

Построение классификаторов аналогичных каскаду Виолы- Джонса с использованием признаков Хаара и искусственных нейронных сетей

Стадник А.В.

"Международный университет природы,
общества и человека "Дубна"

Структура доклада

- Актуальность задачи (1 слайд)
- Каскад Виолы-Джонса, устройство и принцип работы (5 слайдов)
- Фиксированный Хаар-базис и ИНС, описание и результаты (5 слайдов)
- Собственный Хаар-базис и ИНС, описание и результаты (3 слайда)
- Каскад ИНС (2 слайда)
- Сравнение, Выводы, Заключение (2 слайда)

Актуальность задачи

- Computer Vision - развивающаяся область CS, большой потенциал для автоматизации процессов
- Важное значение для Computer Vision
- Детектор объектов – бинарный классификатор (объект – не объект)
- Требование – $\text{false positive} \ll 1$
- Типичное использование – классификатор для сканирующего окна

Принцип работы каскада Виолы-Джонса

- Простые прямоугольные функции, называемые функциями Хаара.
- Интегральное изображение для быстрого обнаружения функции.
- Метод машинного обучения AdaBoost.
- Каскадный классификатор для эффективного совмещения множественных функций.

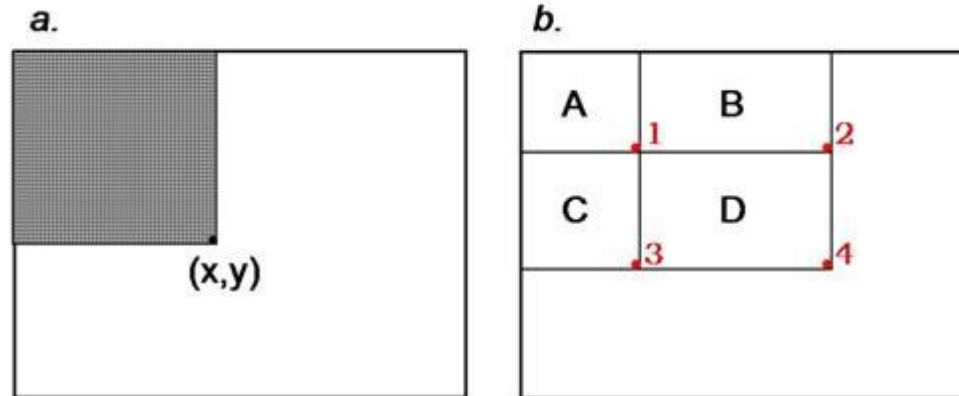
Признаки Хаара (Haar features).

- Численная характеристика изображения, соответствующая конкретному признаку Хаара, определяется посредством вычитания суммарного значения области темных пикселей из суммарного значения области светлых пикселей.



Интегральные изображения

- *Интегральное изображение* – это двумерный массив в ячейках которого хранится сумма пикселей выше и левее текущего, включая текущий.

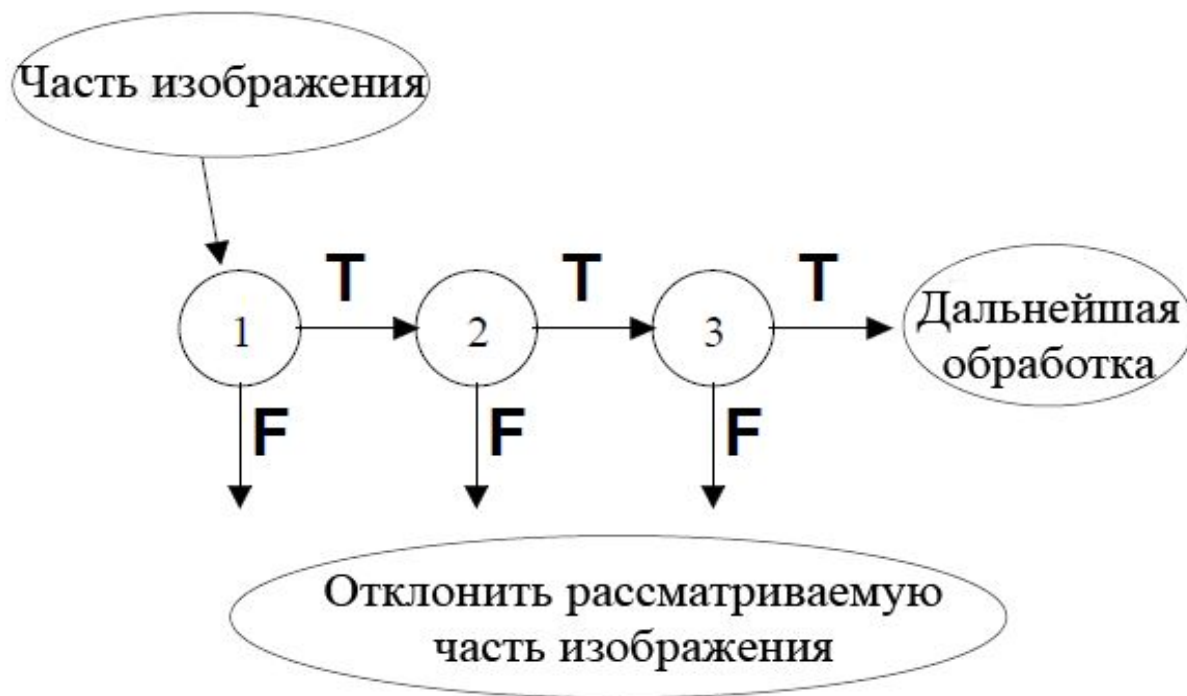


- $A+B+C+D$ является значением Интегрального Изображения в положении 4, $A+B$ есть значение в положении 2, $A+C$ – значение в положении 3, и A – значение в положении 1.
- Сумму для любого прямоугольника в изображении - три целочисленные операции: $(x_4, y_4) - (x_2, y_2) - (x_3, y_3) + (x_1, y_1)$.

Каскад отбраковки (rejection cascade)

Виола и Джонс объединили серии классификаторов AdaBoost как последовательность фильтров, что особенно эффективно для классификации областей изображения.

Каждый фильтр является отдельным классификатором AdaBoost с достаточно небольшим числом слабых классификаторов.



AdaBoost – подбор оптимальных Хаар признаков

- AdaBoost - выбор конкретных используемых функций Хаара и установления пороговых уровней.
- комбинирует много «слабых» классификаторов с целью создания одного «сильного» классификатора.
- «Слабый» классификатор - правильный ответ «ненамного» чаще, чем случайное угадывание.
- Взвешенная комбинация является сильным классификатором.

Фиксированный Хаар-базис и ИНС

Классификаторы, используемые как части каскада Виолы-Джонса достаточно просты - линейное отображение с решающим порогом

Для построения детектора были выбраны два следующих принципа:

- фиксированный базис Хаара, дающий фиксированный вектор признаков в качестве дескриптора входного изображения;
- использование многослойного персептрона в качестве "сильного" классификатора.

Фиксированный Хаар-базис

При выборе фиксированного базиса принимались во внимание следующие факторы:

- соответствие визуальному восприятию изображения,
- доступность для быстрого вычисления через интегральное изображение,
- отсутствие достаточно мелких деталей, несущих высокочастотную информацию не существенную для определения типа объекта.
- удовлетворяющий этим требованиям базис из 112 признаков Хаара



ИНС в качестве классификатора

- Выбранный базис построения дескриптора, совместно с искусственной нейронной сетью (ИНС) конфигурации 112-16-2 (112 входов, 12 нейронов скрытого слоя, 2 нейрона выходного слоя), позволили получить детектор лиц с характеристиками *detection rate* ~ 0.22 и *false positive rate* $\sim 0.3 \cdot 10^{-5}$.
- *detection rate* / *false positive rate*

База данных для обучения

- Для экспериментов в данной работе использовалась база данных изображений CMU Face Database
- <http://vasc.rh.cmu.edu/idb/html/face/>
- 20x20 размер изображения
- Обучающая выборка: 2429 положительных, 4548 отрицательных образцов
- Тестовая выборка: 472 и 23573 соответственно

Результаты

- Оптимизация детектора может быть проведена путем смещения порога срабатывания нейронной сети, который по умолчанию принят за ноль, в ту или иную сторону.



Собственный Хаар-базис

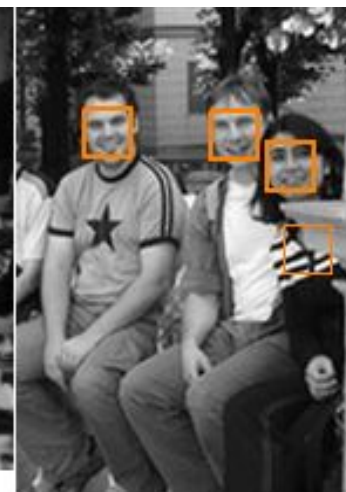
- Очевидный недостаток предыдущего подхода – неоптимальный набор признаков
- Вариант построения адекватного задаче набора Хаар-признаков – собственные вектора, соответствующие максимальным собственным значениям (МГК, РСА)

Собственный Хаар-базис

- Меньшее количество признаков
- Точнее отражают характерные особенности объектов
- 48 собственных векторов
- собственные вектора квантованы, обнулены пиксели, не превышающие среднее значение на $1/2$ среднеквадратичного отклонения.

Собственный Хаар-базис, результаты

- Вероятность детекции 0.27 (detection rate)
- Ошибка второго рода $\sim 10^{(-6)}$

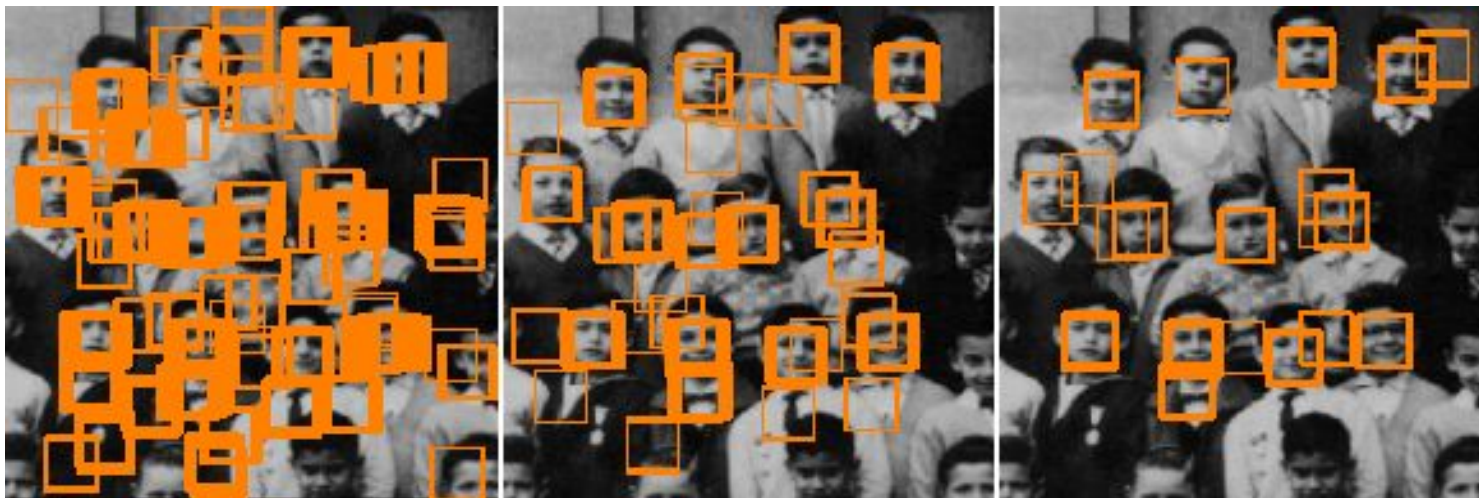


Детектор на основе каскада нейронных сетей

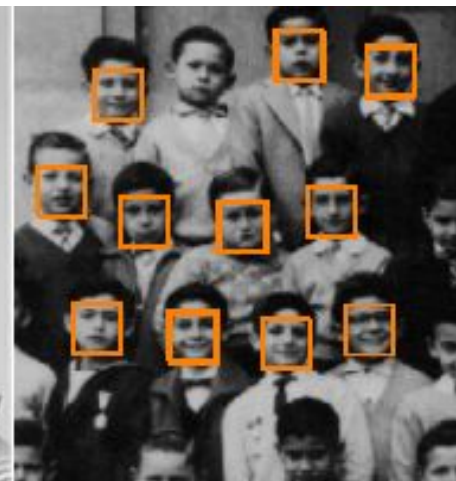
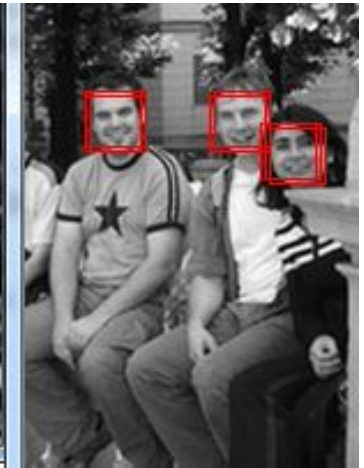
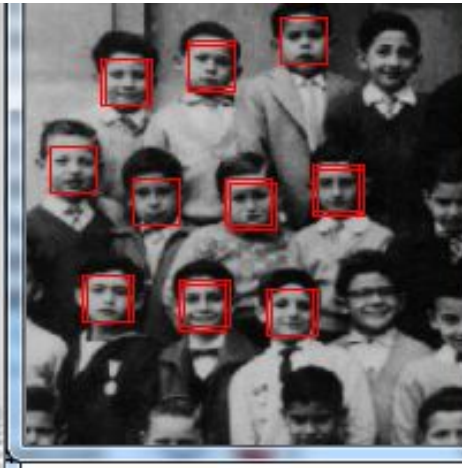
- Аналог rejection каскада – из ИНС разной сложности
- 3 ИНС {8-16-1} {24-16-1} {48-16-1}
- Пороги ИНС₁ и ИНС₂ необходимо занизить
- Из 29929 подокон первой нейросетью каскада было пропущено 2520 (~9%), второй нейросетью - 458 из них, а третьей нейросетью - 46 сработавших сэмплов
- {8-16-1}: вычислительная сложность $(8+1)*16+(16+1)*1 = 161 \text{ float} * + 16 \text{ tanh}$

Детектор на основе каскада нейронных сетей

- Визуальная демонстрация работы сетей из каскада отдельно с заниженным порогом детекции.




Сравнение с каскадом Виолы-Джонса



Направления оптимизации

- Аппроксимация набора собственных признаков – уменьшение числа вычисляемых прямоугольников.
- Вычисление изображений соответствующих каждому из признаков отдельно. Уход от сканирования подокон, вычисление общего изображения результата свёртки. (теорема о свертке, $N \cdot \log(N)$).
- Аналогично для слоя нейронов.
- Использование SIMD.
- Переход от float к fixed.
- Реализация на DSP.



Спасибо за
проявленное
внимание к докладу !