications of Computer Vision (WACV) (2016), Lake Placid, NY, USA. P. 1–9.
4. Martinez B., Valstar M. F., Jiang B., Pantic M. *Automatic Analysis of Facial Actions: A Survey*. In: IEEE transactions on affective computing, 25.07.2017. P. 1–22.
5. Официальный сайт НПЦ «ЭЛВИС» <http://multicore.ru>.
6. 3DIVI Face Recognition SDK. Электронный ресурс <http://3divi.com/face-recognition>.
7. Face Recognition Vendor Test (FRVT) Ongoing. <https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt-ongoing>.
8. Dong Chen, Shaoqing Ren, Yichen Wei, Xudong Cao and Jian Sun. *Joint Cascade Face Detection and Alignment*, 2014.
9. Shengcai Liao, Member, IEEE, Jain Anil K., Fellow, IEEE, and Stan Z. Li, Fellow, IEEE. *A Fast and Accurate Unconstrained Face Detector*, 2014.

