7 Controls

Konrad Baumann

For mechanical or traditional electronic products, the choice and design of controls is determined mainly by technical constraints. For this reason, the function of many traditional appliances could be worked out from their controls. Thus, every telephone had a dial, every audio-amplifier had a control knob operating a potentiometer for volume adjustment, etc. In this way, the functions of the appliances became cultural standards or stereotypes (i.e. part of the general education of the users). Hence the controls were self-explanatory to a high degree, and operating them required relatively little time and attention, even if the degree of automation was low.

7.1 Is the diversity of controls lost?

Since the introduction of processor-controlled user interfaces (UIs), the diversity of controls is often lost to be replaced by a uniform set of keys. The predominance of keys in such appliances primarily has cost reasons; but the wish for visible modernity or aesthetics also plays a role, besides the requirement of low production costs. That is why aesthetical and economical arguments (besides the technical ones) support the trend toward a uniform design of controls as keys.

But while the self-descriptiveness decreases, the number of functions of most appliances increases steadily. The advantages gained through more automation and new functions are frequently neutralized by complex UIs. As a consequence, most of the functions gained by the processor control remain unused. But they are an important incentive for buying the appliance.

However, as soon as the novelty of processors loses its appeal for the customers, the design engineers will have to remember the existing diversity of controls and their respective ranges of applicability. In combination with present-day high-resolution displays, a well-programmed processor and a good industrial design, appliances having a highly visible ergonomic quality can be built. The latter will become an increasingly important buying incentive.

Guideline
- *Use a wide variety of existing control forms to make the UI unmistakable.*

7.2 Control elements for electronic devices

In this chapter the control elements relevant for electronic devices are described and evaluated in detail (Figure 7.1). They will be classified according to the mode of function as recognizable by the user. For that reason devices based on different physical principles may be treated under the same heading. This chapter presents an extensive survey of usable controls. It may serve as a construction set and decision aid in the development of new UIs.

7.3 One-shot key (with key click)

A key is a control which is operated by pressing with a finger and closes a contact as long as the key is pressed. A key with downward displacement or travel may be pressed down and automatically returns to its resting position after release. When operated, the key must have a noticeable click point and produce an acoustic signal (tactile and acoustic feedback) (Figures 7.2–7.4).

Curves A and B in Figure 7.5 represent favourable, C and D unfavourable force functions. The maximum of A and B should be situated between 250 g and 500 g. The steep decrease behind the maximum is essential. The diameter or the side length of a key should be between 12 mm and 20 mm, the distance between the centres of adjacent keys between 18 mm and 20 mm. The operation distance should be between 1.3 mm and 6.4 mm, the optimal value depending on the operation force. The surface of the key should be concave. Closely spaced keys should be separated by bridges (Burandt 1986; Woodson 1987; Gilmore 1989; Grandjean 1993).

7.4 One-shot key (push-button)

Keys without downward displacement or travel do not obey all ergonomic requirements. They are frequently used because they are cheaper, flatter or more attractive and resist environmental influences, but they are poor from the ergonomic point of view.

Because of the missing or inadequate feedback they give rise to more misoperations than keys with a key-click. They demand more attention by the user, because he/she must pay attention to confirmation of the operation.

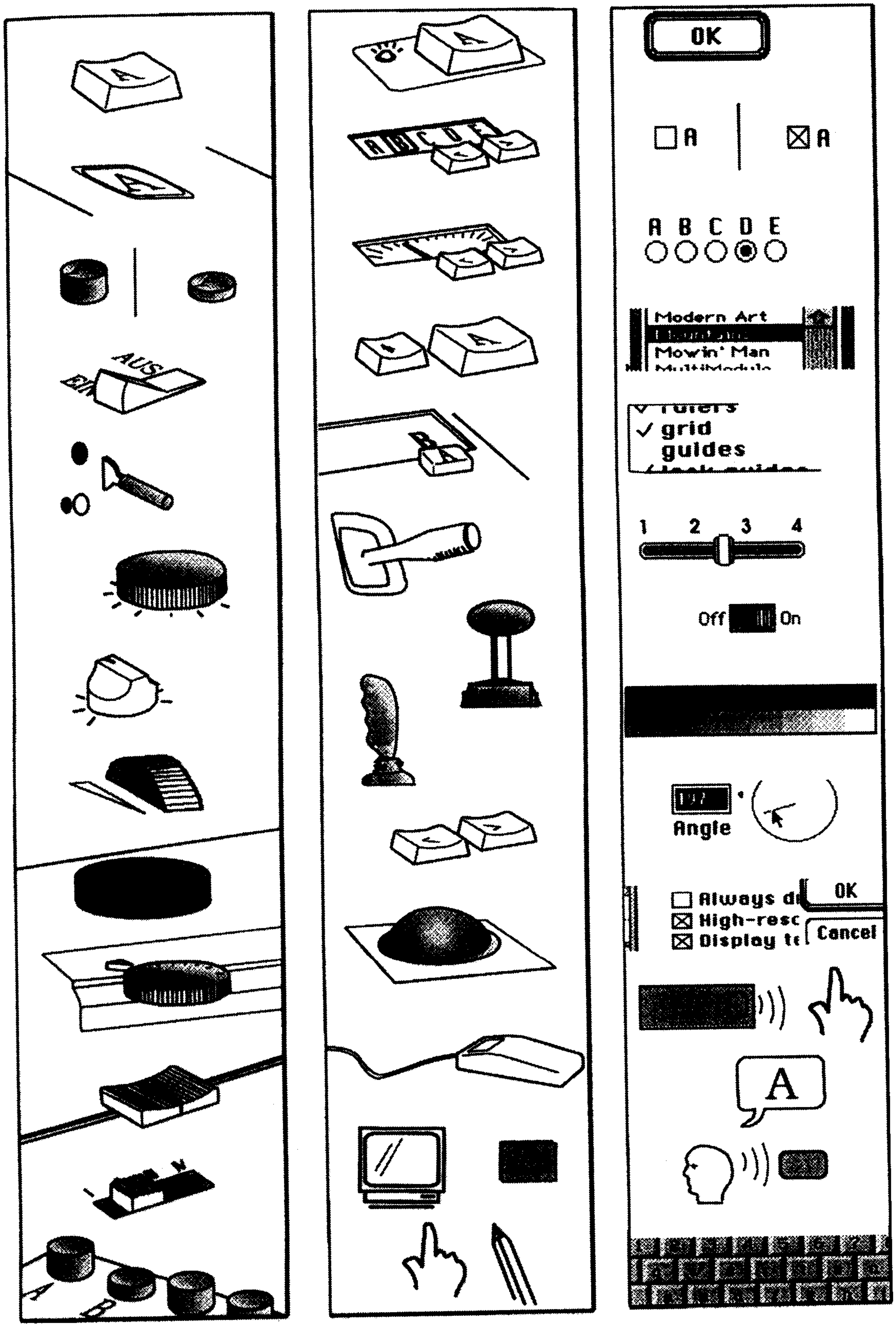


Figure 7.1 Overview of types of control: (left) standalone controls; (centre) controls to be used together with a display; (right) virtual and other controls.

Figure 7.2 Key with key-click.

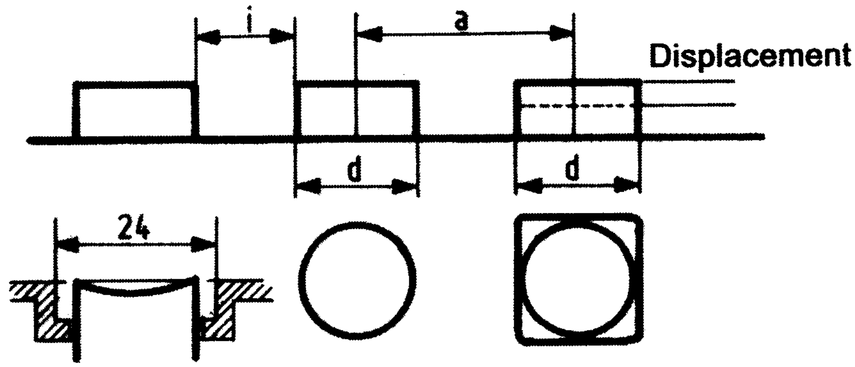


Figure 7.3 Key geometry (Burandt 1986).

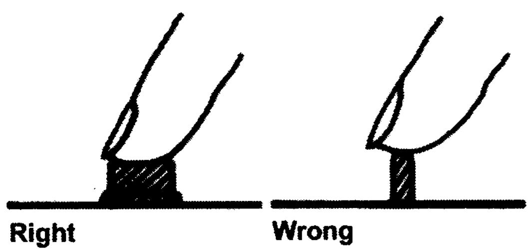


Figure 7.4 Recommended key size (left); and not recommended key size (right) (Grandjean 1991).

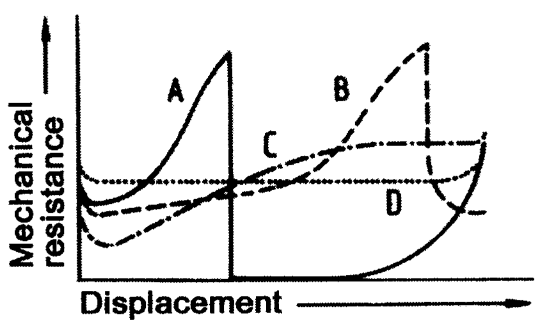


Figure 7.5 Force for pushing a key as a function of the downward displacement or travel (Burandt 1986).

As a compensation for the missing click point a beep is frequently used. In an office environment, this may be disturbing.

Foil keys (Figure 7.6) are operated with more power than keys with operation distance, and they may cause a numb feeling at the fingertips (Woodson 1987). Conducting contact sensors cannot be activated with

Figure 7.6 Foil key.

Figure 7.7 Bistable push-button.

gloves, hence they are inappropriate for public space (although they are often applied there, e.g. in lifts). They also have a minimal indication function and therefore must be made clearly visible.

7.5 Bistable push-button

A push-button is a control with two stable states which changes the state at every operation (switch function) (Figure 7.7). It must have a displacement except when there is a memory and a display for the state (compare with control signals). The positions of the button in the two states should be clearly different since they act as information about the actual state. The lower position belongs to state 'on' or '1'. For dimensions, operation force, click point and acoustic feedback the same recommendations are applicable as for the key with a key-click.

7.6 Bistable momentary toggle switch

The logical function of the toggle switch (Figure 7.8) corresponds to that of the bistable push-button. However, the toggle switch has some great advantages. The actual state, which corresponds to the angle between switch and casing surface, is visually and tactilely even more easily recognizable than with a push-button. The operation is less time consuming, and a very brief 'on' phase can easily be achieved. According to American convention, the 'on' position corresponds to the switch tilted to the right or upward or away from the user; according to European convention to the right or downward. Toggle switches are especially appropriate for

Figure 7.8 Bistable toggle switches (left part from Burandt 1986, middle part from Woodson 1986 and right part from Baumann and Lang 1998).

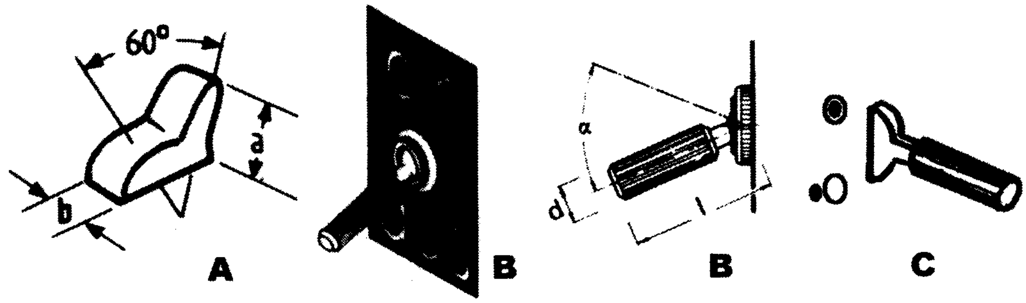


Figure 7.9 Trigger switches (A from Burandt 1986; B from Grandjean 1991; C from Baumann and Lanz 1998).

horizontal casing surfaces and for an arrangement in a row with common rotation axis. In some cases toggle switches with a third, intermediate state are used.

7.7 Trigger switch (bistable or with three switch positions)

The logical function of the trigger switch (Figure 7.9) also corresponds to that of the bistable push-button. When compared with the latter it has the same advantages as the toggle switch. However, it is also well suited for vertical casing surfaces. One disadvantage is that the label cannot be written on the switch itself but only on the casing. If additional trigger switches are closely spaced, confusion can occur. Therefore the labels should exactly obey the guidelines.

Special advantages of the trigger switch result from the larger lever-arm which makes greater operation resistance possible as well as a pronounced tactile and acoustical feedback (click). Hence trigger switches are especially suited for portable electronic appliances which are operated by the user while he/she is moving and not looking at the appliance. Examples are metal detectors and remote controls for model aeroplanes.

7.8 Control knob (continuous)

The control knob (Figure 7.10) is a control for the analogue (infinitely variable) adjustment of a one-dimensional variable. It permits rapidity of the adjustment process and precision over great ranges that is not simultaneously achievable with any other control. In electronic devices, the dimensions of the control knob are such that it can be operated with two or three fingers. Diameter and height should be between 10 mm and 30 mm. The handling surface has to be grooved, especially if the operating resistance (return force) is not small. A rotation angle up to 120° can be performed by hand in one turn. But larger or unrestricted rotation angles may also make sense.

Clockwise rotation should correspond to an increase in the variable. The zero point or the left edge of the adjustment range should be situated

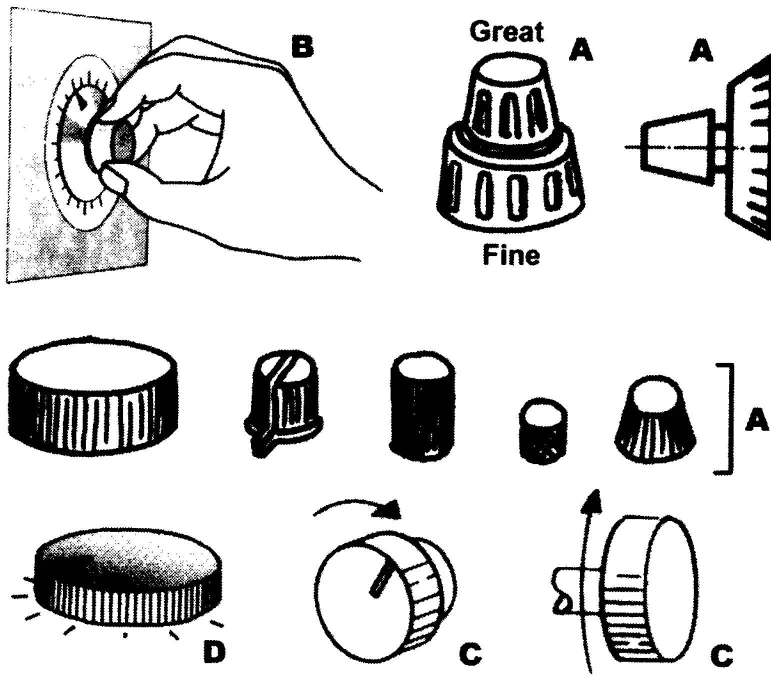


Figure 7.10 Continuous control knobs (A from Burandt 1986; B from Grandjean 1991; C from Woodson 1987; D from Baumann and Lanz 1998).

between the 6 and 9 o'clock position. At the left stop or at some other marked angle the control knob may markedly engage. In this case the device may have an integrated binary switch which switches the device off or brings the control into an undefined state.

The control knob must have a position mark which indicates the variable on a scale printed on the housing of the appliance. This is not necessary if the variable is represented on a separate display. Otherwise a scale label and a position mark on the control knob must be present. They must not be covered by the hand during operation. For that reason, arrow-shaped knobs or knobs with a disk below are especially suited. Control knobs and rotary switches with similar functions may be arranged concentrically above each other. This makes alternate operation without changing the grip possible, and the functional connection is clearly visible.

7.9 Control knob or rotary switch (discrete)

If the variable to be adjusted has three or more discrete states, a rotary switch suggests itself. The latter can be shaped like a control knob for continuous adjustment. However, the operating force should be higher. Correspondingly, strongly grooved shapes or arrow-shaped knobs are preferred. Between two switch positions the operating force as a function of the rotation angle should behave as in Figure 7.11. This will result in a marked tactile feedback.

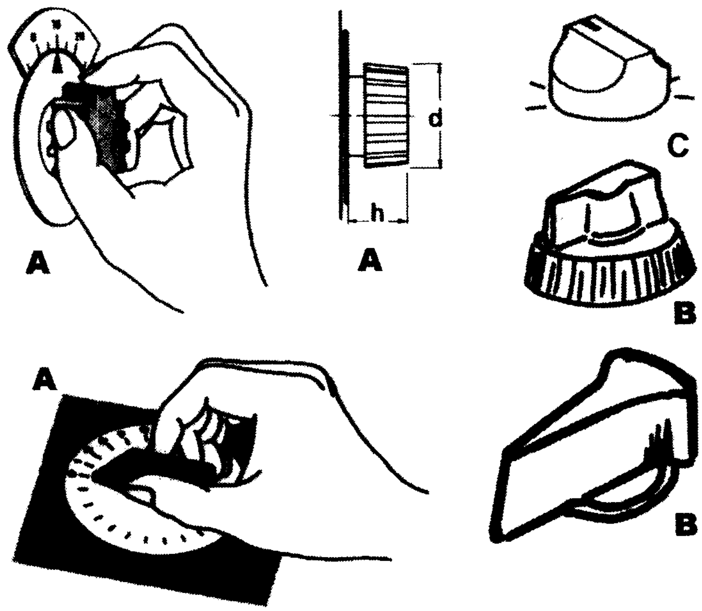


Figure 7.11 Control knobs with discrete positions – rotary switches (A from Grandjean 1991; B from Burandt 1986; C from Baumann and Lanz 1998).

The rotation angle traversed in one switching step should amount to at least 15° if the switch is read only visually, or at least 30° if the switch position is to be touched, without looking at the switch. The zero points or the minimal values of the variables of a switch should be situated between the 6 and 12 o'clock positions. Several rotary switches arranged in a row should have the same zero-position angle. If possible, the switch positions in the 'off' state should be directed along the line connecting the switches or orthogonal to it. In that case this state is recognizable at a glance even if there are many switches ('scanline').

7.10 Thumbwheel and shuttle (side-mounted and front-mounted rotary control)

A thumbwheel is a cylinder standing out from the casing with only part of its mantle showing (Figure 7.12). The mantle of the cylinder must have grooves. The wheel may be positioned at the side of the device and not in front of the user. Depending on the position, it is operated with the tip of the thumb or index finger. The shuttle (Figure 7.13), on the other hand, is a cylinder whose side plane is visible and almost flush with the surface of

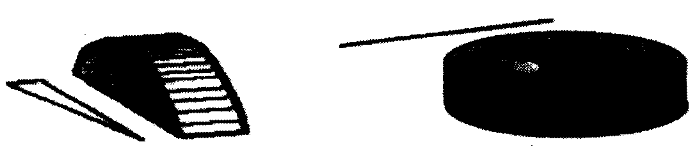


Figure 7.12 Thumbwheel (left) and shuttle (right) on a horizontal device surface.

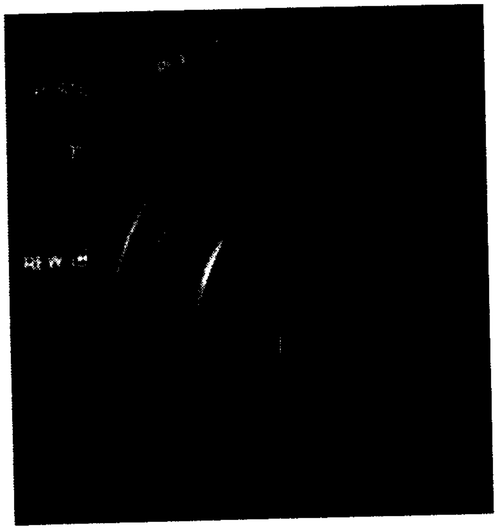


Figure 7.13 Shuttle in a combined control on a vertical device surface.

the device. The side plane of the shuttle has a smooth concave notch for rotating it with the index finger.

Thumbwheel and shuttle preferentially have a large or infinite range of rotation angle and require low operation force. As opposed to the control knob, these two controls are supplied neither with a scale nor with a position mark. The state of the variable must be recognizable for the user through separate feedback (e.g. display brightness, sound volume, a gauge).

The thumbwheel is especially suited for precise adjustment. The shuttle is appropriate for quick adjustment over wide ranges. At one end or in the centre of the range of possible orientations there may exist a marked orientation where the wheel noticeably engages. In this position a binary switch internally connected with the wheel may be operated (e.g. standby switch for the device or bypass switch for the input variable). An increase in the variable should correspond to a turn to the right, upward or away from the user for the thumbwheel, and to a clockwise rotation for the shuttle.

7.11 Thumbwheel (discrete)

For a stepwise adjustment of discrete states only, a thumbwheel must show a visible scale or position mark. Therefore it must be positioned at an edge of the casing in such a way that an inscribable surface is visible. A scale on the mantle is not good enough. In such a case only the actual value, rather than the whole range of values of the variable, would be visible.

7.12 Slider (continuous)

Sliders for analogue (infinitely variable) adjustment have the same function as control knobs (Figure 7.15). They are especially well suited for being mounted on horizontal or slightly inclined device surfaces, and with a

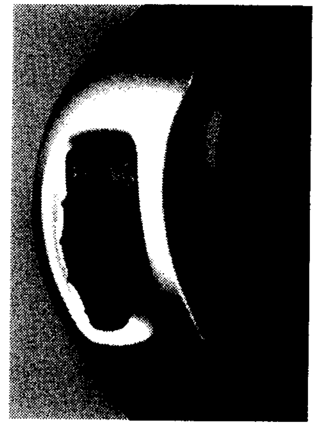



Figure 7.14 (Left) Thumbwheels for setting of time and aperture on the electronic camera Nikon F-401. Both wheels are locked automatically when turned into their normal position and can be unlocked by pushing the small button between them. (Right) Thumbwheel with three positions (up/down/neutral) used for menu selection tasks in the mobile phone Sony CMD-Z1. At the same time, the wheel serves as a monostable key for confirmation.

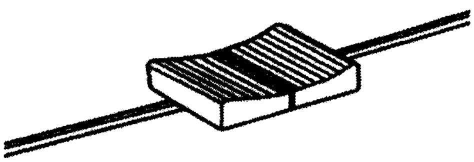


Figure 7.15 Continuous slider.

sliding direction parallel to the user's line of sight. The reason is that pushing sideways or upwards is less powerful and less precise. The advantages of sliders take effect if at least two of them are parallel. In that case they can be moved individually, in pairs or synchronously in groups. With some practise, simultaneous opposite motions are also easily done. Therefore sliders are particularly used in audio mixers. In spite of their often large number they are readily surveyable. Rows of sliders may favourably be applied in combination with displays which represent block diagrams (creation of analogies).

Throttle levers in ships and aeroplanes are also frequently arranged parallel in pairs and can be moved simultaneously by hand. To be precise, these are rotatable levers with a similar appearance to sliders.

7.13 Sliding switch (discrete)

Sliding switches (Figure 7.16) for stepwise adjustment have the same function as rotary switches. However, adjustment is not possible as unerringly and precisely as that of rotatable controls. Moreover, positions cannot be judged quickly, neither visually nor tactilely. When adjusting discrete values, the user must look at the switch or count; however, a

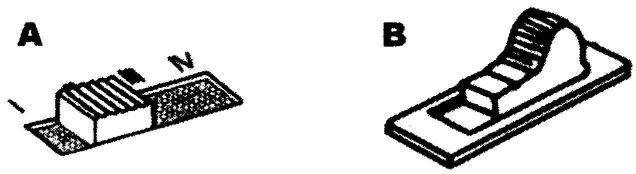


Figure 7.16 Discrete sliding switches (A from Baumann and Lanz 1998; B from Burandt 1986).

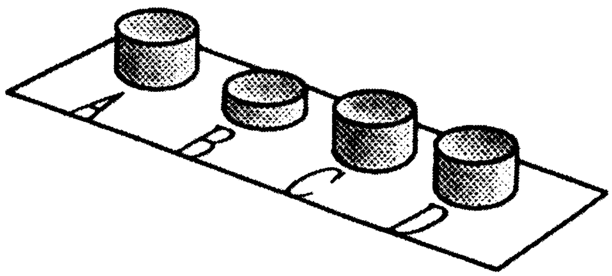


Figure 7.17 Radio buttons (key row with alternating function).

skilled user can operate a rotary switch with steps of 30° without looking at it.

7.14 Key row (alternating bistable keys, radio buttons)

A key row (Figure 7.17) consisting of a number of *n* keys with simple choice is a control with *n* discrete states. Thus the number of keys corresponds to the number of states. At any time only one key can be pressed (i.e. be in the 'on' state). By operating a key the state of the key operated immediately before automatically returns to the 'off' state. The key row does not correspond to *n* independent controls, each representing a two-valued variable, but to one single control with one *n*-valued variable. Key rows are frequently used in radio sets for the selection of stored frequencies. The advantage in comparison with the rotary switch is the possibility to switch from one state to any other without having to run through the intermediate states. Furthermore, each key may be arranged within a visible functional unit (modular structure) (e.g. together with a separate display).

7.15 One-shot key with control signals

There is a possibility to make controls with several stable states out of one-shot keys by (Figure 7.18) combining them with a memory and one or several display elements. The latter indicate the state of the control variable currently in memory. Light-emitting diodes (LEDs) are suited as a display as well as symbols on a liquid-crystal display (LCD) or keys illuminated from within. It is quite important that the connection between the key and the corresponding display element is clearly recognizable by the user.

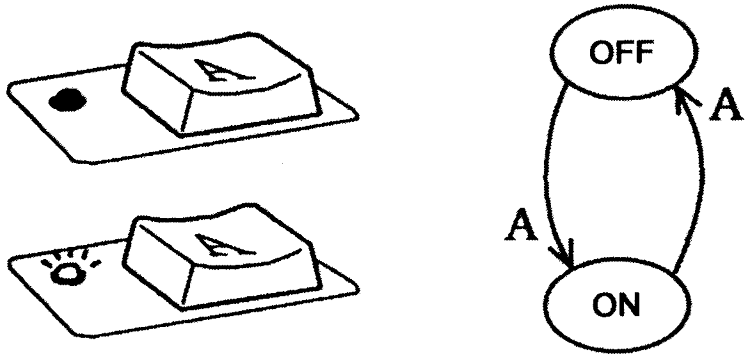


Figure 7.18 One-shot key with control signal in both states and corresponding state diagram.

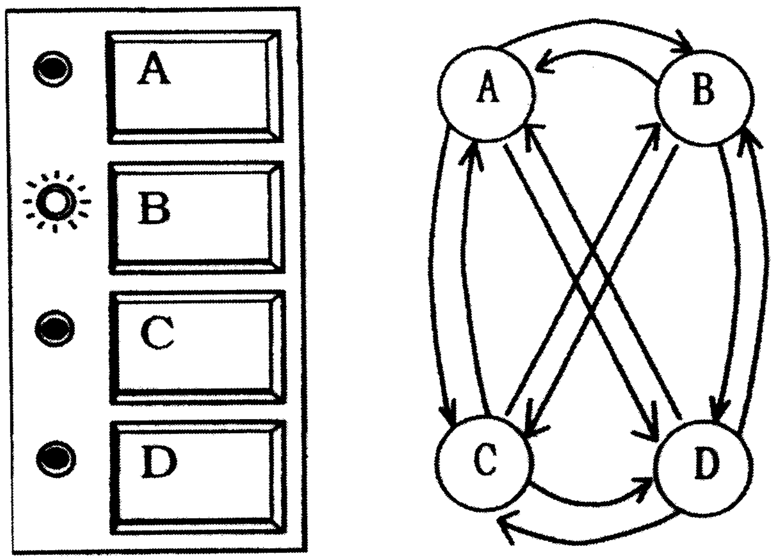


Figure 7.19 Key row (pseudo radio buttons) consisting of one-shot keys with alternating control signals and corresponding state diagram.

In the examples shown in Figures 7.18 and 7.19 controls with two and four stable states are constructed from LEDs, keys and a coloured border. They correspond in their function to a bistable key and to a key row with alternating function (simple selection), respectively. Hence the respective state diagrams are also the same.

The differently coloured LEDs must be positioned to the left of the keys or above the keys in order not to be covered by the right hand during operation. One advantage of these electronically controlled controls is that they can be switched by the device itself. The displays may also fulfil additional functions (e.g. a light diode may additionally serve as a warning signal by blinking).

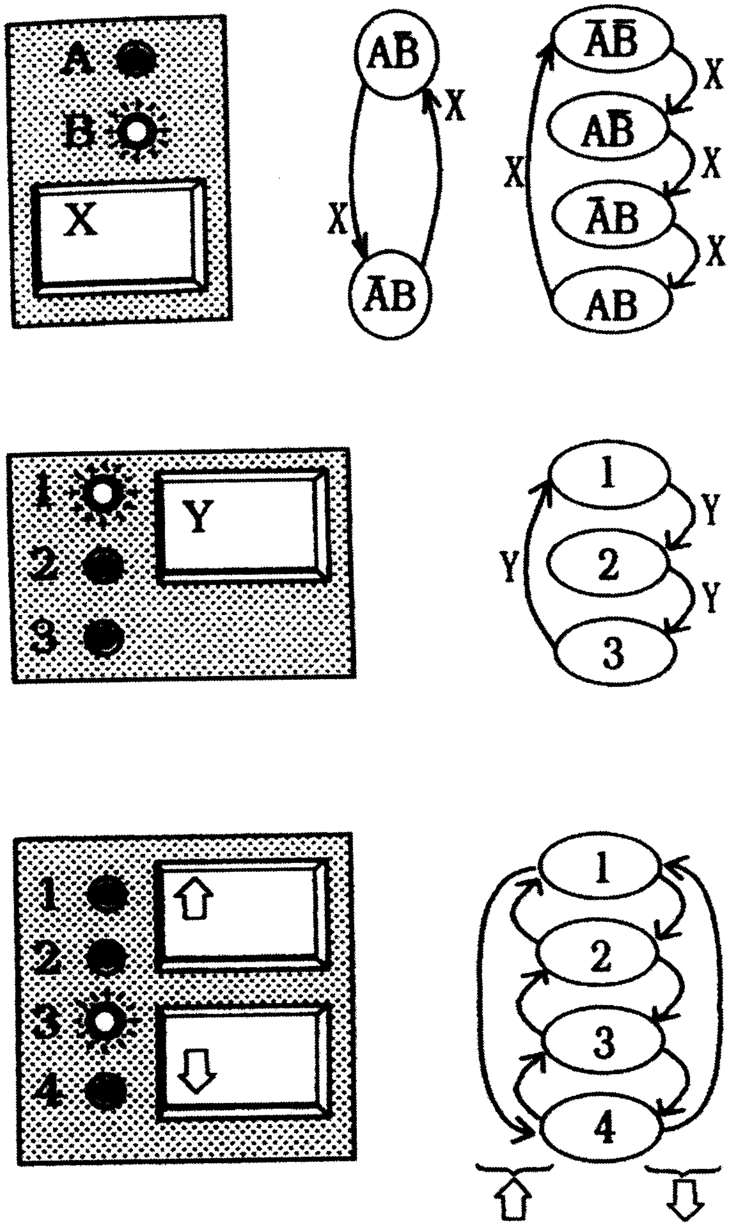


Figure 7.20 Possibilities to imitate rotary switches or toggle switches by the use of one-shot keys and control signals.

7.16 Arrow key(s) with alternating control signals

The function of a rotary switch (Figure 7.20) may be imitated by the use of one-shot keys, a memory and display elements. A rotary switch is operated by repeated rotation into a neighbouring state. Here, the two arrow keys correspond to the two rotation directions of the rotary switch; the states of the displays (the light diodes, in this case) correspond to the stable switch positions.

The control can have a cyclic state diagram, just like a rotary switch: a closed circle of states. In this case a single arrow key suffices. But if the switch is replaced by keys and if the number of keys is reduced, the

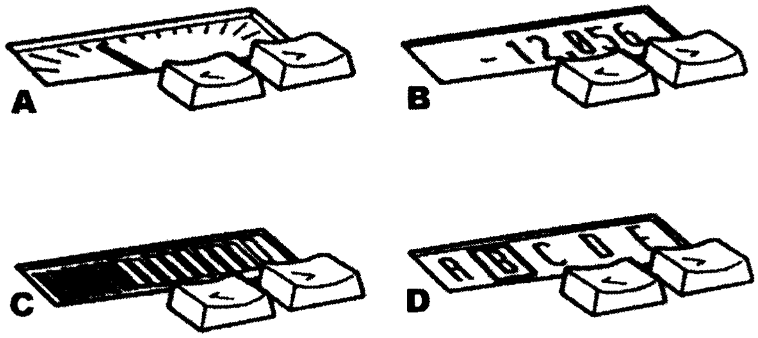


Figure 7.21 Arrow keys with (a) analogue, (b) digital, (c) pseudo-analogue, (d) alternating control signals.

operation of the control becomes lengthy. In order to pass through several states the user has to operate a key several times. However, it is possible to install a repetition function which is activated by prolonged pressing. Nevertheless, the operation of arrow keys generally requires more time and attention than the operation of rotary switches. A tactile feedback of the actual state is impossible; hence, during adjustment the user must permanently compare the actual state with the desired one.

7.17 Arrow keys with display element

Arrow keys (Figure 7.21) controlling a continuous or nearly continuous variable have the same function as a control knob or a slider. When they control a variable with discrete values they resemble a rotary or a sliding switch. The advantages and disadvantages of the respective electronic designs of the control are identical to those described in the previous paragraph. The electronic controls are more flexible, cheaper and controllable by the appliance, while the mechanical versions are quicker, less error prone and more comfortable. Thus, ergonomic considerations favour the conventional devices.

The advantages of both variants are combined in control knobs and sliders equipped with a motor for (remote) control. They are used, for example, in professional audio mixers for recording studios and as a volume control for audio amplifiers.

Arrow keys with display elements, in principle, have a similarity with the virtual controls to be described in Section 7.27. Arrow keys do not directly control the state of the variable but rather the speed of its change. The proper choice in speed of change will be considered later.

7.18 Multifunctional keys and shift keys

Monostable multifunctional keys are mainly used in keyboards (Figure 7.22). They act together with at least one shift key which switches the function of several keys at the same time. An example is the typewriter

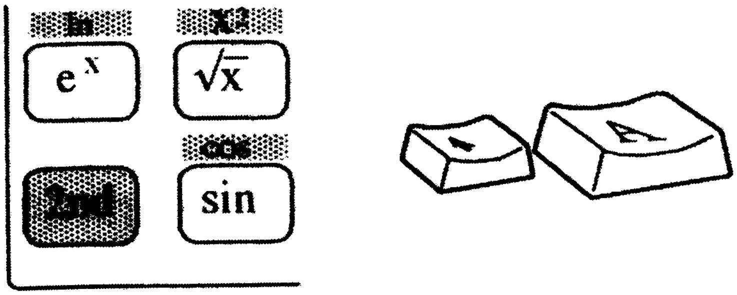


Figure 7.22 Multifunctional keys and corresponding shift keys.

keyboard, with one monostable and one bistable shift key for the capital letters and the computer keyboard having in addition 'control' and 'alternate' keys. Furthermore, some pocket calculators have keys with two or three functions generally labelled in colours.

More than three labelled functions per key are not advisable because in this case the keyboard would become difficult to survey. A multiple function without labelling contradicts the principle of self-explainability. But it may be assigned for the acceleration of frequently used processes ('hot key'). However, the same function must also be addressable explicitly (e.g. via a menu).

7.19 Soft key (software controlled key)

A soft key (Figure 7.23) is a semi-virtual control; that is, the labelling and the function of the one shot key do not exist physically. They are not part of the hardware but belong to the software of the appliance and may be changed according to its state. Soft keys are preferentially arranged so that the soft label and the hard keys can be associated.

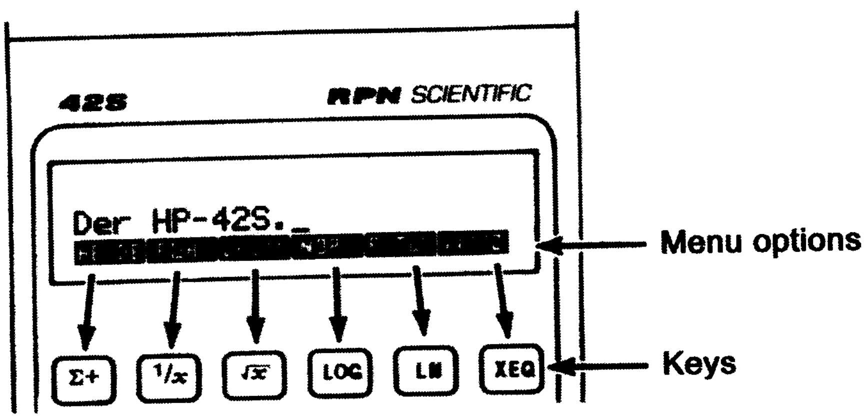


Figure 7.23 Soft keys in the upper line of a calculator keyboard (Hewlett-Packard).

Figure 7.24 (Left) Two-dimensional isometric joystick or trackpoint built into a keyboard; (right) two-dimensional isotonic joystick in a model aeroplane remote control.

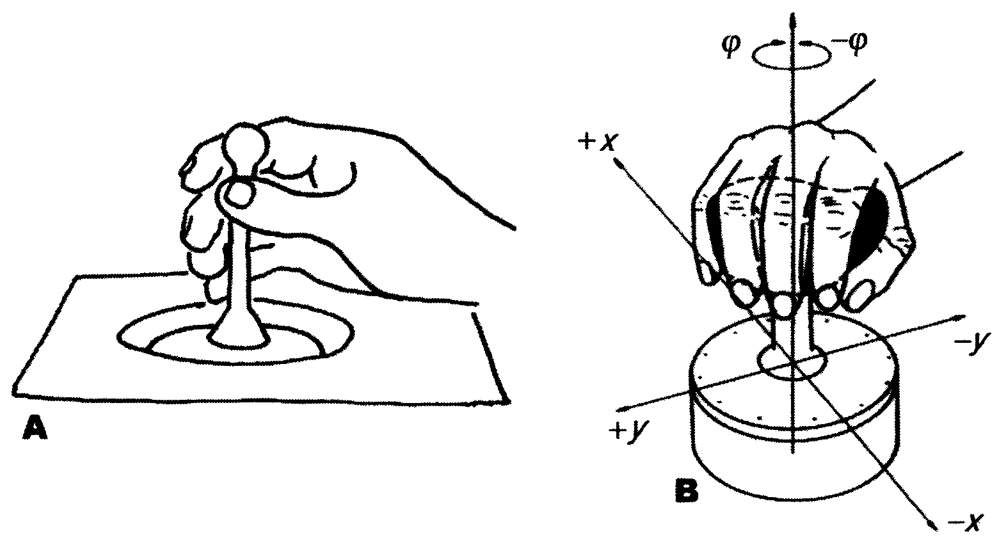


Figure 7.25 (Left) Two-dimensional and (right) three-dimensional isometric joystick (A from Baumann and Lanz 1998; B from Schmidtke 1993).

7.20 Isotonic joystick

An isotonic joystick 'is a control with the shape of a lever whose original position is vertical and which may be displaced in different directions' (Geiser 1990). It affords controlling one or two continuous variables and automatically returns to the vertical position when released.

The values of the variables are proportional to the angle of the original position. The name isotonic indicates that the return force (resistance) is approximately equally large for each displacement. The isotonic joystick is a quickly adjustable control but not a very precise one. Its advantages are small space consumption, easily and intuitively learnable handling and good tactile feedback about the state of variables.

It is appropriate for control and tracking tasks (e.g. steering aeroplanes, model planes, computer-simulated planes, cranes and lifting ramps). Its use is especially appropriate where the movement of the joystick is immediately analogous to the movement of an appliance part or to a visual display (Figures 7.24 and 7.25).

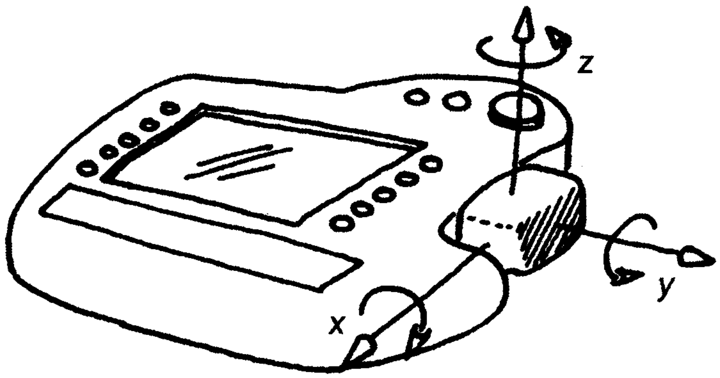


Figure 7.26 Six-dimensional control ('6D mouse') for steering an industrial robot. For every one of the three axes linear movement and rotation can be applied. The device is built into a remote control with screen and keys (Kuka robots). (The Kuka robot control shown has been redrawn by K. Baumann after Zühlke (1996b.)

7.21 Isometric joystick (force pick-up)

An isometric joystick is an analogue, self-centring control which measures the forces applied on it and yields up to three variables. Its properties correspond to those of the isotonic joystick. The name isometric indicates that the displacement is negligible. The third variable is gained preferably from the rotation around the lever axis. The force pick-up method is not used as frequently as the distance pick-up because it has no purely mechanically working counterpart.

A two-dimensional isometric joystick built into a keyboard is called a track-point. It has the advantage of small size and does not even require any space for being bent sideways when used. With an isometric joystick one can maximally control six variables – three translatory ones and three by rotation. With a control of this type the arm of an industrial robot may well be controlled (creation of analogies) (Figure 7.26).

7.22 Control column (discrete joystick)

For a digital control column, a three-valued variable with a zero state and two displaced states is related to each dimension. Without operation the variable is in the zero state. Operation causes a switch to be closed which yields an acoustical and tactile feedback (click). This type of control column may be considered as a discrete isometric joystick or as a discrete isotonic joystick (Figure 7.27).

Its two three-valued variables are much less precise than the continuous variables of joysticks. Nevertheless there are applications where the use of a digital control column is advantageous. For regulations with multiple integrating control circuits (e.g. large mass inertia in an acceleration process) a stable system is possible only if the user mentally integrates once or

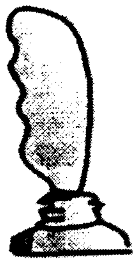


Figure 7.27 Two-dimensional discrete joystick.

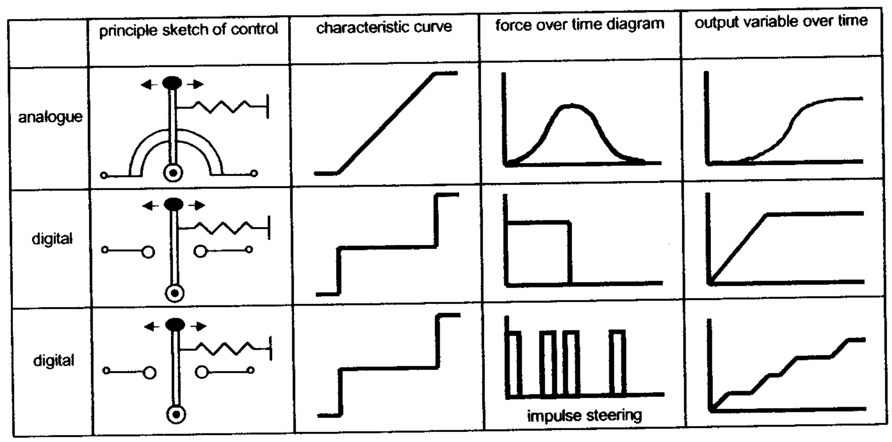


Figure 7.28 Principles of steering with analogue and digital controls. (Redrawn by K. Baumann after Schmidke 1993.)

repeatedly the time course of the control signal. This is much more easily done by counting control impulses than by multiplying duration and value of a continuous variable (Schmidtke 1993). Figures 7.28 and 7.29 compare analogue and digital controls in principle and by means of an example.

7.23 Arrow keys (one-dimensional or two-dimensional)

An arrow key is a one-shot key marked by an arrow (Figure 7.30). The use of one-, two- or four-arrow keys makes sense. The arrow keys should be arranged in such a way that the backward extended arrows meet in a point. Another optional arrangement is to have the arrow keys pointing left, downward and right in one row and the upward pointing arrow key above them in the middle. The control consisting of a combination of two- or four-arrow keys corresponds in its function to a one- or two-dimensional discrete control column.

Arrow keys have less favourable ergonomic properties regarding input speed and error security than the discrete control column but they are usable for the same tasks. For reasons of cost and space they are often applied instead of a trackball or a mouse. However, this is possible only

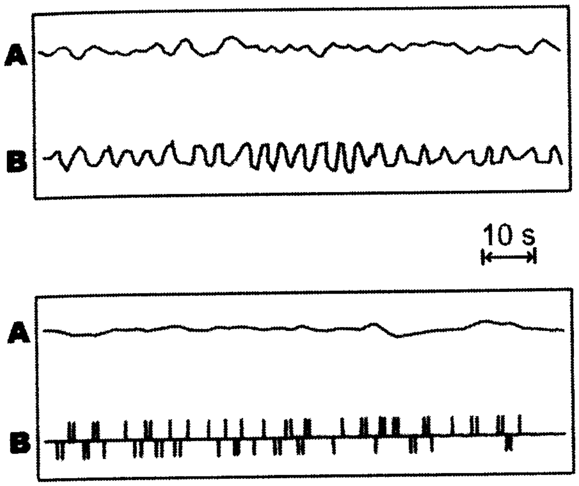


Figure 7.29 Control error over time A and control force over time B in an acceleration system using an analogue (upper diagram) and a digital (lower diagram) control. It can be seen in this case that better results are achieved with the digital control. (Redrawn by K. Baumann after Schmidke 1993.)

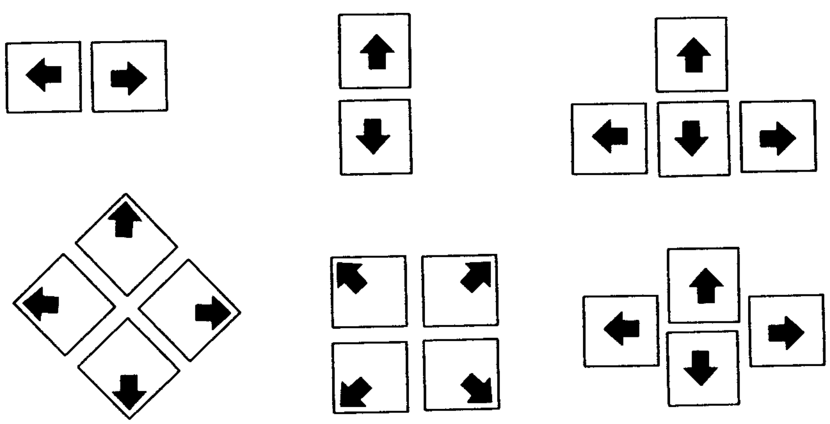


Figure 7.30 Arrow keys, possible arrangements.

at the price of a considerable loss of input speed or precision and requires more time and attention by the user. Only if the number of states of the controlled variable, or its speed of change is strongly restricted for other reasons, the use of arrow keys suffices for the control.

The processing of the input variables by the appliance is especially important when using arrow keys (repetition function, feedback). This will be described in detail in Chapter 8 on keyboards.

7.24 Trackball

A trackball is a two-dimensional, continuous control consisting of a freely revolving sphere and two mechanical or optical sensors (Figure 7.31). The sensors measure rotations of the sphere with respect to two orthogonal axes, thus describing a two-dimensional continuous variable. The diameter

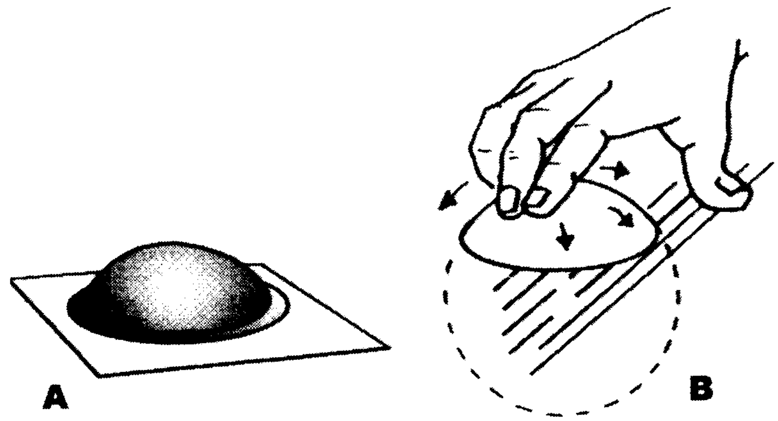


Figure 7.31 Trackball (A from Baumann and Lanz 1998; B from Woodson 1987).

of the sphere is between 2 cm and 10 cm. About one-third of the sphere stands out from a case surface. Mass and friction should be chosen in such a way that the sphere in normal motion immediately stops after releasing, but briefly (up to 1 s) rolls on when in quick motion.

The two-dimensional variable of a trackball can be adjusted very precisely, and in a restricted range also very quickly. As opposed to a joystick the trackball is not self-centring. It can stay in every chosen state. The two-dimensional range of adjustment is unlimited. The trackball may be integrated into an inclined case surface. Besides application of the trackball as a computer input device its traditional application is in flight control systems. As with the mouse the trackball needs to be combined with at least one monostable key.

7.25 Mouse

A mouse (Figure 7.32) is a two-dimensional, continuous control measuring the motion relative to a smooth surface (e.g. a table top), and in this way produces a two-dimensional continuous variable. To achieve this, mechanical or optical sensors may be used. The optical mouse is more precise and reliable than the mechanical one. But it only works on a special support and in the correct orientation.

The mouse is a quick and precise control that can be stationary in every state of the unlimited adjustment range. Disadvantages are the necessity of a table top and the possibility of the mouse detaching from the appliance. Usually the mouse is combined with at least one monostable key (one-

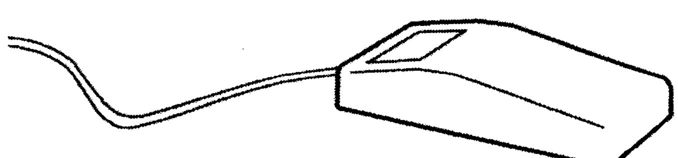


Figure 7.32 Mouse.

shot key) and so is well suited for pointing, selection and drag-and-drop operations on a screen-based graphical user interface.

With both trackball and mouse it is possible to process the variable in a velocity-dependent way. If the motion is quick then the variable changes more over a set distance than at slow motion.

Like the trackball, the mouse requires a display for representation of its variable. Ideally this is a graphical display in connection with a cursor or cross-hair. One problem with the mechanical mouse, as well as with the trackball, is the fact that the sphere may get dirty. For this reason their use in production plants is less advisable than in offices or laboratories.

7.26 Touch-sensitive devices

There is a variety of two-dimensional controls based on flat touch-sensitive devices. Feedback about the input process takes place via a display. The input surface is a pad, a tablet or a screen which produces a two-dimensional variable (i.e. a pair of coordinates) by touching with the position indicator. Depending on the design, the position indicator may be a special device (i.e. a pen) or the user's finger. There are many adaptations for such input tools, which use electric resistance, capacity, light or sound for recording the position.

The input surface is primarily suited for the input of coordinates (i.e. for drawing and for operating virtual controls), with appropriate programs also for input by handwriting. In the following subsections four basic types of touch-sensitive device are described and compared (Table 7.1).

7.26.1 Touch screen

A touch screen is a complete synthesis of control and display. It is especially suited for appliances with changing, novice users. It offers a maximum of flexibility in limited space.

Since the functional principle is practically unrecognizable (i.e. the device is not self-explanatory), the control must be made clear by the program belonging to the touch screen (the graphical user interface). Once the user has understood the principle he will enjoy the excellent visual feedback

Table 7.1 Four classes of touch-sensitive input device.

		Pointer	
		Finger	Pointing device
Input area	On the display	Touch screen	Touch screen with pointing stick
	Separated from display	Trackpad	Graphics tablet

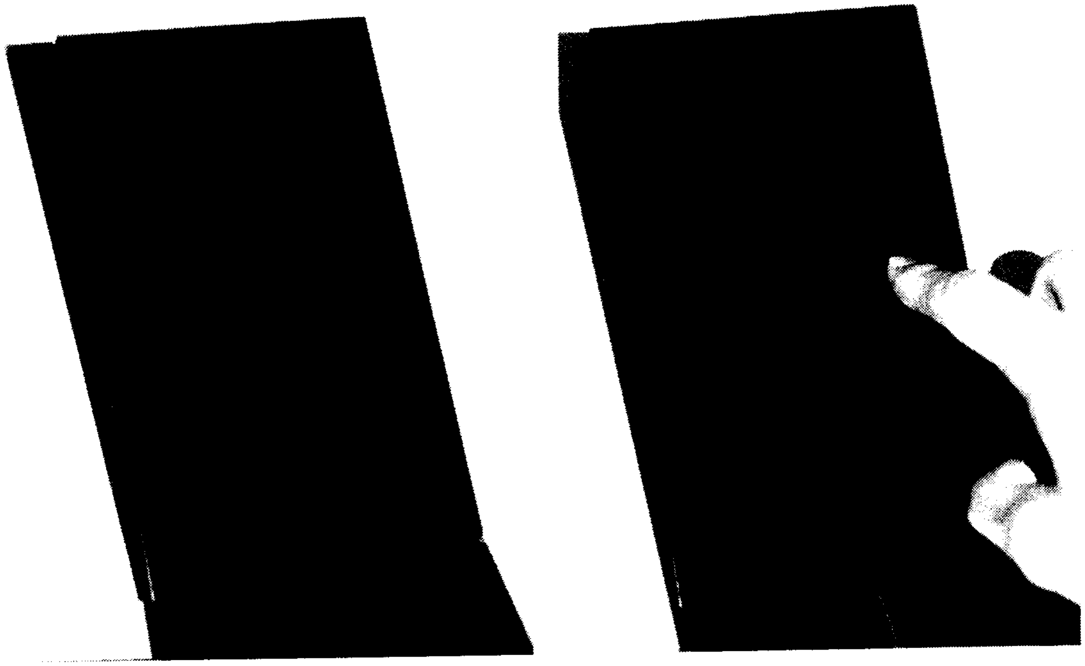


Figure 7.33 Bang and Olufsen audio/video remote control. The glass panel is touch sensitive on the whole surface. An LCD is built into the upper and LEDs into the lower part of the panel.

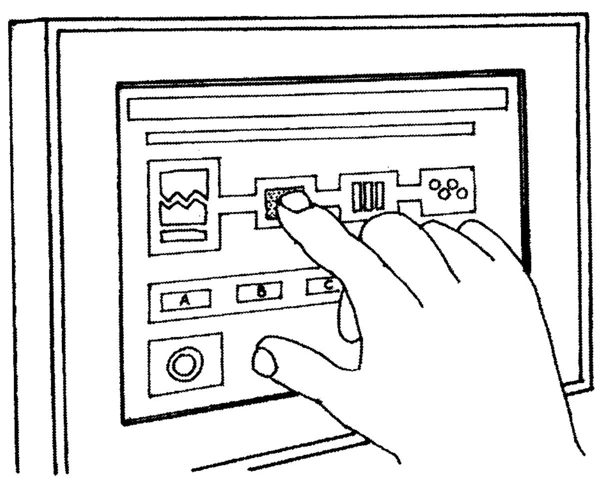


Figure 7.34 Control unit for industrial applications (Siemens Simatic operator panel) that has a 10.4 inch TFT (a type of LCD) touch-sensitive screen and an integrated Pentium processor without fan, and is resistant against moisture, dust, etc.

during operation. The touch screen is quick but imprecise. It is well suited for selection processes by the aid of virtual controls, menus and symbols.

If the screen is fixed vertically at face level the user's arms may become tired after prolonged operation. Furthermore, the user's hand can easily cover parts of the screen. With horizontally positioned screens in the workplace there is the possibility of operation by mistake. Since touch-screens have no movable parts they are frequently used in appliances in public spaces, like automatic teller machines (ATMs) or public information kiosks.

The *dimensions of hot spots* (virtual keys) shown on a finger-operated touch screen are generally the same as the pitch of hard keys, but 2 mm should be added to the diameter to account for parallax when aiming at the hot spot. Thus, the minimum dimension is 12 mm (vertical) and 14 mm (horizontal). Pheasant (1986) recommends that keys have a minimum diameter of 6 mm, and that there should be a minimum distance between keys of 6 mm. This gives a minimum pitch of 12 mm in all directions. However, this assumes either square or circular keys. In fact, the human finger is larger in the horizontal dimension than the vertical, and it is reasonable to reduce the vertical pitch by up to 2 mm (our own trial-and-error results) resulting in a minimum vertical pitch of 10 mm. For touch screens, we need to allow a larger pitch to allow for targeting errors (parallax, etc.). Adding 2 mm then gives 12 mm and 14 mm.

7.26.2 *Touch screen with pointer*

The first input technology used for this purpose were light pens in conjunction with a cathode-ray tube (CRT) display. Recognition of the position was taken over by the hardware part controlling the screen. Today, touch-sensitive devices are used in combination with a flat screen (LCD) and a cheap plastic pen (Figure 7.35).

This control is much more precise than its counterpart without a pointer; thus, its resolution is much better. The pen should have an easily accessible place for quick storage – otherwise it can be easily lost. Besides selection processes, the touch screen with pen is very well suited for drawing and for input by handwriting.

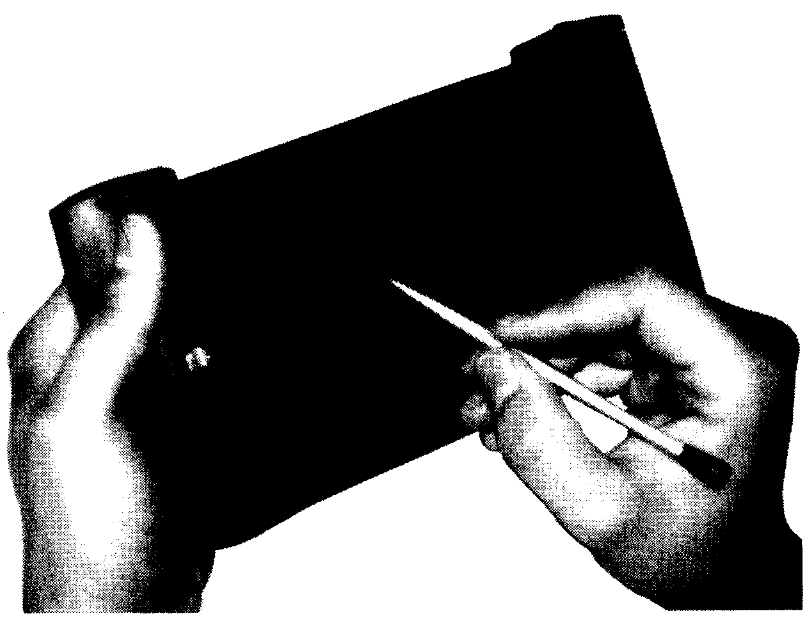


Figure 7.35 Apple Newton (trademarks) personal digital assistant.

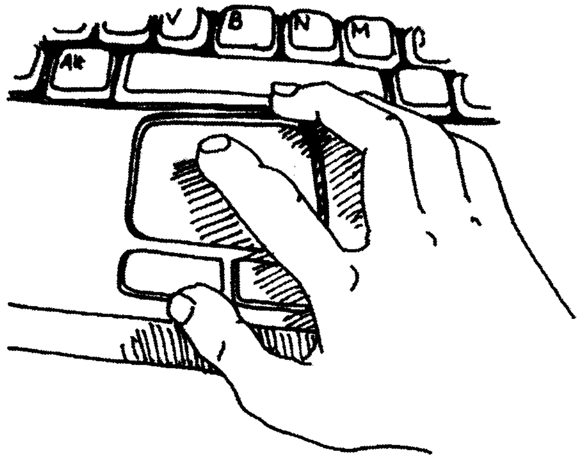


Figure 7.36 Trackpad built into a laptop computer.

For more details about handwriting recognition see Chapter 9.

7.26.3 Trackpad

A trackpad (Figure 7.36) is an input area separated from the display, which is operated like a touch screen with a finger. It may even be separated from the appliance and be designed in the form of a tablet. The corresponding program may control the variable of the control element in different ways. Helander (ed.) (1988) refers to three modes: absolute, relative and joystick.

In the *absolute mode* of the control element, the coordinates of the screen correspond to those of the input surface or a section thereof. Touching the surface with the finger puts the variable onto the absolute coordinates of the touching point. Then, the program generally puts the cursor on the display onto the corresponding position. Thus, the absolute mode imitates the function of a touch screen.

In the *relative mode*, on the other hand, an initial touching of the input area does *not* change the variable. Only a movement of the finger along the surface moves the variable into the corresponding direction. Thus, the relative mode imitates the function of a mouse or a trackball.

Joystick mode: Finally, the input surface may also imitate the function of an isotonic joystick. Touching the surface corresponds to a displacement of the joystick to the respective position.

The centre of the quadratic or circular input surface corresponds to the resting position of the joystick. Without touching, the variable of the device returns to the central resting position.

The last two modes are well suited for small input surfaces whereas the absolute mode requires a larger surface. For selection processes the control element must be combined with a monostable key. This key is preferably positioned below the input surface, i.e. on that side of it facing the user. In this way the device may be operated by the index and the key by the thumb of the same hand or by the index of the other hand.

For a touch-sensitive input device there may be a need to have additional input information which replaces the separate monostable key for selection processes. This may be achieved by distinguishing between three discrete touch values 'no touch / pressure / strong pressure'. More commonly the monostable key-press is emulated by making the touch-sensitive device able to detect a click (short touch), double-click (two short touches in the same area within a short time interval) and grasp (short touch immediately followed by a long-lasting touch in the same area) operation. The grasp operation may be followed by a move and a release operation which results in a drag-and-drop operation commonly used in graphical user interfaces (GUIs).

The GUI developer must be aware of the fact that this software-aided functionality enhancement of a touch-sensitive device is not common knowledge of the average user. If no other way to perform the click, double-click, grasp and drag-and-drop operation is provided there is a need for special training or assistance for novice users.

The visual feedback about the actual state of the variable (e.g. about the position of the cursor) is much more important in the case of the described touch-sensitive devices than in the case of the touch screen. Thus, if a touch-sensitive device is used, it is advisable to make the cursor conspicuous during operation. For a touch screen, on the other hand, the cursor is unimportant during an operation because it disappears behind the position indicator (hand or pen). Therefore the selected symbols and menu options should be rendered prominent.

A small finger-operated touch-sensitive area is relatively imprecise because of the relative magnitude of the finger. In addition, lifting the finger can easily produce a small shift of the variable's position which must be corrected by the program if more accuracy is necessary.

7.26.4 Graphics tablet

A graphics tablet (Figure 7.37) is a pad of relatively large size on which the variable is entered by the aid of a mouse-like positioning device. The absolute position of the mouse is measured by the tablet's built-in hardware via inductive coupling. The mouse has at least one key and a transparent part provided with cross-hairs that makes exact positioning possible. The tablet has a grid with scales for the input of coordinates. The input of the coordinates is done by measuring the actual position when a mouse key is pressed and not by shifting the mouse on the surface.

The graphics tablet has higher precision than trackpads. For this reason it is mainly used as an input device for computer-aided design (CAD) software. Secondary (or multiple) input layers or input areas filled with virtual keys may be added to the tablet's primary grid of coordinates.

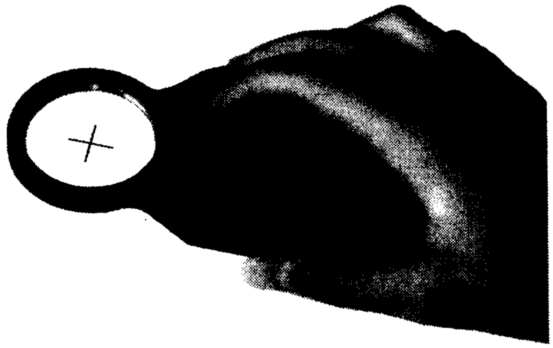


Figure 7.37 Positioning device used with a graphics tablet.

Specific functions or elements may be assigned to these keys (e.g. commands or predefined construction elements – shapes, mechanical parts, electronic devices, architectural construction elements) for CAD. They are printed on the tablet surface and can be selected by pressing keys on the mouse. Moreover, additional keys may be mounted on the mouse (e.g. a numeric keyboard).

Because the graphics tablet is one of the input devices with the most features, using it requires more time and training to get acquainted with than a trackpad, trackball or mouse.

7.27 Virtual controls

Virtual controls are generated by means of software and presented on a display screen. Their function, however, is usually analogous to real controls (like the ones presented earlier in this chapter). The transition from real to virtual controls is shown in Figure 7.38. In every row of Figure 7.38 there are three controls having the same functionality and the corresponding state diagram. In the leftmost column there is the 'real' control showing the selected state by the physical position of the moving part of the control. This gives immediate all-time visual and tactile feedback and is a real benefit for the user – not just the most expensive way of implementing the input device.

The middle column shows the version of the control storing the state of the variable electronically and giving visual – but no tactile – feedback via a control signal (LED or LCD). The next column shows the corresponding virtual representation of the control. It does not exist physically at all, but is implemented in software and represented on screen. It is operated by means of a cursor – the virtual finger of the user – and a separate input device (e.g. a trackball or a mouse with a key for confirmation of the selection).

Virtual controls have the highest flexibility – they can be hidden when not needed. However, to hide a control means that the user cannot have any feedback about the state of it. It is a challenge for the UI designer to overcome this flaw and make it a virtue.

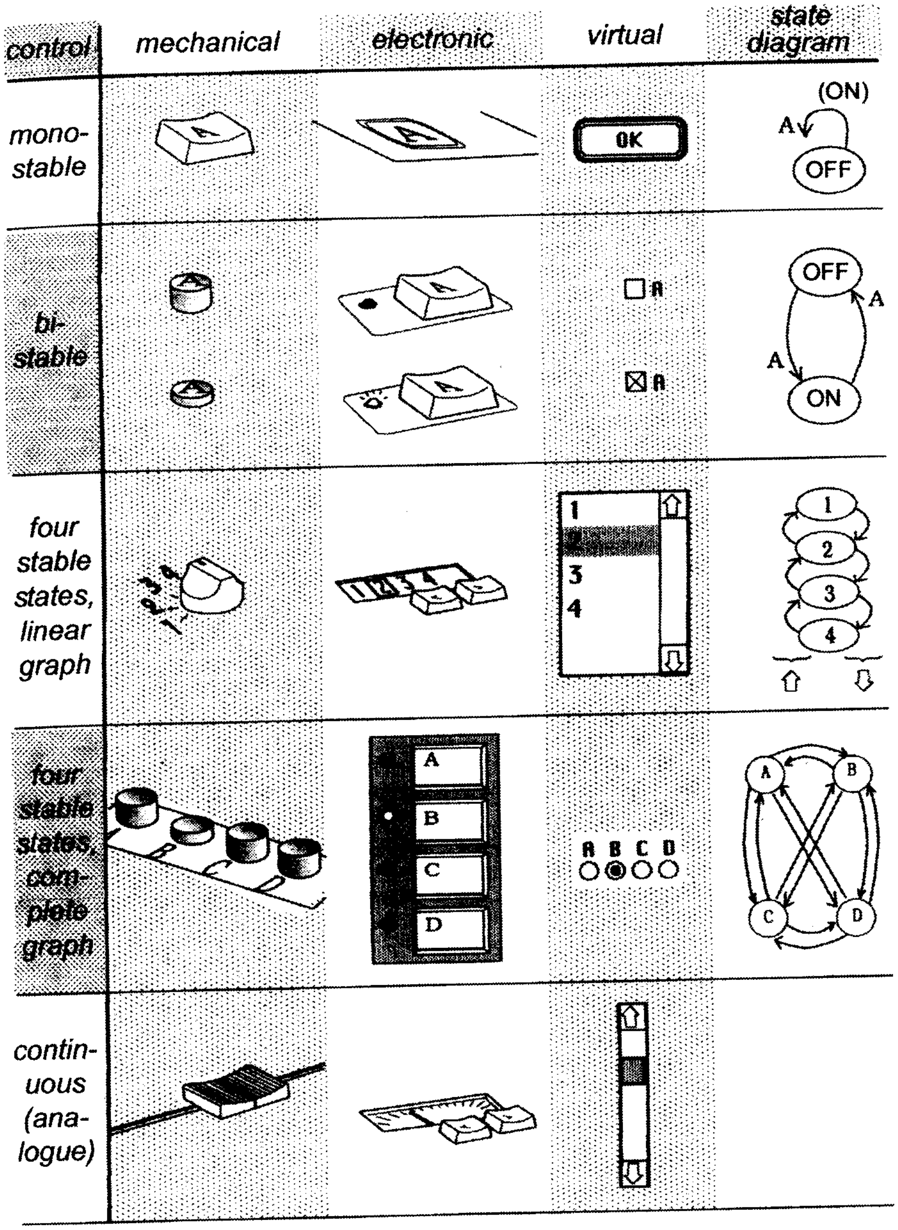


Figure 7.38 Transition from real to virtual controls giving five examples.

The rightmost column shows the state diagram corresponding to all three (real, electronic, virtual) representations of the control.

7.28 Safety and security

For security reasons it may be necessary to protect a control against accidental use (e.g. by a child). This can be done by making it difficult to operate the control, so that special knowledge or skills would be needed (e.g. some kitchen ovens having controls which the user has to push before use – only then do they come out of the surface of the appliance and can be turned on). Another example is the reset key of most laptop computers. As these appliances are designed for mobile use, it can happen that conventional keys are pressed by accident. To obviate this, the reset key is hidden under the housing and can only be operated by pushing a needle or a similar object through a hole.

In case of emergency switches protecting against injury, e.g. the emergency shut-off for laboratory equipment, quick and easy operation must be ensured. So such a control is protected against accidental use by a label and bright red colour, making it obvious that it is not a normal control.

7.29 Summary controls

Figure 7.39 summarizes the controls described in this chapter and Table 7.2 outlines some properties of input devices. More details about other input devices and methods will be given in Chapters 8–11.

Table 7.2 Summary of input devices to be used with displays and their properties (Greenstein and Arnaut in Helander 1988).

	Touch screen	*Touch screen with pointing*	*Graphics tablet*	*Mouse*	*Trackball*	*Joystick*
Eye–hand coordination	+	+	O	O	O	O
Free view to the display	–	–	+	+	+	+
Common housing	+	O	+	O	+	+
No parallax problems	–	–	+	+	+	+
Resolution (accuracy)	–	–	+	+	+	+
Flexible positioning	–	–	O	O	+	+
Small footprint	+	+	–	–	+	+
Short training time	+	O	O	O	O	O
Comfortable for constant use	–	–	O	O	+	+
Absolute coordinate input	+	+	+	–	–	O
Relative coordinate input	–	–	+	+	+	O
Can emulate other input devices	–	–	+	–	–	–
Suitable for:						
Pointing	+	+	+	+	+	–
Quick pointing	+	+	O	O	O	–
Pointing with confirmation	–	O	O	+	O	–
Drawing	–	–	+	O	–	–
Accurate drawing	–	–	+	–	–	–
Tracking of slowly moving targets	O	O	+	+	+	–
Tracking of quickly moving targets	–	–	O	O	O	+

control	principal sketch	max. dimension, analogue/digital	number of stable states	preciseness error tolerance	speed
key (monostable, with path)		D 1	1	+	+
foil key (monostable, without path)		D 1	1	–	+
bistable push-button		D 1	2	+	O
bistable toggle switch		D 1	2	+	+
trigger switch		D 1	2..3	+	+
continuous control knob		K 1	∞	+	O
discrete control knob – rotary switch		D 1	2..10	+	O
continuous thumb wheel		K 1	∞	+	–
shuttle		K 1	∞	–	+
discrete thumb wheel		D 1	2..20	O	–
slider		K 1	∞	O	+
discrete sliding switch		D 1	2..5	O	+
key row (radio buttons)		D 1	2..10	+	+

comments:

maximum dimension = maximum number of variables	'K' = continuous, analogue 'D' = discrete, digital	'+' = good 'O' = neutral '–' = bad

Figure 7.39 Summary of standalone controls and their properties.

control	principal sketch				
one-shot key with control system		D 1	2	+	+
arrow key(s) with alternating control signal		D 1	2..5	+	–
arrow keys with analogue display		K 1	30..∞	–	–
multifunctional key with shift key		D 3	1	O	–
soft key		D n	1	O	O
isotonic joystick		K 2	1	O	+
isometric joystick		K 3	1	O	+
discrete joystick		K 2	1	O	+
arrow keys		D 2	1	–	–
trackball		K 2	∞	+	O
mouse		K 2	∞	+	O
touch screen		K 2	0	–	+
touch screen with pointer		K 2	0	O	O
trackpad		K 2	0	–	O
graphics tablet		K 2	0	+	–

Figure 7.40 Summary of controls to be used together with a display.

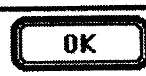
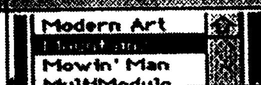
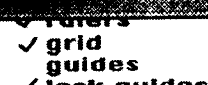
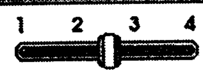
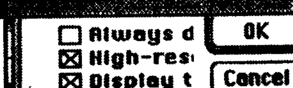
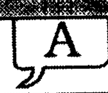
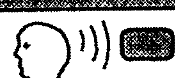
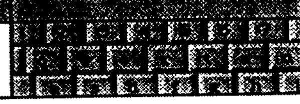
control	principal sketch				
virtual monostable key		D 1	1	+	+
virtual bistable push-button		D1	2	+	+
virtual key row		D1	2.10	+	+
menu with alternating items		D 1	2.30	+	O
menu with multiple choice		D n	je 2	+	O
virtual analogue slider		K 1	∞	O	+
virtual sliding switch		D 1	1	+	+
virtual input area		K 2	0	O	+
virtual shuttle		K/D 1	∞	+	+
dialogue box		D n	je 1..2	+	+
movement sensor		K 3	0	−	+
voice input		K 2	0	−	−
EEG brainwave sensor		K	0	−	−
keyboard		see Chapter 8			

Figure 7.41 Summary of virtual and other controls and their properties.