

МЕТОД ОПРЕДЕЛЕНИЯ СХОДСТВА ЖЕСТОВ НА ОСНОВЕ СРАВНЕНИЯ ВИДЕО

Серов Сергей Сергеевич

Студент

Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: sergey.s.serov@gmail.com

Научный руководитель — *Местецкий Леонид Моисеевич*

Задача сравнения жестов на видео является актуальной задачей обработки изображений, так как её решение необходимо для построения алгоритма распознавания жестов. В отличие от других работ, в данном исследовании предпринята попытка построить универсальную метрику сравнения жестов по видео, непосредственно использующую данные о перемещении объектов в кадре и не опирающуюся на семантику конкретных жестов [2].

Основными признаками, различающими жесты между собой, являются траектория движения и динамика формы кистей рук и лица. Целью данного исследования является разработка метрики сравнения жестов на видео, учитывающей эти ключевые факторы, но универсальной для разных жестов. В качестве экспериментальной базы исследования используется один из индийских языков жестов.

Идея предлагаемого метода состоит в декомпозиции задачи на две подзадачи: обработка кадра и межкадровая обработка.

На этапе обработки кадра необходимо произвести сегментацию лица и кистей рук, а также определить их местоположение. Сегментация может быть проведена с помощью выделения на кадре соответствующих пятен. В случае записи трехканального RGB-видео для этого можно использовать бинаризацию, а в случае записи видео с данными о глубине изображения RGB-d — карту глубины.

Для регистрации траекторий движения и формы пятен будем выполнять построение *морфологических скелетов* пятен [1]. Каждое пятно может быть представлено в виде графа, для которого известны координаты x, y вершин, рёбра и радиусы r максимальных кругов, соответствующих вершинам.

На этапе межкадровой обработки, если известны местоположения пятен кистей рук и лица на отдельных кадрах, становится возможным вычисление траекторий этих пятен на видео (рис. 1). Каждая траектория в таком случае представляется в виде множества упорядоченных троек вида (x_i, y_i, r_i) . На основе этой информации сформируем признаковое описание каждого жеста.

После регистрации траекторий движения пятен для сравнения признаков описаний жестов предлагается следующая метрика. Для каждой пары соответствующих пятен двух жестов будем выполнять попарное *выравнивание* их траекторий с помощью метода *динамического программирования*. В качестве расстояния между пятнами выберем сумму евклидовых расстояний между точками выровненных траекторий. Расстоянием между жестами будем считать сумму расстояний между пятнами кистей рук.

Построенная метрика сравнения признаков описаний жестов теперь позволяет провести их классификацию с использованием базы эталонов. Разработка метода распознавания жестов на видео, а также регистрация и учёт положения пальцев на основе построенных морфологических скелетов продолжают исследования в этом направлении.

Исследование поддержано грантом РФФИ 20-01-00664.

Иллюстрации



Рис. 1. Пример зарегистрированных на видео траекторий и окружностей максимальных кругов пятен лица и кистей рук

Литература

1. Местецкий Л. М. Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры. М.: Физматлит, 2009
2. Bandini A., Zariffa J. Analysis of the hands in egocentric vision: A survey. 2019. ArXiv preprint arXiv:1912.10867