

Технология распознавания жестов в видеопотоке с использованием методов машинного обучения на примере азбуки глухонемых

Землянская С.А.

СПбГУ, Математико-механический факультет

Научный руководитель: д.ф.-м.н., проф. Граничин О.Н.

Рецензент: Кривоконь Д.С.

Санкт-Петербург, 2013



Существующие решения

Работы по данной теме можно условно разделить на несколько групп:

- Использующие карту глубины,
- Применяющие маркерные перчатки,
- Производящие локализацию ладони по цвету кожи,
- Распознающие динамические жесты, как последовательность сменяемых состояний.

Входные данные

На вход подаётся видео с одной камеры:

- Видео содержит последовательную смену жестов
- Фон не обязательно однородный, может содержать неактивное или периодичное движение

Выходные данные

Видео разделяется на отдельные жесты, каждый из которых классифицируется в рамках выбранного алфавита.

Контрольный набор

В качестве контрольного набора жестов было выбрано подмножество из международного алфавита для глухонемых.



Рис.: Международный язык жестов (a-p)



Рис.: Международный язык жестов (r-z)

- 1 Предварительная обработка видео
 - Представление видео в виде последовательности изображений
 - Вычитание неподвижного фона
- 2 Разделение видео на жесты
 - Отслеживание "активности" движения на видео
 - Нахождение моментов смены жеста
- 3 Получение "признаков" жеста и классификация



Рис.: Отслеживание движения на видео

- 1 Предварительная обработка видео
 - Представление видео в виде последовательности изображений
 - Вычитание неподвижного фона
- 2 Разделение видео на жесты
 - Отслеживание "активности" движения на видео
 - Нахождение моментов смены жеста
- 3 Получение "признаков" жеста и классификация

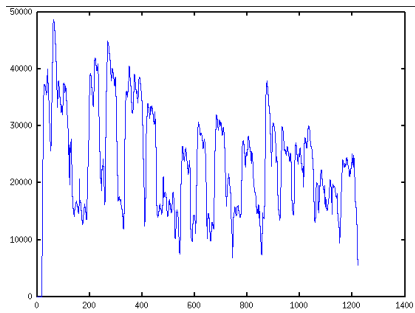
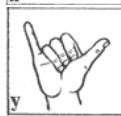
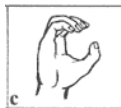
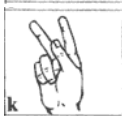


Рис.: Изменение активности движения на видео

- 1 Предварительная обработка видео
 - Представление видео в виде последовательности изображений
 - Вычитание неподвижного фона
- 2 Разделение видео на жесты
 - Отслеживание "активности" движения на видео
 - Нахождение моментов смены жеста
- 3 Получение "признаков" жеста и классификация

Контрольный набор

В качестве контрольного набора жестов было выбрано подмножество из международного алфавита для глухонемых



Признаки жеста:

- Гистограмма направленных градиентов
- Габариты ладони + Моменты контура

Классификатор:

Метод опорных векторов

Классификация жеста

Зафиксированный жест классифицируется как принадлежащий контрольному множеству, если более 60% кадров с ним классифицированы одним образом и не принадлежащим контрольному множеству в противном случае.

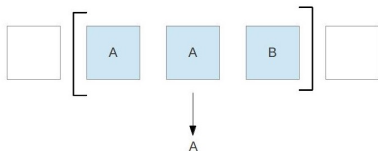


Рис.: Распознавание жеста

Реализация

Приложение было реализовано на языке Octave

Обучение производилось:

- При различном освещении
- Менялся фон на заднем плане
- Обучение производилось под конкретного человека

Метрики

Для оценки классификации использовались метрики:

- $precision = \frac{tp}{tp+fp}$ - точность,
- $recall = \frac{tp}{tp+fn}$ - полнота,
- $percent$ - процент верно классифицированных жестов.

Метрика /Признаки	Моменты контура	HOG
Precision	0.90643	0.96685
Recall	0.91716	0.97765
Percent	82.7%	93.5%

Таблица: Значения метрик

Итоги

В ходе проделанной работы:

- Произведён анализ существующих систем по распознаванию жестов.
- Разработан и реализован алгоритм для распознавания статических жестов по видео с одной камеры.
- Произведено обучение алгоритма для множества из 6-ти жестов азбуки глухонемых.