

Интерфейс на основе жестов для взаимодействия с виртуальными средами
А.В. Зырянов, В.Л. Авербух
Уральский государственный университет, ИММ УрО РАН Екатеринбург

В настоящее время для взаимодействия с компьютером используются, в основном, два вида интерфейса: командная строка и графический интерфейс. И тот и другой являются, по сути, искусственными языками, а значит, для работы с компьютером пользователь должен овладеть новым для него языком. Причем использование таких языков в виртуальных средах затруднительно. Решить эту проблему можно путем обучения компьютера пониманию естественных языков и, в частности, языка жестов.

Основными преимуществами языка жестов по сравнению с другими естественными языками является:

1. универсальность (жест действия, например, вырезания объекта при помощи виртуального ножа, не зависит от национальной культуры человека);
2. возможность проецировать повседневный опыт на виртуальные среды (скажем, если человек возьмет виртуальный нож, то он будет точно знать, как нужно им пользоваться и каких результатов следует ожидать).

Для захвата жестов я разработал технологию, основанную на единственной веб-камере и обыкновенном фонарике, который пользователь держит в руке. Изображение, получаемое с веб-камеры, подвергается анализу: изменение положения светового пятна позволяет определить две координаты, на основе изменений размера светового пятна определяется третья, а путем анализа геометрической формы светового пятна вычисляются углы наклона фонарика. Таким образом, система может определять не только трехмерное положение фонарика, но и направление, в котором он смотрит, что позволяет использовать фонарик в качестве трехмерной мыши и трехмерного указателя. Преимуществом данной системы является простота использования (нет нужды расставлять множество камер и прикреплять специальные маркеры к руке человека), а также низкая стоимость.

Для распознавания жестов производится анализ изменения координат фонарика во времени. Получаемая многомерная траектория сравнивается с набором базовых траекторий, каждая из которых связана с определённым жестом. При обнаружении соответствия считается, что пользователь ввёл соответствующий жест. При этом отсутствует необходимость как-то обозначать начало и конец жеста, а также делать паузу между жестами (т.е. система способна выделить в пользовательском вводе несколько подряд идущих жестов). Также стоит отметить, что алгоритм анализа достаточно производителен и может работать в реальном времени.

В настоящее время создан прототипный вариант описываемой системы. Работа над проектом продолжается.

The Gestures-Based interface for Interacting with Virtual Environments
Alexander Zyryanov, Vladimir Averbukh
Ural State University, IMM UrB RAS Yekaterinburg

Currently two types of interface are using for human-computer interaction: the command line and the graphical interface. Both are, in fact, artificial languages, so user must learn a completely new language to interact with computer. And these artificial languages are not suitable for virtual

environments. To solve this problem, we should teach computer to understand human language. In particular, gesture language.

The main advantages of gesture language over other natural languages are:

1. universality (action-gesture, like cutting with a knife, does not depend on the human's national culture);
2. possibility of projecting the daily experience into virtual environments (say, if a person takes a virtual knife, he would know exactly how to use it and what to expect in results).

For capturing gestures I developed a technology based on a single web camera and an ordinary flashlight, which the user holds in his hand. The program analyse web camera image: from position change of light spots the first two coordinates are determined, from the changes of the light spot's size we determine the third, and from geometric shape analysis program calculates the angles of the flashlight. Thus, the system may determine not only the three-dimensional position of the flashlight, but also the direction in which it looks. So flashlight can be used as a three-dimensional mouse and as a three-dimensional pointer. The advantage of this technology is its simplicity (no need to calibrate many cameras and to attach special markers to the user's hand) as well as low cost.

For gesture recognition, coordinates change over time is analysed. The program compare resultant multi-dimensional trajectory with a set of basic trajectories, each of which is linked with a certain gesture. If we find correspondence, we decide that the user has entered the corresponding gesture. It should be mentioned that there is no need to designate the beginning and end of gesture as well as make a pause between gestures (the system is able to recognise several gestures in continuous user input). Also the analysis algorithm is fast enough to work in real time. The prototype variant of described system is realized now, but the system is still under development.