



159159

Methods and devices for inputting data to a computer system

Field of the invention

The present invention relates to methods and devices for inputting data to a computer system using a pointing device, such as a stick or even a user's
5 finger or arm.



G00002

Background of Invention

Various visual methods have been proposed to allow data to be input to a computer by a user making a pointing action using a special pointing device, such as a device which generates a laser beam. On certain occasions, such
10 as when making a presentation to a large audience using a display controlled by the computer, these methods enable more convenient inputting of data to the computer than by using a mouse, a joy-stick or a key-board in certain circumstances. Examples include US patent applications 20010030668 and 20010022575.

15 In such existing systems the pointing device emits a directed electromagnetic beam into the display area and optical devices and visual methods are used to observe the intersection of the beam and the display area. The location of this intersection point is used to obtain input data to the computer, for example to trigger operations of the computer.

20 Since in these arrangements a special electronic pointing device is required, the operation is different from the ways in which people naturally indicate a location: by pointing to that location using a stick, a finger or an arm.

Also, all the known methods require the beam from the electronic device to physically intersect the display area and this sometimes causes discomfort for
25 the audience. For example, some people do not like the glinting laser beam-spot moving on the display screen.

Summary of the invention

The present invention aims to provide new and useful methods and systems for inputting data to a computer system.

- 5 In general terms, the present invention proposes that a user employs a pointing device having an axis to point at a display. The pointing device may be a simple stick, or even a user's finger or arm. Images of the pointing device are taken by at least two cameras. Each of these images contains a line representing the axis in the two-dimensional image plane of the
- 10 corresponding camera. The intersection between these lines in a common two dimensional co-ordinate plane is determined, corresponding to the position in that common co-ordinate frame of the point of intersection of the axis of the pointing device with the display. The co-ordinates of the point of intersection of the lines in the common two-dimensional co-ordinate frame may thus be
- 15 transformed by a predetermined co-ordinate transformation into the plane of the display, and/or into a different co-ordinate plane determined by the computer, to generate input data for the software of the computer.

Preferably, the common two-dimensional co-ordinate frame is the image plane of one of the cameras.

20 Brief Description of The Figures

Preferred features of the invention will now be described, for the sake of illustration only, with reference to the following figures in which:

Fig. 1 shows the units of a data input system which is an embodiment of the invention; and

- 25 Fig. 2 shows the steps of an algorithm performed by embodiment of Fig. 1.

Detailed Description of the embodiments

Referring firstly to Fig. 1, the overall configuration is shown of a system which is an embodiment of the invention. The system includes a PC having a
 5 computer unit 1 and a display 2 operated by the computer unit 1. Note that the display 2 may be the monitor of the PC, or alternatively may be a separate display, such as a display for a large audience.

The system further includes a pointing device 3 which may be a straight stick
 10 or a user's arm or finger, two cameras 5, 7, and a processing unit 13 which receives the outputs of the cameras 5, 7 via cables 11. The processing unit 13 generates a signal which is transmitted to the computing unit 1 of the PC, which produces a feedback on the display 2. Note that the processing unit 13 may, as shown in Fig. 1, be a separate physical unit from the PC, but
 15 alternatively it may be implemented by hardware or software within the PC computing unit 1. In either case, its function is generating input data for the computer unit 1 (i.e. input data for software running on the computer unit 1).

The pointing unit 3 is directed by a user to point its axis at a position 17 on the
 20 display 2. The processing unit 13 detects the co-ordinates of the pointing position 17 in the display area. These co-ordinates are referred to here as (x_{P-disp} , y_{P-disp}) in the co-ordinate system (x_{disp} , y_{disp}) of the display area.

Optionally, the processing unit further normally determines the co-ordinates of a corresponding position in a second co-ordinate system (x_{PC} , y_{PC}) which is
 25 used by the software of the PC. This second position is referred to here as (x_{P-PC} , y_{P-PC}), and may for example be the co-ordinates in of a corresponding point on the display area of a monitor of the PC which, like the display 2, is controlled by the computing unit 1. For example, it may be a point on an integer matrix of the PC display with predefined x and y directions, e.g. an
 30 800x600 matrix, based on which the software can direct and manipulate the

display of the PC. Note that, as compared to the PC co-ordinate system, the display co-ordinate system may have different units and a different x-y definition. Moreover, there may be “distortion” between the two sets of co-ordinate systems (e.g. a projective transformation), for example when the display is generated by a projector. Thus, generally speaking, a transformation matrix is needed to define the correspondence between these two co-ordinate systems, though it might as easily be a scaling matrix, or even an identity matrix in some cases as mentioned below.

If the embodiment had only included one camera, it would have been impossible to detect the pointing position when the stick (finger/arm) does not touch the display plane. Even using two cameras it would be very troublesome to detect the pointing position using a conventional 3D vision method which requires complex and restrictive calibration and high cost hardware to perform real-time processing to generate a 3D depth map.

Instead, the embodiment an algorithm subjects the outputs of the two cameras 5, 7 to 2D vision algorithms to determine the pointing position in an easy-to-implement way. Note that the algorithm does not require the pointing device 3 to touch the display plane physically.

It is well known in machine vision theory that there is a projective transformation between the co-ordinates in a planar surface (X,Y) and the co-ordinates in an image plane (x, y) . This transformation can be written as:

$$\begin{bmatrix} sx \\ sy \\ s \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (1)$$

Here the matrix M (m_{ij} , $i=1,2,3$ and $j=1,2,3$) can be determined by a calibration method, and “s” is a scalar used in machine vision techniques.

In the system of the present embodiment there is a similar relationship between the co-ordinate system (x_{disp}, y_{disp}) on the display surface 2 and the co-ordinate system (x_{PC}, y_{PC}) of the PC. This relationship may be expressed by the matrix $M_{PC-to-disp}$. This is because the projection system is similar to a machine vision imaging system. Note that in the case that the display 2 actually is the monitor of the computer then the two co-ordinate systems are the same, and in this case the matrix $M_{PC-to-disp}$ can be considered as an identity matrix with 1 for the diagonal elements and 0 for other elements. More generally, however:

$$\begin{bmatrix} sX_{disp} \\ sy_{disp} \\ s \end{bmatrix} = M_{PC-to-disp} \begin{bmatrix} x_{PC} \\ y_{PC} \\ 1 \end{bmatrix} \quad (2)$$

Two further co-ordinate systems of interest are the two-dimensional image planes of the left and right camera 5, 7, written here respectively as I_L and I_R . Points in these co-ordinate systems are written as (x_L, y_L) and (x_R, y_R) respectively. There are relationships between images taken by the left and right cameras given by matrices denoted as M_L and M_R respectively. These relationships are as follows:

$$\begin{bmatrix} sX_L \\ sy_L \\ s \end{bmatrix} = M_L \begin{bmatrix} x_{disp} \\ y_{disp} \\ 1 \end{bmatrix} \quad \begin{bmatrix} sX_R \\ sy_R \\ s \end{bmatrix} = M_R \begin{bmatrix} x_{disp} \\ y_{disp} \\ 1 \end{bmatrix} \quad (3)$$

This means that there is matrix M_{L-to-R} which defines a projective transformation between the left (x_L, y_L) and right image planes (x_R, y_R) , according to the formula:

$$\begin{bmatrix} sX_R \\ sy_R \\ s \end{bmatrix} = M_{L-to-R} \begin{bmatrix} x_L \\ y_L \\ 1 \end{bmatrix} \quad (4)$$

- 5 Thus for every point in the display plane, there is a projective co-ordinate transformation of the corresponding co-ordinates between the left and right image planes.

Also from machine vision theory it is known that a straight line (e.g. the axis
 10 of the pointing device 3) in 3D space will be projected to a straight line in the image plane of a camera. For our system, there will be two corresponding lines, one in each image plane. The lines (written as L_L and L_R respectively) can be detected by use of line detection (e.g. a Hough transformation) or with the aid of color segmentation (e.g. using a neural network method to detect
 15 the specific color of the pointing device 3, e.g. skin color in the case that the pointing device 3 is a finger or arm). If the pointing device 3 lay entirely in the display plane 2, there would be transformation according to Eqn. (4) for each point on the pointing device 3. But if the pointing device 3 is not within the display plane 2, there does not exist such a relationship for all points on the
 20 pointing device and this is the case we need to deal with.

Our algorithm is derived from Eqn. (4) and consists of the following 3 steps illustrated in Fig. 2.

- 25 In step 1, the line representing the pointing device 3 in a first of the image planes is transformed into the other image plane using a pre-determined matrix. For example, writing the line in the image plane of the left camera as L_L we can use Eqn. (4) to transform this line in the left image plane I_L into the right image plane I_R , to form a line in that right image plane we refer to as L_{L-to-R} , as shown in Fig. 2.
 30

In step 2, the point of intersection of the two lines in the second image plane is determined. For example, if following step 1 the right image plane I_R now contains 2 lines L_R and L_{L-to-R} , then the crossing point P of the two lines in the right image plane is found, as shown in Fig. 2. This point P corresponds to the pointing point 17. Its co-ordinates in the right image plane are written as (X_{PR}, Y_{PR}) .

In step 3, the co-ordinates of the intersection point P are transformed from the co-ordinate system of the second image plane into the co-ordinate system of the display and/or of the PC using two pre-determined matrices. For example, if the second image plane is the right image plane I_R then the co-ordinates of the intersection point in the right image plane (X_{PR}, Y_{PR}) are transformed into the corresponding co-ordinates in display plane, i.e. (X_{P-disp}, Y_{P-disp}) , using a pre-calibrated relationship determined by Eqn. (3). These co-ordinates are just the pointing position 17 of the pointing device 3 in the display plane 2. Finally these co-ordinates are optionally further translated to the PC co-ordinate system using the pre-calibrated matrix $M_{disp-PC}$ of Eqn. (2).

The reasoning is as follows: due to the manipulation of step 1 above, this final point should be applicable to transformation described by Eqn. (4), thus the corresponding point in the 3D space should lie in the display plane. Also, this point must be on the axis of the pointing device 3 in 3D space, since it is on the line in one of the image planes (e.g. it is on L_R). So this point is exactly the intersection point of the axis of the pointing device 3 and the display plane 2, i.e. the pointing position 17 on the display plane 2.

The pointing position in the PC co-ordinate system can be used to control software, just as a mouse controls software by indicating a point in the PC co-ordinate system. No modification of the software is required, since the input

data to the software is a point on the display area of the monitor, just as if that point had been selected by a mouse. The software may be designed to add an overlapped indicator in the PC display memory, to be displayed together with the existing content. Thus the display 2 is modified.

5

Furthermore, the software may be configured to permit the user to control other PC operations by moving the pointing position 17 in pre-selected ways. For example, the software may be such that the user can make a circular movement of the pointing device 3 to move the pointing position 17 in a circle and thereby select a corresponding part of the content in the display 2.

10

Alternatively, the user may keep the pointing position 17 constant for a predetermined time to highlight the corresponding content. Alternatively the user may point to a predetermined location inside or outside the display 2 to activate a menu.

15

Although only a single embodiment of the invention has been described above, many variations are possible within the scope of the invention as will be clear to an expert.

Claims

1. A method of inputting data to a computer including:
 - controlling the position of a pointing device having an axis so as to direct the axis towards a display controlled by the computer;
- 5 generating using at least two cameras respective two-dimensional images including images of the pointing device;
 - deriving from the images corresponding lines in a common two-dimensional co-ordinate system, and determining a point of intersection of the lines in the common two-dimensional co-ordinate system; and
- 10 obtaining input data for the computer from the point of intersection of the lines.
2. A method according to claim 1 in which the computer varies the image displayed in the display area based on the input data.
3. A method according to claim 1 or claim 2 in which the common two-

15 dimensional co-ordinate system is the image plane of one of the cameras.
4. A method according to claim 1, claim 2 or claim 3 in which the step of obtaining input data includes performing a predetermined transformation of the point of intersection of the lines in the common two-dimensional co-ordinate system to determine the point of intersection of the axis of the

20 pointing device with the display in the co-ordinate system of the display.
5. A method according to any preceding claim in which the step of obtaining input data includes performing a predetermined transformation of the point of intersection of the lines in the common two-dimensional co-ordinate system to determine a point in a further two-dimensional co-ordinate

system, the input data being obtained from the co-ordinate values in the further two-dimensional co-ordinate system.

6. A method according to any preceding claim in which the pointing device is a finger or arm of the user.

5 7. A method according to any of claims 1 to 5 in which the pointing device is a straight indicating stick.

8. A data input system for a computer comprising:

a display controlled by a computer to display images in a display area;

10 at least two cameras for registering respective two-dimensional images of a pointing device defining an axis;

a processing unit for deriving from the image corresponding lines in a common two-dimensional co-ordinate system, and determining a point of intersection of the lines.

15 9. A data input system according to claim 8 in which the common two-dimensional co-ordinate system is the image plane of one of the cameras.

20 10. A data input system according to claim 8 or claim 9 in which the processing unit performs a predetermined transformation of the point of intersection of the lines in the common two-dimensional co-ordinate system to determine the point of intersection of the axis of the pointing device with the display in the co-ordinate system of the display.

11. A data input device according to claim 8, claim 9 or claim 10 in which the processing unit performs a predetermined transformation of the point of intersection of the lines in the common two-dimensional co-ordinate system to determine a point in a further two-dimensional co-ordinate system, the

computing unit performing said variation of the images using the co-ordinate values in the further two-dimensional co-ordinate system.

12. A display system comprising a data input system according to any of claims 8 to 11 and a computer arranged to receive the input data, the
5 computer unit being arranged to vary the images displayed in the display area based on the input data.