

## Секция «Вычислительная математика и кибернетика»

### Автоматическое распознавание статичных жестов с помощью Kinect на примере дактильной азбуки

*Артюхин Станислав Геннадьевич*

*Студент*

*Московский государственный университет имени М.В. Ломоносова, Факультет вычислительной математики и кибернетики, Москва, Россия*

*E-mail: stas.artuhin@gmail.com*

Жест (от лат. *gestus* — движение тела) — некоторое действие или движение человеческого тела или его части, имеющее определённое значение или смысл, то есть являющееся знаком или символом. В последнее время все больше внимания уделяется автоматическому распознаванию жестов. Это обусловлено тем, что жесты - удобный способ ввода информации в компьютер. Наличие почти в каждом компьютере веб-камеры приемлемого качества, а также вычислительная мощность современных персональных систем позволяет решать задачи обработки изображения в реальном времени.

Важной задачей распознавания жестов является задача распознавания жестов языка глухонемых. В языках глухонемых большинство жестов динамичные. Более простым языком для распознавания является дактилирование букв и цифр. В этом языке для каждой буквы и цифры есть свой жест, как правило, статичный.

Рассмотрена задача распознавания жестов на основе данных, полученных с веб-камеры и камеры глубины. Жест является последовательностью пары изображений

$$G = ((I_1, D_1), (I_2, D_2), \dots, (I_{|G|}, D_{|G|})).$$

Статичный жест задается одной парой изображений. Задача формулируется следующим образом. Пусть дана база статичных жестов  $G_i = (I_i, D_i)$ . Требуется по данному видеоряду

$$V = ((I_1, D_1), (I_2, D_2), \dots, (I_k, D_k))$$

определить

- количество жестов присутствующих в видеоряде
- для каждого найденного жеста определить начальный и конечный кадр видеопоследовательности

Для решения задачи распознавания жестов необходимо научиться выделять руку человека. Были рассмотрены несколько методов автоматического выделения руки. В данный момент в работе задача решена двумя способами: с помощью черной перчатки, которую можно эффективно выделять с высокой точностью, и с помощью камеры глубины Kinect и цветовой модели HSV.

В данной работе используется выделение геометрических признаков. По двумерной проекции руки строится ее медиальное представление. Затем осуществляется генерация различных признаков скелета. Полученные признаки используются для нахождения ближайшего эталона. Хранение эталонов в виде признаков, а также устойчивость признаков позволяет уменьшить объем требуемой памяти и вычислительных затрат.

### **Литература**

1. Местецкий Л.М. Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры, Москва: Физматлит. 2009
2. Gudmundsson S. A., Sveinsson J. R., Pard'as M. et al. Model-Based Hand Gesture Tracking in ToF Image Sequence, 6th International Conference on Articulated motion and deformable objects (AMDO). 2010.
3. Wang R. Y., Popovic J. Real-time hand-tracking with a color glove, ACM Transactions on Graphics. 2009. Vol. 28, no. 3.
4. Dhawale P., Masoodian M., Rogers B. Bare-hand 3D gesture input to interactive systems, CHINZ'06: Proceedings of the 7th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: design centered HCI. New York, NY, USA: ACM, 2006. P. 25–32.
5. Van den Bergh M., Van Gool L. Combining RGB and ToF cameras for real-time 3D hand gesture interaction, IEEE Workshop on Applications of Computer Vision (WACV). 2011