

The book “Computer Peripherals” is a second revised edition. The book contains the whole experience of the author gained in the area of computer peripheral devices and their use in CAD environment (Project Automation Systems). It offers a classification and defines which the basic technical parameters are. The most common modern interfaces, Input/Output graphic devices are presented. Graphic language for plotters` control HP – GL is presented in a short way. A special part of this book concerns 3D devices and some elements of multimedia.

The book “Computer Peripherals” is intended to be used as a reference material and textbook for students at university level, studying Computer Science. It will also be used for training of professional groups for whom the training in Computer Peripherals is desirable as well as for PhD students and researchers, working in different areas of science.

© Prof. DSc Raycho Todorov Ilarionov – author, 2013
© University publishing house “Vasil Aprilov”- Gabrovo, 2013

ISBN

Preface to the Second Edition

The second edition of the book Computer Peripherals is an extensively revised model of the first one. Modern interfaces and new technological peripherals are included in it. The chapter on analogue input-output information has been deleted entirely. Its place is taken by 3D input-output devices.

Chapter 2 has been extended with additional information about Centronics interface and IEEE 1284 standard for parallel interface has been dealt with. In the section about serial exchange, attention is paid to the universal serial bus (USB). IEEE 1394 interface, optical infrared wireless interface (IrDA) and radio interface Bluetooth are described in detail. Topological diagrams and ways of communications are supplied.

In Chapter 3 new information has been added about modern displays based on flat panels (TFT LCD). Other display technologies are also discussed. The operating principle of multimedia projectors is concerned.

Chapter 4 has been renovated in the section on interactive manipulators and the description of modern optical 'mice'. Various types of scanners for black-and-white and colour copying have been added. The structures of the so called CCD matrices on which digital cameras are based have been described in detail. In the section on optical carriers new technologies for optical recording and reading CD and DVD – single-layer and multi-layer structures, standards, capacity, etc. have been proposed.

Chapter 5 is the only one that has remained almost unchanged. Output devices for 2D graphic information have not been further developed technologically in recent years. Only some cosmetic changes in their appearance and optimization of certain parameters have taken place.

Chapter 6 is on printing devices. Besides the early technological environments given in the first edition, more attention has been paid to jet technologies based on piezo-modulator and thermo-head. The operating principle of laser printing based on the electrostatic method of building the image is discussed in detail.

Finally, Chapter 7 is completely new. It covers 3D input-output devices which are associated with modern CAD. A model of 3D scanner is described, which inputs the relief of 3D graphical objects in computer environment. A kinematic model of X, Y, Z positioning system with laser measurement and axis drive has been covered. The functions of the system have been defined. An algorithm is given for converting the information obtained about the object in CAD format.

Devices for 3D information output are known as devices for rapid prototyping. Several methods of building 3D objects from materials with structure close to plastics are described in this chapter.

Preface to the First Edition

In the recent decade the work of designers was aided by the rapidly developing computer technologies. This resulted in abrupt enhancement of the quality of the designed products.

Computer design is based on the implementation of the so called design automation systems. They are applied in various branches of industry (mechanical engineering, electronics, power engineering, architecture, instrument making, biology, medicine, chemistry, agriculture, etc.). Each system consists of computing environment and peripherals. The computing environment is a computer configuration with software and data base. The peripherals are a set of devices which provide graphical input and output of the system.

Input devices for graphical information are described in Chapter 4. They are classified into: interactive manipulators, digitizers, scanners, cameras, etc. Interactive manipulators directly control the display cursor. Joy stick, light pen and trackball are discussed as representatives of this class. In the section on digitizers, the inductive method of digitalization is described in detail and a formal model of functions is given.

Devices for graphical output described in this book are: displays, plotters and printers.

Displays are dealt with in Chapter 3. They are classified and described in detail: devices for colour and monochromatic images, vector displays, graphical stations, advertisement display boards, etc.

Chapter 6 is devoted to plotters. Besides complete classification, the most common mechanical constructions have been concerned. Methods of controlling positioning systems have been discussed and the most widely used language for plotter control - HP-GL has been described.

Chapter 7 deals with printers. Daisy-wheel and matrix models are presented as printing devices. Modern technologies for laser and jet printers have also been included.

Interfaces for connection between the devices in the design automation systems are covered in Chapter 2. Parallel and serial exchange in the most common standard forms (Centronics, IEEE-488, RS-232, V24, RS-423, RS-485, I²C и D²B) has been discussed.

In computer control of manufacturing, it is necessary to monitor information received from various sensors as well as to control executive mechanisms. Analogue input-output signals are used to this end. The ways of analogue-to-digital and digital-to-analogue conversion of information have been described in Chapter 5.

This book should be of value to students studying computer engineering.

1. CAD/CAM SYSTEMS AND THE CONCEPT OF INTERACTIVE COMPUTER GRAPHICS. TOOLS. TOOL REQUIREMENTS

1.1. CAD/CAM Systems and the Concept of Interactive Computer Graphics

One of the main applications of computer technology is the Computer Aided Design (CAD), side by side with the Computer Aided Manufacturing (CAM).

Computer-aided design incorporates computing systems that help engineers in developing, modification, analysis and optimizing of various design solutions. An electronic computing system involves hardware and software tools which perform specialized functions in the designing process. Generally, the hardware consists of computers, one or several graphic displays, keyboards and peripherals including input/output devices of 3D objects. The software consists of computer programs, ensuring the work with graphic terminals, and application software which implement design and construction functions, specific for a definite user. Examples for application software are the programs for analysis of strength and stresses in elements of different mechanical structures; calculations of dynamic characteristics of mechanisms and calculations of heat transfer parameters; analysis and synthesis of electronic circuits; integral microcircuits design, etc. The set of specific application software is changed, modified and improved depending on user's demands and the available hardware for design implementation.

The term Computer Aided Design characterizes any design activity which incorporates an electronic computing machine or computer in the process of development, analysis or modification of technical design solutions. Modern CAD systems are based on the vast usage of Interactive Computer Graphics (ICG) tools. The term refers to user-oriented graphic systems designed for data generation, conversion and visual presentation.

The graphics system user is the designer who defines relevant data and commands on the computer by means of one of the available data input devices of the system set. The user interacts with the computer through the display screen. The designer creates images on the screen applying commands which address standard subprograms stored in the memory of the computing environment [7].

In most CAD systems the image on the screen is constructed with standard geometrical elements – points, lines, circles, etc. The generated image can be further modified by the user with the help of addressed commands, i.e. the image can be zoomed-in, zoomed-out, moved, rotated and subjected to other transformations. In the process of executing different manipulations on the image the user constructs its specific details.

The typical system of interactive computer graphics is composed of hardware and software tools. The hardware comprises central processing unit; one or more work stations (incl. graphic display terminals), and a set of peripherals (printers, digitizers, drawing equipment, etc.).

The software of ICG systems incorporates computer programs for processing graphical data, and special additional (not included in the system

set) application software intended to implement definite design functions as required by different users (Fig.1.1).

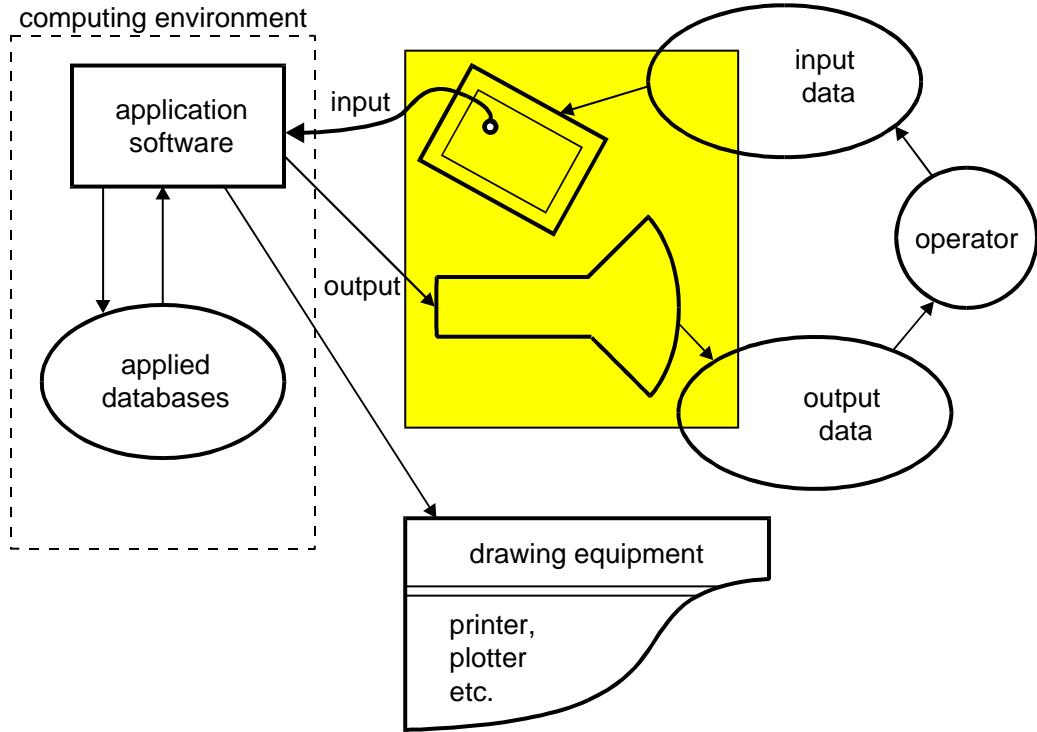


Fig. 1.1. ICG system programs set

It should be observed that the ICG subsystem incorporates only one of the CAD components. Another essential component is the designer who uses the ICG as instrument for solving engineering tasks. The ICG's tools enhance designer's creativity providing the so called synergy effect. Synergy is the phenomenon presenting the beneficial effect formed by the coordinated work of all system's components that is greater than the plain summed effect of each component. Within the "human-computer" entity the designer performs part of the designing work that corresponds to the utmost of his/hers intellectual abilities (conceptual project, independent thinking) while the tasks assigned to the computing environment demand high execution speed, visual data presentation and storage of vast data amount, adapted to computer solutions. As a result of such harmonious interaction between man and computer, the efficiency achieved by the design solutions is greater than the sum of the effects from the separate work of man and computing environment [6, 7].

The advantages of using CAD systems are determined by a number of basic factors [7].

1. Increased efficiency of designers' work. It is achieved with the help of the computing environment which provides visual presentation of the designed object, components and specific details. Moreover, the computer shortens the time for synthesis, analysis and documenting of the design solutions. The higher efficiency reduces not only the design costs but the time schedules as well.

2. Higher performance. The CAD systems open up broad opportunities for a thorough engineer analysis and examination of a great number of alternative design solutions. Due to the CAD higher accuracy the number of errors is reduced. As a result, the conditions for designing are practically improved.

3. Better communication between users. Drawings made with CAD systems are of higher quality, they fully comply with standards' requirements incorporated within the system, and the developed project becomes better documented. The user is prevented from making technical mistakes, the drawings become more comprehensive and legible.

4. Organizing databases for designing systems. The database contains the whole information required for the successful completion of the design. It may also include databases created for previous projects with their descriptions. Thus, previous projects may be used for references, or as starting points of new designs. For instance, a new object can be designed on the basis of an existing old object.

1.2. Technical Tools of CAD Systems

The hardware used by a CAD system can be divided into several groups (Fig.1.2):

1. Computing systems. They involve one or several computing machines connected in a network. Each of them can be:

- general purpose computers (universal), or
- specialized computers.

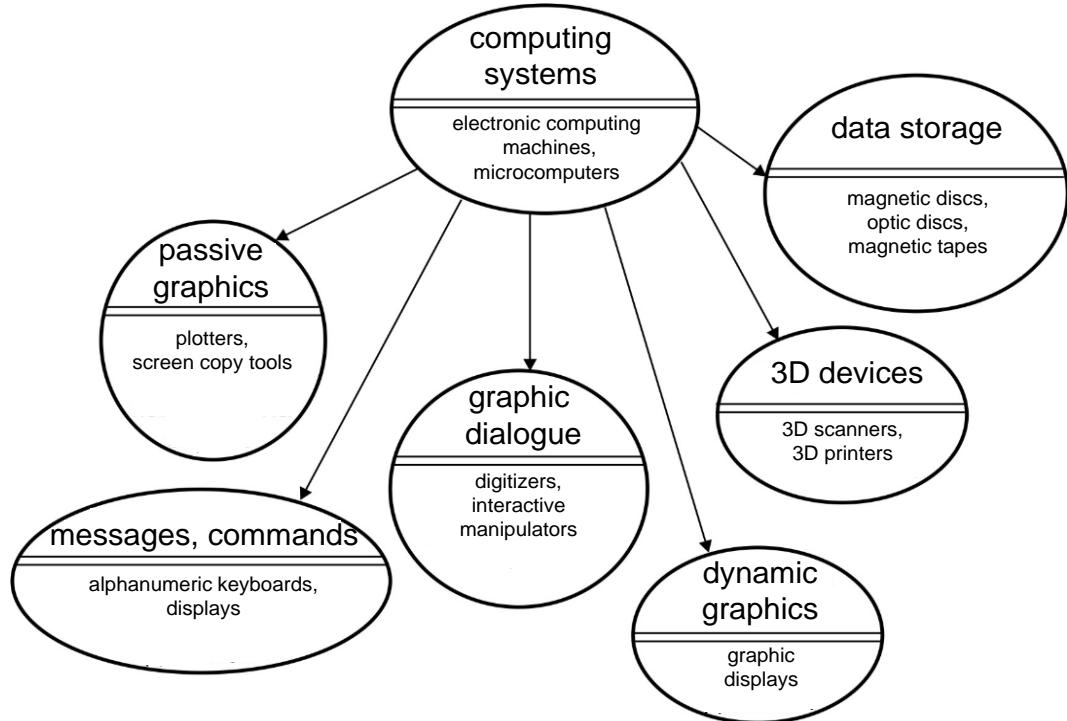


Fig. 1.2. CAD hardware

Mostly, the CAD systems use general-purpose computers. Specialized electronic computing machines are intended to solve a range of tasks in

association with the designing of specific technical objects. For example, modeling computers are used only for logic modeling of specific devices [8]. For the time being engineers are not provided with an objective criterion for identifying the type of computers regarding the nature of tasks solved with CAD. Generally, users solve the problem having in mind all the tasks connected with the automation of engineer's work, and certainly their financial funds. That circumstance has led lately to the wide usage of smart terminals and personal computers [8].

2. Data storage devices. They provide the possibility to process a huge amount of information. Recently, devices of high storage capacity (magnetic and optical discs, diskettes, magnetic tapes, USB memory, etc.) are widely introduced and applied. Their usage in CAD systems is determined by a number of factors:

- the enormous amount of data that in most cases has to be processed from the design start up to the production stage;
- the method of database creation and usage;
- the architecture of computer systems' networks faces the problem of database organization (local data, common data storage for groups or fields, etc.) [8].

3. Graphic data visualization devices. They are subdivided into:

- passive graphic devices: for creating drawings and pictures which cannot be changed;
- dynamic graphic devices: for creating graphic objects that can be deleted or changed at any time

4. Input data devices. On the other hand they are subdivided into graphic devices (devices for data transfer as coordinate values) and non-graphic devices (alphanumeric displays and keyboards).

5. 3D devices. Mainly, they are used for developing CAD models. The 3D scanners are devices with the help of which the relief of a geographical object can be digitally presented, while the common scanning devices just copy the color information of the image. The 3D printers belong to the class of rapid prototyping devices. They construct the 3D information in succession, generating interconnected points. Thus a conceptual model of 3D CAD file is quickly created without additional technological tools.

The designer operates with these devices at his working place by using relevant commands, graphic and non-graphic data for implementing the project. Lately, the gravity centre of CAD technical devices has been shifted to its peripherals, as they determine both the intellectual level of the technical provision, and the efficiency of engineer's work. From that point of view, the operational characteristics of a workplace are estimated by means of the following parameters:

- number and type of displays (alphanumeric and/or graphic);
- diversity of input and output devices available to the operator;
- ergonomic factors which are taken into consideration when organizing a workplace (working space, arrangement of devices, screen quality) [6].

The proper compatibility of all devices within the set, assigned to fulfill a task, depends not only on the hardware but to a great extent on the software, and therefore on the procedures to be carried out by the designer.

1.3. Requirements for Graphic Peripherals

Requirements for Passive Graphics Devices

The category of graphic peripherals includes devices which produce graphic images on a permanent carrier. The distinctive feature is that the produced image cannot be changed and the carrier should not be re-used.

The most widespread devices of this type are plotters. They are automatic drawing machines converting digitally coded data into images without operator's direct actions. Depending on the principle of the image generation plotters are divided into vector or raster type.

Classic plotters implement the vector principle for image building in compliance with the objects' specification. In fact, the plotter makes a motion according to the preset parameters of each vector, and the writing tool (pen) leaves a trace on the carrier equivalent to the motion.

Raster plotters can be classified as mechanical, electro-spark type, electro-chemical type, thermal, optical, magnetic and inkjet type.

Depending on the type of static carriage and the writing pen structure, the vector plotters are classified as flatbed, drum, roll-on and roller-frictional plotters.

Regarding functionality plotters are divided into two main groups:

First group – includes the following functions [1]:

- servo control on two-step or direct-current motors;
- maintenance of constant drawing speed;
- writing tool control;
- linear and circular interpolators;
- character generator;
- working with different number of pens;
- drawing different types of lines;
- scaling and rotation of drawings to angles multiple of 90°;
- availability of several tables with different characters;
- characters scaling and rotation to a certain angle;
- inclined characters drawing;
- setting up of plotting area and cutting off elements falling outside it;
- powerful language for program control providing remote access to all foregoing functions.

Second group – includes 17 functions in addition to the above mentioned functions:

- drawing rotation to a random angle (drawing transformation into the carrier coordinate system);
- text rotation to a random angle;
- text writing at adaptable distance between characters;
- polygon filling in a pattern (full filling, hatching, etc.);
- exchange different protocols with the computing environment.

Basic technical features of plotters:

1. Drawing size – set up as machine drawing paper sizes (A0 ÷ A4) in mm or inches.
2. Resolution – the smallest distance to which the writing element can move. It is measured in mm and is usually within 0,0125 ÷ 0,1 mm.
3. Accuracy – the error resulting from the movement of the writing tool from point to point at a random distance. It may have positive or negative

value. Usually the accuracy varies within 0,05 - 0,25 mm. For some non-flat vector plotter the accuracy is set up separately for X and Y coordinates.

4. Reiteration – the error resulting from the multiple motion of the writing tool to one and the same point.

5. Drawing speed – represents the plotter's speed and is usually within $25 \div 90$ cm/s.

6. Functions:

- linear interpolator;
- circular interpolator;
- character generator;
- geometric primitives;
- types of lines;
- translation;
- scaling;
- rotation.

7. Programming language

8. Types of interfaces – series RS 232 and parallel IEEE 488.

Requirements for Dynamic Graphics Devices

Displays are devices for showing a generated image that may be subject to further modifications without changing the information carrier. Displays can be of two types:

- alphanumeric:- the displayed information is a succession of alphanumeric characters;
- graphic:- the displayed information is both alphanumeric and graphic.

Displays are characterized by the following parameters:

- amount of displayed data;
- size of screen's working area;
- number of characters displayed on the screen;
- image refresh rate;
- possibility of working with definite number of colors;
- image display quality;
- options of screen division into random data zones.

The amount of information for alphanumeric displays is measured by the maximum number of characters set on the screen and it is defined as the product of the maximum number of characters in a line multiplied by the number of lines displayed on the screen. Depending on the type of graphic displays, the amount of displayed information is represented either by the number of address points, or as a summed up length of the graphic image vectors. For alphanumeric displays the number of displayed characters is within 128 – 160. As for graphic displays, the standard set of characters can be extended with specialized symbols which are frequently used in reproducing specific graphic images [5].

According to image regeneration method, graphic displays are divided into vector and raster types.

In vector displays the image is constructed by a controllable coordinate motion of the electron beam in consistency with the vector representation of

the graphic object. The more detailed graphic image, the longer is the time of object generation [1].

In raster displays regeneration does not depend on the image content. Therefore these displays are faster than vector ones. They have high resolution performance and are usually designed on the basis of Cathode Ray Tubes (CRT).

The display refresh rate is determined by the rate of image refreshing. This parameter is important in case of large and fast-changing amount of data, characteristic of alphanumeric displays. By rule, they are constructed with CRT, have short screen persistence time and maximum refresh rate [6].

The quality of the represented information is an essential ergonomic feature of a display. It is determined by the size of the image components, their brightness and contrast, as well as the lack of parasite-noise-images, etc. [6].

For enhancing the data processing capabilities, displays operating in autonomous mode have been developed (graphic stations). Their controlling device (display controller) is designed on the basis of microprocessor, mostly intended to carry out vector-raster conversion. They are called smart displays and possess a number of functions for graphic processing. The main functions are:

- scaling;
- translation;
- rotation;
- character generation;
- area filling with color shades;
- area hatching;
- marker motion as a function of interactive means.

Graphic displays are also evaluated with regard to interactivity level that is specified in two operations:

- possibility for deleting parts of the graphic image and modifying it further;
- time of display response to operator's query [6].

Requirements for Graphic Data Input

Graphic data input devices serve for message and data transfer in computing environment. Depending on their functionality these devices form several groups (fig.1.3).

1. Devices which allow data transfer by means of keys (keyboards). Keyboards can be alphanumeric or functional. Functional keyboards represent a set of keys that may cancel relevant operation when pressed, while the number of the pressed key is conveyed to a computer system. The software product always prompts the keys which can be used at a given moment [6, 15].

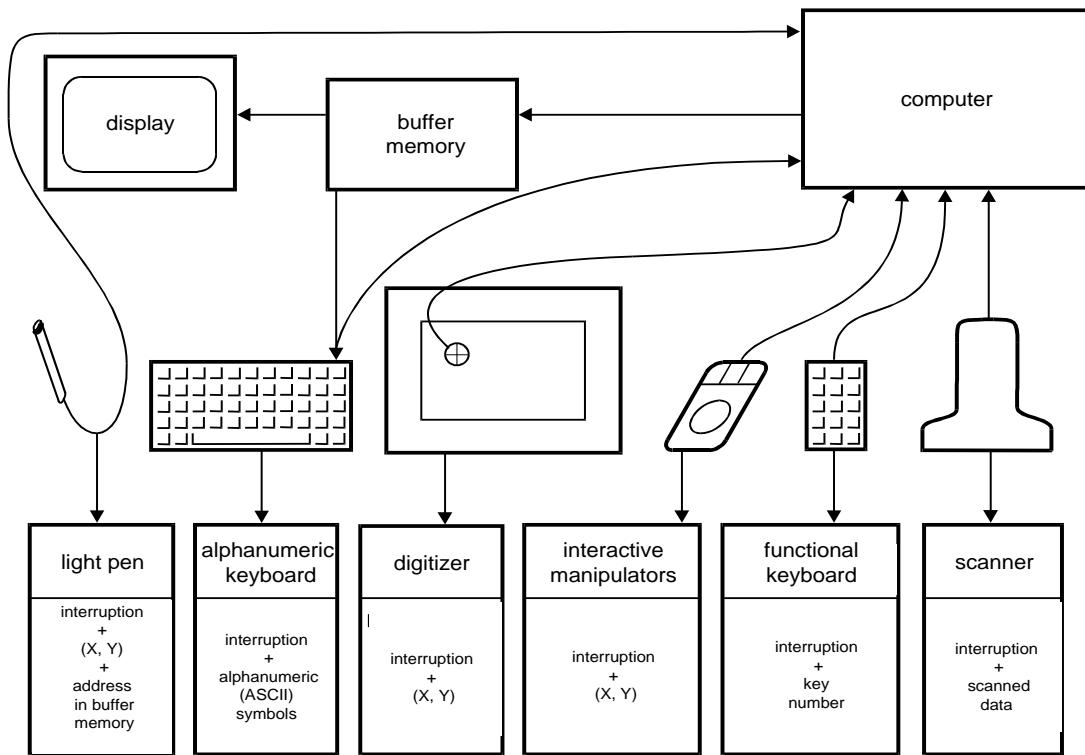


Fig. 1.3. Data input devices

2. Devices for direct specification of a point on the screen. The light pen is a representative of the group. The response of its operation is the initialization of a specified point (X, Y) on the screen. The light pen's accuracy is not high and it partially depends on the parallax [6, 13, 15].

3. Devices for indirect specification which allow the setting up of coordinates. This group involves different types of digitizers, tablets, mice and track balls. The last three are also called interactive manipulators [1, 14, 15].

The process of converting graphic data into digital code includes two stages: reading and encoding [8].

During the reading process the graphic element (point, line, elementary fragment) is recognized and its coordinates are relevantly specified in the assumed coordinated system. The encoding process performs data conversion into a digital code in accordance with the established rules.

Two-coordinate digitizers are used for input of complex graphic images, such as machine construction drawings. Graphic data reading is implemented by the operator by means of a positioning element called vizor or sensor. Digitizer's configuration involves working area for placing the document and a menu for choosing different functions and for alphanumeric data input. Digitizers are evaluated functionally according to their characteristics.

1. Working area size – it is set up in accordance with the machine drawing paper size (A0 ÷ A4) in millimeters or inches.

2. Resolution – the smallest step of positioning element's movement that leads to increase and decrease in the current coordinates. Most frequently digitizers' resolution is within 0,1 ÷ 0,025 mm.

3. Accuracy – the error that occurs during the movement of the positioning element from one point to another at a random distance in the working area. It may get positive or negative values. Generally the accuracy varies within $0,075 \div 0,625$ mm.

4. Speed of coordinate points output – number of points per second (P/s). This parameter conveys information about the device speed, and usually it is within $60 \div 200$ points per second.

5. Input mode

6. Functions

7. Programming language

8. Menu

9. Types of interfaces - series RS 232 or parallel

The rate of automation defines two types of digitizers – automatic and semi-automatic. In the first case, graphic data reading is implemented automatically (scanners). The second case includes human participation.

Scanners are designed for automatic data input of printed text or handwriting, graphics, drawings, photos. Data conversion into digital form is executed with the help of photosensitive sensors. There are three basic types of scanners – hand-held, desktop (flatbed) and drum type. The image, scanned by hand-held and flatbed scanners, is static and the photosensitive sensors are moving while scanning it completely. For flatbed scanners this is implemented by an automatic mechanism, for hand-held scanners – by the operator's hand. As for mechanically driven scanners (drum type), the paper, where the image is positioned, rolls over a rolling drum that moves it along static photosensitive sensors.

The main technical parameters of scanners are:

- maximum resolution along X and Y axes – it is given by the number of points per inch;
- scanning mode – contrast, semi-shaded, etc. depending on the scanner;
- contrast control – given with the number of grades;
- brightness control – given with the number of grades;
- possibility of changing the scanner's resolution;
- maximum scanning speed;
- scanning area;
- interface for data transfer (SCSI, IEEE-488 - GPIB, RS232).

Requirements for 3D Devices.

3D scanner automatically collects 3D coordinates from the object surface in a specified succession. It is characterized by the following basic technical parameters:

- accuracy – 0,025 mm;
- resolution – 0,005 mm;
- maximum motion along X, Y and Z axes;
- scanning methods (pre-set description of surfaces, or automatic);

- contact or contactless scanning method;
- scanning speed (number of points per second);
- output file format.

Devices for rapid prototype generation of 3D objects are known in practice as 3D plotters. Their main features are:

- maximum volume of generated figure (set along X, Y and Z axes);
- speed of image generation;
- type of technology used;
- accuracy of generated image;
- resolution – apprx. 0,1 mm;
- possibility of generating color images.

2. INTERFACES

2.1. Parallel Interface

Via parallel interface the data is transmitted by bytes in a parallel code from/to the microprocessor, and between the microprocessor and the memory. Generally, the standard specifies an 8-bit-data bus to which different external devices are connected.

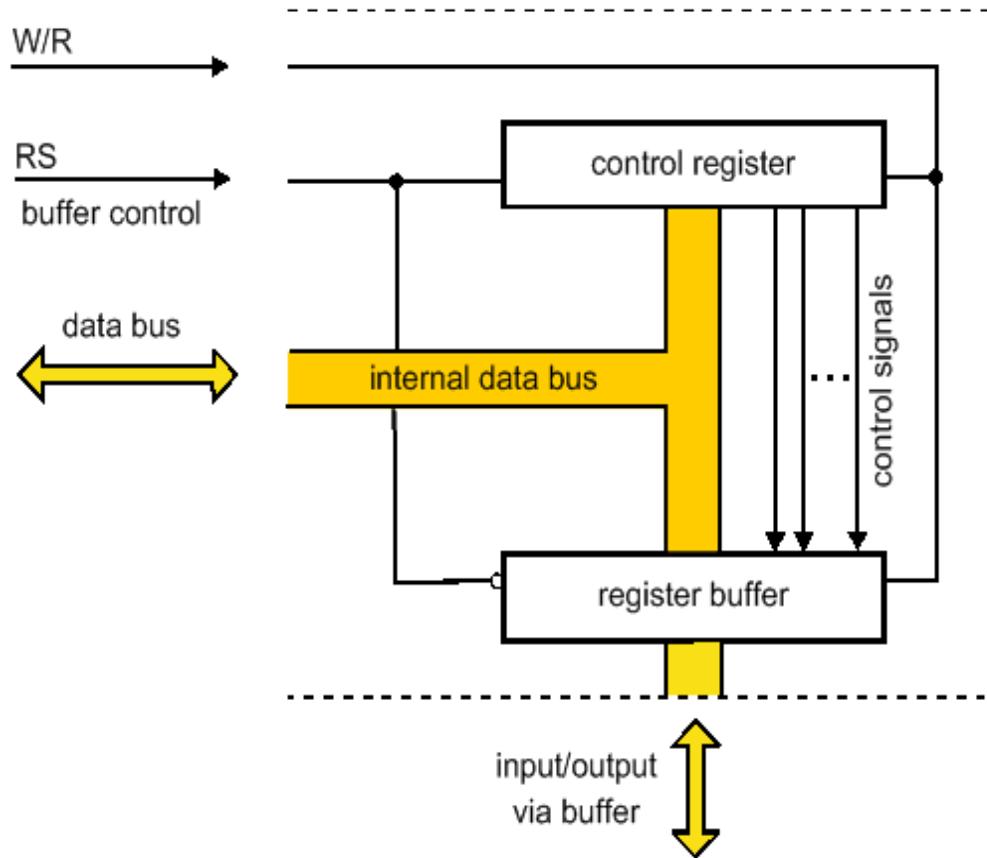


Fig. 2.1. Structural diagram of a device for parallel data transmission

The typical diagram for a parallel data transmission device (Fig.2.1) includes a control register and a register port (buffer). The data is loaded into a register buffer and the control register assigns the direction of the data flow - to the input/output devices. The selection of a register buffer or an I/O device is made by a decoder controlled by the address bus and the input/output signals. Therefore only one device can be selected at a time.

The signals in the input/output systems at complex data transfer are DR (Data Ready) and IDA (Information for Data Acknowledgement). They control the data transmission between the two systems (Fig.2.2 and Fig. 2.3). The Source transforms the data values at the input line and after a short time delay informs the Host that the data is ready (signal DR). The Host accepts the data and sends a signal (IDA) with a positive polarity.

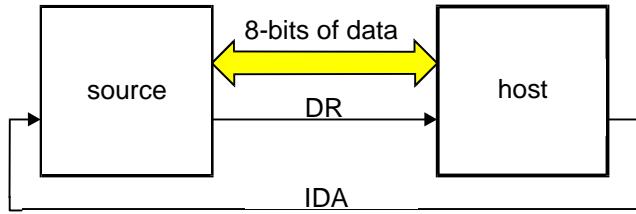


Fig. 2.2. Parallel interface signal lines

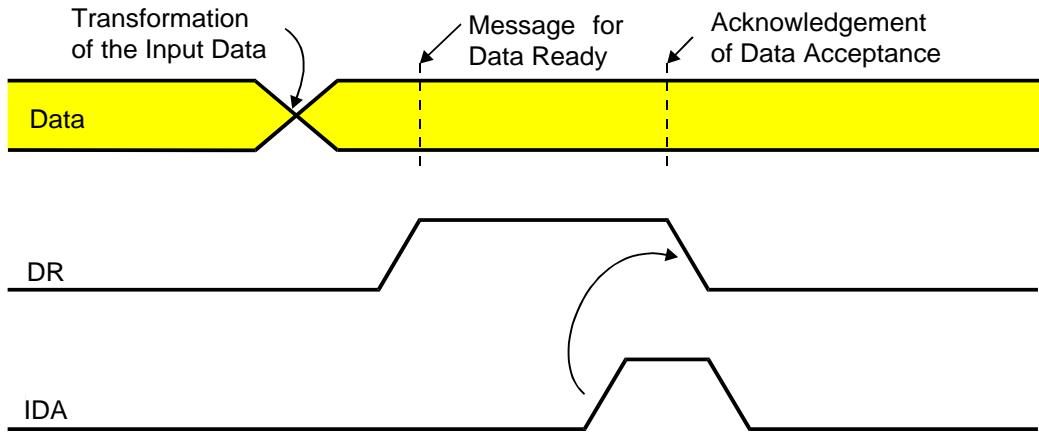


Fig. 2.3. Timing diagrams at data transmission over parallel interface

2.2. Centronics Interface

Centronics is a standard parallel interface designed for data transmission from a microprocessor system to a printing device (printer). Data transmission is analogous to the above described, and it differs only in the active values of the signals which respond to low levels.

The simplest Centronics interface model contains 7 or 8 data lines and two control signal lines: STR – data strobe and ACK – acknowledgment for data acceptance. Each signal line has its self-reverse connection and is connected to a definite pin of a standard connector. The organization of connections is shown in Fig.2.4.

Signals D0 - D7 on a personal computer correspond to pins 2, 3, ..., 9, and signals STR and ACK respectively correspond to pins 1 and 10. Data transmission is executed in the following sequence (Fig. 2.5):

1. The source activates the data on the data feed line D0 - D7.

2. Data strobe values STR are acknowledged valid.

3. The positive front of STR signal initiates the host for data acceptance (the host contains a logic scheme).

4. When the host is ready for a new data acceptance, it sends an ACK signal. After the positive front of ACK, a new STR signal can be produced.

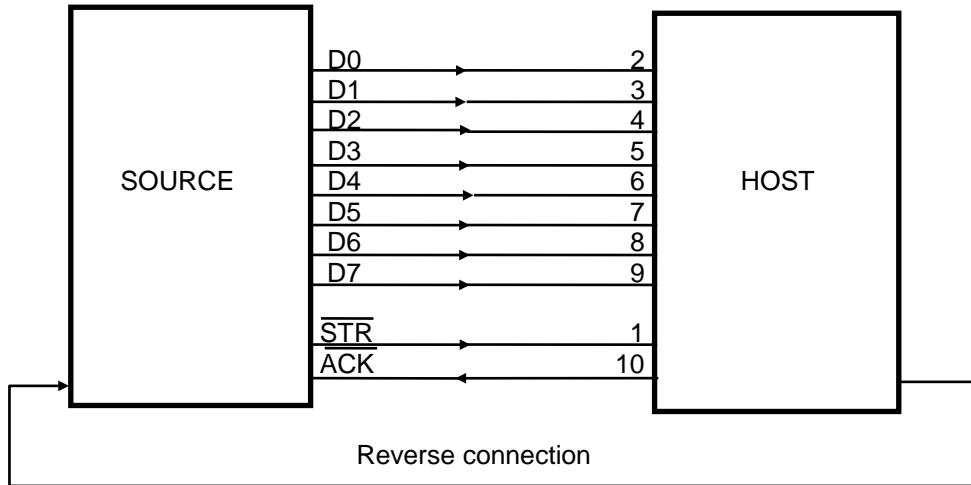


Fig. 2.4. Organization of connections for data transmission over Centronics interface

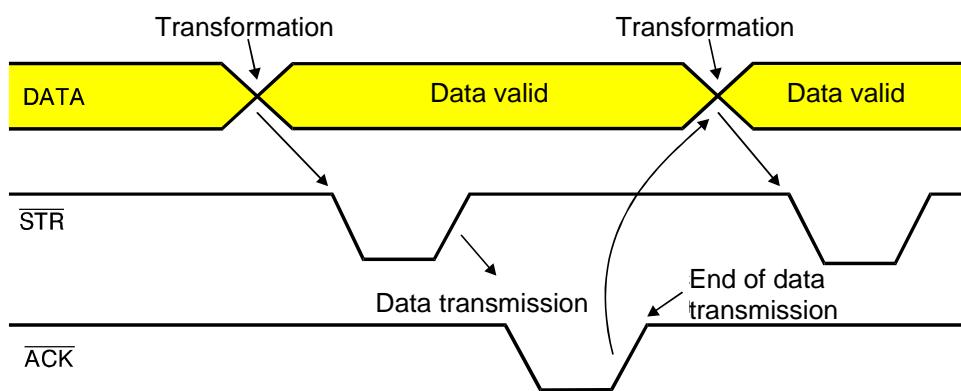


Fig. 2.5. Timing diagrams for Centronics interface data transmission

This simple system of data transmission excludes the occurrence of errors, like host switching-off, or end of paper on the printer. In order to enable data transmission at such events, the Centronics interface must be provided with additional control signals (Tabl.2.1):

1. BUSY – a signal of high level in an active state. It indicates that the data cannot be accepted. The signal is usually generated in a system with a buffer input when the buffer is filled up (pin 11).
2. PRIME – a signal of low level in an active state. It serves for initializing the host's logic scheme (pin 16).
3. PAPER END – printer's output signal of high level in an active state. It indicates end of paper or printer's button OFF position (pin 12).
4. SELECTED – a signal of high active level. It indicates that printer's button is in ON position. The alternative of this signal is DESEL (pin 13).
5. FAULT - a signal of low active level. It indicates either end of paper, or device switched off, or interface cable not connected (pin 15).
6. SELECT - a signal of low active level. It allows printer operation (pin 17).
7. AUTO LF - a signal of low active level. The printer passes to a new line after printing the current one (pin 14).

Table 2.1 lists the pins of Centronics interface with designations of a standard 25-pin connector of the personal computer.

pin	signal
1	DATA STROBE (STR)
2	data bit 1
3	data bit 2
4	data bit 3
5	data bit 4
6	data bit 5
7	data bit 6
8	data bit 7
9	data bit 8
10	ACKNOWLEDGE (ACK)
11	BUSY
12	PAPER END (PE)
13	SELECT
14	AUTO LF
15	FAULT
16	PRIME
17	SELECT PRINTER
18÷25	GROUND

Table 2.1. Correspondence between Centronics interface signals and the connector pins of a computer

The controller of Centronics type interface includes a data register, an interface control register and a status register. The registers are controlled by a decoding system that responds to three addresses (ports) of the I/O memory in a computer – two addresses for writing and reading, and one only for reading.

The data register is assigned to address 378 (278/3BC) in the I/O memory. It is accessible for reading and writing. Table 2.2 shows the register bits and their correspondence to the connector pins.

Register bit	b7	b6	b5	b4	b3	b2	b1	b0
Signal	D7	D6	D5	D4	D3	D2	D1	D0
Pin №	9	8	7	6	5	4	3	2

Table 2.2. Data register of Centronics parallel interface

At the write command from the system bus, the data that will be transferred to the peripheral device is recorded in the data register. Then the data is sent to the relevant interface pins.

At the read command from the system bus, the data recorded in the same address is read. If the peripheral device connected to the socket is capable of data transmission, from that register it is possible to read the result from the “AND” operation between the data of the register and the peripheral device.

The interface control register is assigned to address 37A (27A/3BE) in the I/O memory. Table 2.3 shows the register bits and their correspondence to connector's pins.

Register bit	b4	b3	b2	b1	b0
Signal	IRQ mask	SELECT	PRIME	AUTO LF	STR
Pin №	-	17*	16	14*	1*

Remark: '*' designates that the register of the interface status contains inversive values.

Table 2.3. Control register of Centronics interface

Bits b5÷b7 of the control register are not in use. Bit b4 is the mask for interruption from the parallel interface (IRQ) and it is not led to the connector. At read command from the system bus, the result obtained is from the AND command between b0÷b3 bits contents and the signals from the peripheral device.

The status register is accessible only to the read command from the system bus at I/O address 379 (279 / 3BD). The register stores the records of the current state of the interface signals sent by the peripheral device. Table 2.4 shows the register bits and their correspondence to connector's pins.

Register bit	b7	b6	b5	b4	b3	b2	b1	b0
Signal	BUSY	ACK	PE	SELECTED	FAULT	-	-	-
Pin №	11	10	12	13	15	-	-	-

Table 2.4. Status register of Centronics interface

2.2.1. IEEE 1284 Standard for Parallel Interface (Centronics)

The IEEE 1284 standard (Standard Signaling Method) defines a bi-directional peripheral parallel interface prevalently for connecting a computer with a printer. It specifies the physical features of a parallel port for multi-mode data transfer. The IEEE 1284 standard provides a high data capacity of the connection between a COMPUTER and a printer, or between two computing environments. The interface line length is about that of a standard printer cable, but the standard permits a twisted pair which makes the connection more reliable and faultless.

The IEEE 1284 standard specifies five modes of operation including the high speed modes EPP and ECP. Some operational modes are only for data input, others – for data output. The combination of five modes leads to the implementation of four types of ports (Table 2.5).

Parallel port type	Comment
SPP (Standard Parallel Port)	4-bit input, 8-bit output
Bi-directional	8-bit I/O
EPP (Enhanced Parallel Port)	8-bit I/O
ECP (Enhanced Capabilities Port)	8-bit I/O, uses DMA

Table 2.5. Ports defined by IEEE 1284 standard

IEEE 1284 defines also the parallel port modes which indicate the approximate data transfer speed (Table 2.6).

Parallel port mode	Direction	Data transfer speed
Nibble (4-bit)	Input only	50Kbps
Byte (8-bit)	Input only	150Kbps
Compatible	Output only	150Kbps
EPP (Enhanced Parallel Port)	Input/Output	500Kbps2.77Mbps
ECP (Enhanced Capabilities Port)	Input/Output	500Kbps2.77Mbps

Table 2.6. Data transfer speed in different modes as defined by IEEE 1284

Standard Parallel Port (SPP)

Older computers do not have some of the parallel port types. The single parallel port is used only for data transmission from the computer to devices, as the printer. Though it was not envisaged to be used as an input, a scheme is provided wherein four of the signal lines can be initialized as such. In this way the signal lines operate as an 8-bit output and a 4-bit input, ensuring a maximum output transfer of 150 Kbps and a maximum input transfer of 50 Kbps.

Enhanced Parallel Port (EPP)

Intel, Xircom and Zenith Data System developed and presented the EPP in October 1991. At present, most of the systems include a multi-mode parallel port that supports the EPP. It operates almost at a speed of an ISA bus and offers ten times higher data capacity than the conventional parallel port. It is designed particularly for peripherals which use parallel port such as net adapters, disc and tape devices, and is defined by the IEEE 1284 standard. The transfer speed is up to 2,77 Mbps [31].

The hardware protocol is described in Table 2.7. It is important to note the difference between the lines' names of SPP and EPP.

Pin	SPP Signal	EPP Signal	IN/OUT	Function
1	Strobe	Write	Out	Low level – write command, high level – read command.
2-9	Data 0-7	Data 0-7	In-Out	Data bus – bidirectional
10	Ack	Interrupt	In	Line Interrupt. Interruption at positive front.
11	Busy	Wait	In	Used at handshake. EPP cycle can begin at low level and end at high level.
12	Paper Out/End	Spare	In	Spare – not used by EEP.
13	Select	Spare	In	Spare – not used by EEP.
14	Auto Linefeed	Data Strobe	Out	Low level – indicates data transfer.
15	Error/Fault	Spare	In	Spare – not used by EEP.
16	Initialize	Reset	Out	Reset – active low level
17	Select Printer	Address Strobe	Out	Low level – indicates address transfer.
18-25	Ground	Ground	GND	Ground

Table 2.7. Lines' functionality of Enhanced Parallel Port Connector

Output lines Paper Out, Select and Error are not used in the EPP mode of command. These lines can be of any service to the user via an assignment within the SPP Status register.

EPP Handshake (negotiation)

To achieve a valid data transfer it is necessary to follow the EPP handshake between devices. For each EPP cycle the software must perform the I/O command to the relevant EPP register.

EPP data write cycle (Fig.2.6):

1. Data is recorded in the EPP data register (base +4).
2. The Write line passes to low level (low level indicates write command).
3. The data is sent to data lines 0 – 7.
4. Data Strobe line is assigned if the Wait line is on low level (beginning of cycle).

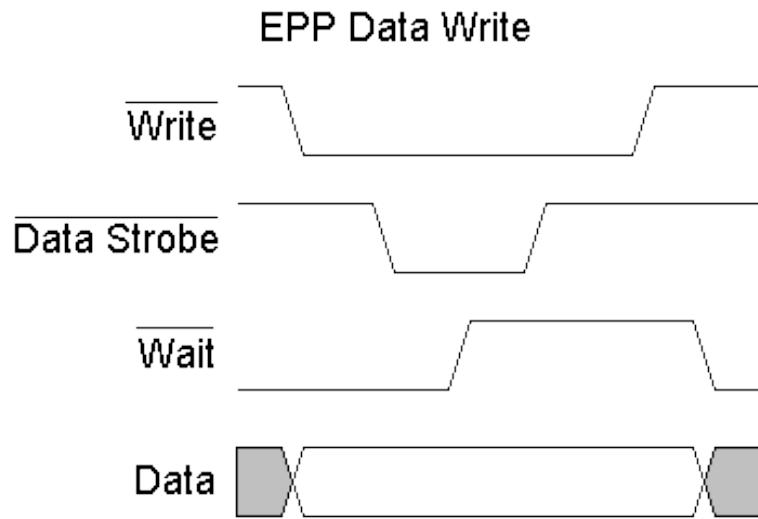


Fig 2.6. EPP Data Write Cycle

5. The host waits for acknowledgement by setting the Wait line on a high level (cycle end).
 6. Data Strobe line is set in initial status.
 7. Data Write cycle ends.
- EPP Address Write Cycle (Fig. 2.7):

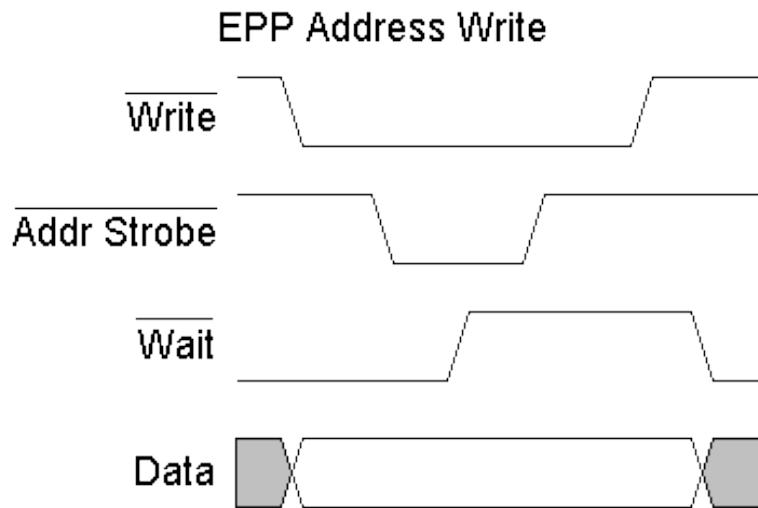


Fig 2.7. EPP Address Write cycle

1. The address is recorded in the EPP address register (base +3).
2. The Write line is on low level (Low level indicates Write command).
3. The address is assigned to the data lines 0 – 7.
4. The Address Strobe line is assigned if the Wait line is on low level (beginning of cycle).
5. The host waits for acknowledgement with the setting of the Wait line on high level (end of cycle).
6. Data Strobe line is set in initial status.
7. Address Write cycle ends.

EPP Data Read Cycle (Fig. 2.8):

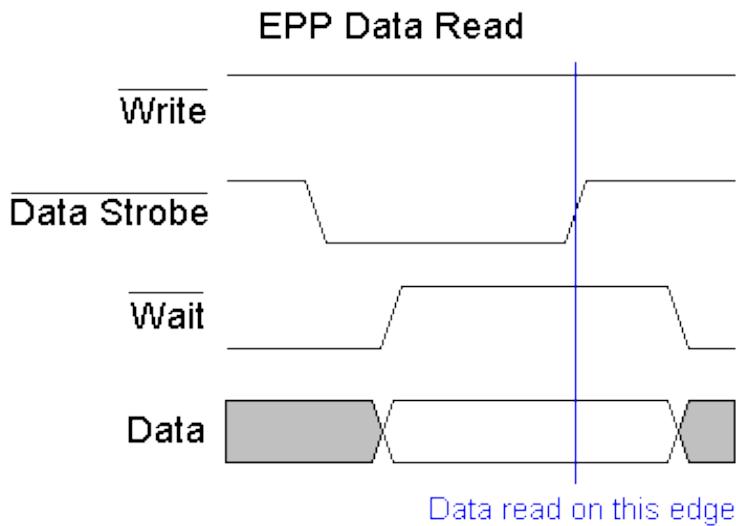


Fig 2.8. EPP Data Read cycle

1. The program reads the EPP data register (base +4).
2. The Data Strobe line is assigned if the Wait line is on low level (beginning of cycle).
3. The host waits for acknowledgement with the setting of the Wait line on high level.
4. The data is read from the parallel port lines.
5. Data Strobe line is set in initial status.
6. Data Read cycle ends.

EPP Address Read Cycle (Fig. 2.9):

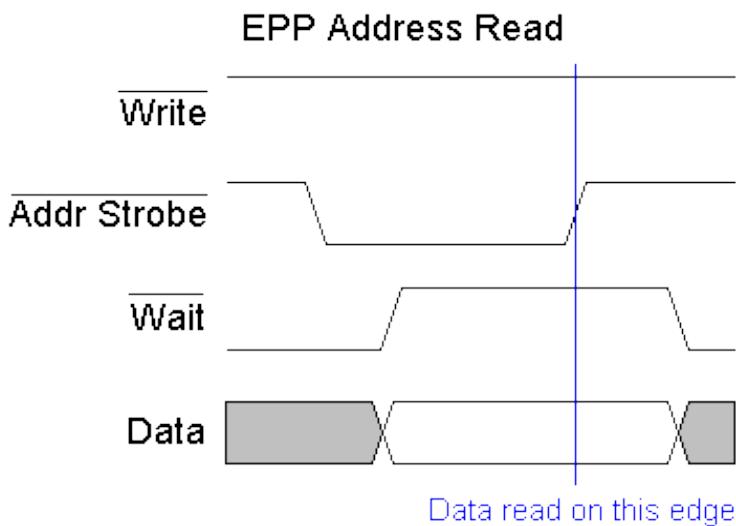


Fig 2.9. EPP address read cycle

1. The program reads the EPP address register (base +3).
2. The Address Strobe line is assigned if the Wait line is on low level (beginning of cycle).

3. The host waits for acknowledgement with the setting of the Wait line on high level (end of cycle).
4. The data is read from the parallel port lines.
5. Data Strobe line is set in initial status.
6. Address Read cycle ends.

EPP Software Registers

The EPP mode includes a new set of registers. Three of them are inherited from the standard parallel (Table 2.8).

Address	Port name	Read/Write
Base + 0	Data Port (SPP)	Write
Base + 1	Status Port (SPP)	Read
Base + 2	Control Port (SPP)	Write
Base + 3	Address Port (EPP)	Read/Write
Base + 4	Data Port (EPP)	Read/Write
Base + 5	Undefined (16/32bit Transfers)	-
Base + 6	Undefined (32bit Transfers)	-
Base + 7	Undefined (32bit Transfers)	-

Table 2.8. EPP software registers

The first three addresses are like the addresses of the standard parallel port and have the same destination. Therefore, if the EPP is used the data can be transferred via the base +0 in the same way as with the SPP.

For communicating with devices which support EPP mode, it is required to write the data in the EPP data register (base +4). Then the controller generates all necessary signals. By analogy with it, if an address is assigned to a given device, the address must be written in the EPP address register (shift +3).

Both registers (data and addresses) can be used for reading and writing, yet the EPP printer's controller initializes a read cycle after both lines – Data Strobe and Address Strobe are outputs. The device generates a reading request via interruption and the Interrupt Service Request (ISR) performs the read command.

The status register has a slight modification. Bit 0 that is reserved in the SPP set of registers becomes a bit for time-out in the EPP. The bit is assigned when the EPP time-out occurs. This happens when the Wait line does not return to the initial position for about 10 µs (depending on the port) after the lines IOW or IOR have been assigned. The IOW and IOR are I/O lines for reading and writing from the ISA bus.

The EPP mode depends on the operational time of the ISA bus. In the Read mode, the port performs adequate handshake and returns the data in the relevant ISA cycle. This operation, however, is not within one ISA cycle. In the meantime the port uses the I/O Channel Ready (IOCHRDY) of the ISA bus to show the Wait status until the cycle ends.

The three registers, base +5, base +6 and base +7 are used for 16- and 32-bit operations for reading and writing, providing the relevant port supports them. In this way the I/O operations are reduced. The parallel port transmits only 8 bits per time unit. Thus, each 32- or 16-bit-word written for the parallel port is divided into blocks with the size of a byte and then it is transmitted along the data lines.

Program Considerations

Prior to the beginning of an EPP cycle for reading or writing, the EPP data and address registers must be properly configd. In the initial status the EPP sets the Address Strobe, Data strobe, Write and Reset lines in inactive state – high level. Some of the ports require this setting before each EPP cycle. Therefore the first task is to manually initiate these lines using the SPP registers.

Enhanced Capabilities Port (ECP)

In 1992 Microsoft and Hewlett-Packard presented the Enhanced Capabilities Port (ECP) as another type of a high speed parallel port. Similar to EPP, the ECP offers improved capacity for a parallel port and requires special hardware logics.

ECP is included in the IEEE 1284. Distinctive from the EPP, the ECP is not designed to support portable devices. It supports expensive, high capacity printers and scanners. Besides, the ECP mode requires the use of a Direct Memory Access (DMA) channel that is not specified by the EPP. Hence some problem conflicts may occur with other devices in the computer configuration which use DMA (some sound cards or SCSI adapters). Most of the recent computers support both ECP and EPP modes.

The assignment of DMA channels used for ECP mode of the built-in parallel port is generally executed by the BIOS boot up program. Some older systems however may need configuration of jumpers on the motherboard itself.

Hardware Protocol

The ECP uses a standard D25 connector, as the EPP does. The ECP covers both SPP and EPP. When ECP operates like the SPP mode each line functions as shown on Table 2.7 above.

When it operates like the ECP mode the interface connector's pins have different assignment - see Table 2.9.

HostAck and PeriphAck are lines which indicate the signal type, i.e. data or commands. If the lines are in high level, that is an indication of data availability on the lines. When a command cycle is executed the lines are in low level. If the host sends commands, the HostAck line is in low level. If a peripheral sends commands, the PeriphAck line is in low level.

A command cycle is implemented when the RLE data or addresses are transmitted (Run Length Encoding – bits number). This is defined by bit 7 of the data lines (pin 9). If bit 7 is ‘zero’, then the remaining bits (from 0 to 6) are the RLE, which is used in the compressed data scheme. If bit 7 is ‘one’, then the remaining data lines transmit an address. Without the configuration of bit 7, the transmitted value is within the interval of 0-127.

Pin	SPP signal	ECP signal	IN/OUT	Function
1	Strobe	HostCLK	Out	Low level on this line indicates data valid at the host. When this pin is not assigned the high potential of the clock cycle is used for data shift to the device.
2-9	Data 0-7	Data 0-7	In/Out	Data bus, bi-directional.
10	Ack	PeriphCLK	In	Low level on this line indicates data valid at the device. When this pin is not assigned the high potential of the clock cycle is used for data shift to the host.
11	Busy	PeriphAck	In	In reverse direction the high level indicates data, the low level – command cycle. In forward direction it functions as PeriphAck (Peripheral Acknowledgement).
12	Paper Out/End	nAckReverse	In	Low level indicates reverse request for device acknowledgement.
13	Select	X-Flag	In	Additional flag.
14	Auto Linefeed	Host Ack	Out	In forward direction the high level indicates data, the low level – command cycle. In reverse direction it functions as HostAck (Host Acknowledgement).
15	Error / Fault	PeriphRequest	In	Low level assigned by the device indicates data availability from the device.
16	Initialize	nReverseRequest	Out	Low level indicates data in reverse direction.
17	Select Printer	1284 Active	Out	High level indicates Host is in 1284 mode of transmission. For interruption – it passes to low level.
18-25	Ground	Ground	GND	Ground

Table 2.9. Description of the outputs for Extended Capabilities Parallel Port.

ECP Handshake (negotiation)

ECP handshake is different from that of a standard parallel port. The most obvious difference is that the ECP has the ability to transmit data at any time in both directions which requires additional signals. The diagrams below show the ECP handshake in both directions – forward and backward [32].

ECP forward data cycle (Fig. 2.10):

1. Host sets the data to the data lines.
2. Host indicates a data cycle by setting the HostAck line.
3. Host indicates valid data by setting the HostClk low.
4. Peripheral sends its acknowledgement of valid data by setting the PeriphAck.
5. Host resets the HostClk in initial status (high) that is used to shift data to the peripheral.
6. Peripheral sends its acknowledgement of the byte by resetting the PeriphAck line in initial status.

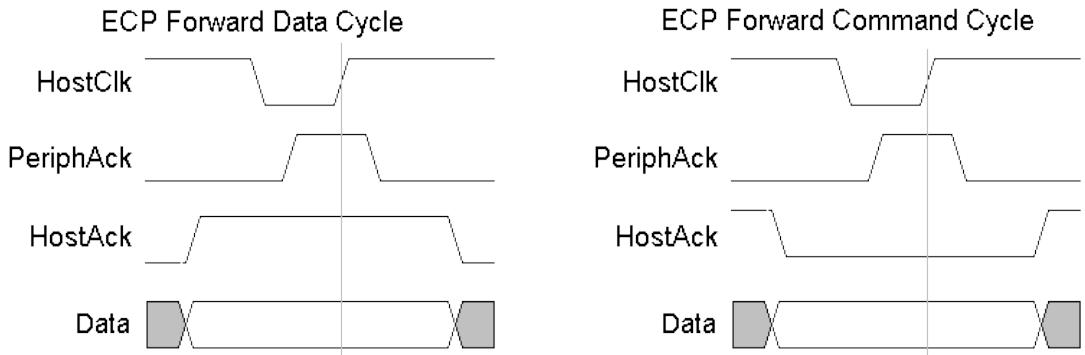


Fig 2.10. ECP forward data cycle

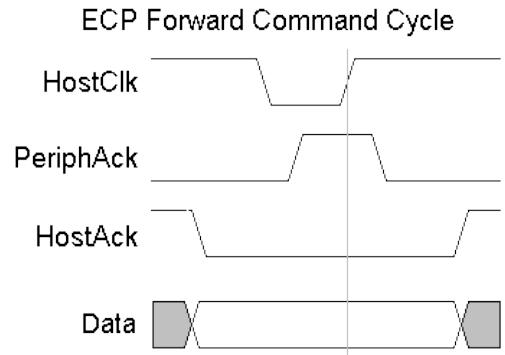


Fig 2.11. EPP forward command cycle

ECP forward command cycle (Fig.2.11):

1. The host sets the data to the data lines.
2. Host indicates the command cycle by setting the HostAck line in low level.
3. Host indicates valid data by setting the HostClk line in low level.
4. Peripheral sends its acknowledgement of valid data by asserting the PeriphAck.
5. Host resets in initial status the HostClk high that is used for data shift to the peripheral.
6. Peripheral sends its acknowledgement of the byte via resetting PeriphAck line in initial status.

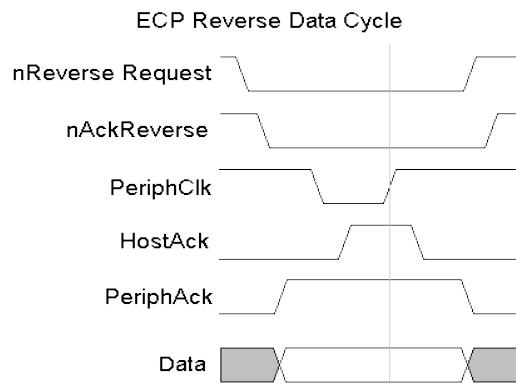


Fig 2.12. EPP reverse data cycle

ECP reverse data cycle (Fig. 2.12):

1. Host sets the ReverseRequest line in low level to request a reverse channel.
2. Peripheral acknowledges reverse channel request by setting the AckReverse line in low level.
3. Data is set to the data lines by peripheral.
4. Data cycle is selected by peripheral by setting the PeriphAck in high level.
5. Peripheral indicates valid data by setting PeriphClk signal in low level.
6. Host sends its acknowledgement of valid data via HostAck going high.

7. Peripheral device sets the line PeriphClk in high level, used to shift data to the host.

8. Host sends its acknowledgement of the byte by resetting the HostAck line in low level.

ECP reverse command cycle (Fig. 2.13):

1. Host sets ReverseRequest low to request a reverse channel.

2. Peripheral acknowledges reverse channel request by setting the AckReverse line in low level.

3. Data is placed on data lines by the peripheral.

4. Command cycle is selected by peripheral by setting the PeriphAck in high level.

5. Peripheral indicates valid data by setting PeriphClk signal in low level.

6. Host sends its acknowledgement of valid data via HostAck going high.

7. The peripheral device sets the line PeriphClk in high level, used to shift data to the host.

8. Host sends its acknowledgement of the byte by resetting the HostAck line in low level.

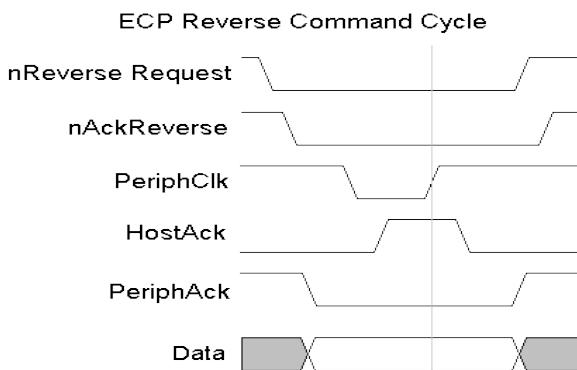


Fig 2.13. EPP reverse command cycle

ECP and SPP Handshake

At standard parallel port the handshake takes just 5 steps, while at ECP it is carried out by a longer sequence. It might be suggested that the ECP is much slower than the SPP, but practically it is not, as all steps are under the I/O hardware control. If the handshake is implemented by software then it will be executed more slowly than at standard parallel port.

Run Length Encoding (RLE)

RLE is a simple scheme for data compression used by the ECP protocol. This scheme can support a maximum compression ratio of 64:1. It functions by transmitting the run count of repetitive signal bytes, followed by one copy of the byte.

For instance, if a string of 25 "A" is to be transmitted, a run count byte of 24 is sent first, followed by the "A" byte. The receiving peripheral expands the next byte adequately to the length defined by the number of the signal bytes.

The byte indicating the length (a run count) must be distinguished from the other bytes in the data path. It is transmitted as a command to

the ECP's address FIFO port. The bytes transmitted to that register indicate either an address, or length (number of runs) depending on the status of bit 7. If bit 7 is set "1", the remaining 6 bits indicate a channel address. If bit 7 is set "0", the least significant 7 bits indicate a run length count.

ECP Software Registers

The first three registers are the same as those of the standard parallel port. It should be noted that bit 5 of the control register allows bi-directional port. This bit reflects the functioning direction of the ECP port and has an effect on bits FIFO Full and FIFO Empty of the ECP register (Table 2.10).

Address	Register name	Read/Write
Base+0	Data Port (SPP)	Write
	ECP Address FIFO (ECP MODE)	Read/Write
Base+1	Status Port (All Modes)	Read/Write
Base+2	Control Port (All Modes)	Read/Write
Base+400h	Data FIFO (Parallel Port FIFO Mode)	Read/Write
	Data FIFO (ECP Mode)	Read/Write
	Test FIFO (Test Mode)	Read/Write
	Configuration Register A (Configuration Mode)	Read/Write
Base+401h	Configuration Register B (Configuration Mode)	Read/Write
Base+402h	Extended Control Register (Used by all modes)	Read/Write

Table 2.10. ECP registers

ECP's Extended Control Register

The most important register with an ECP is the Extended Control Register (ECR), see Table 2.11. This register sets up the mode in which the ECP is to run and assigns the status of the ECP queue (FIFO).

Bit	Function
7-5	Selects Current Mode of Operation
000	Standard Mode
001	Byte Mode
010	Parallel Port FIFO Mode
011	ECP FIFO Mode
100	EPP Mode
101	Reserved
110	FIFO Test Mode
111	Configuration Mode
4	ECP Interrupt Bit
3	DMA Enable Bit

2	ECP Service Bit
1	FIFO Full
0	FIFO Empty

Table 2.11. ECR - Extended Control Register

The three most significant bits of the extended control register (ECR) define eight modes:

1. Standard mode. Activates ECP as a standard parallel port (SPP) without bi-directional functionality.

2. Byte /PS/2 mode. Behaves as a SPP in bi-directional (reverse) mode.

3. Parallel port FIFO mode. All the data written in the data queue are sent to the periphery based on the SPP handshake. The hardware generates relevant signals for the handshake. This mode is useful for devices, which do not support ECP mode. It has some of the ECP features like FIFO buffers and hardware generation of handshake, but with SPP instead of ECP handshake.

4. ECP FIFO mode. This is a standard mode for ECP and uses ECP handshake.

5. EPP mode/reserved. For some controllers this mode allows the use of EPP, for others – it is reserved.

6. Reserved. Currently reserved.

7. FIFO test mode is used for testing FIFO buffer capacity, as well as the correct running. Each byte written in the Test FIFO register (base +400h) is placed in the FIFO buffer and each byte, read by that register, is taken from the FIFO buffer. Bits FIFO Full and FIFO Empty of the Extended Control Register are used for specifying FIFO buffer capacity – normally it is of about 16 bytes.

8. Configuration mode. The two configuration registers cnfgA and cnfgB become accessible via their addresses.

The other bits of ECR bits have important role for the ECR port operation as well. Bit 4 (interrupt bit) allows the usage of interruptions; bit 3 allows Direct Memory Access (DMA). Bit 2 (ECR Service) shows whether an interrupt request is initiated. Its setting in initial status is different for different chipsets.

Bits FIFO Full (bit 1) and FIFO Empty (bit 0) show the status of the FIFO buffer. They depend on the direction indicated in the control register's bit 5. If bit 0 is assigned, the FIFO buffer is completely empty. If bit 1 is assigned, the FIFO buffer is full. If neither bit 0, nor bit 1 is assigned, the queue collects data but is still not full.

ECP's Configuration Register A (cnfgA)

The configuration register A is one of the two registers of the ECP. These registers are accessible when the ECP port is in configuration mode. CnfgA is accessible on address base +400h.

Reading the configuration register gives more comprehensive information about the ECP (Table 2.12). The most significant bit shows whether the card generates level interruption, or front interruption. This state depends on the bus type, used by the card. Bits 4 to 6 show the bus

width in the card. Some cards have only 8 bits for data; others may have 32 or 16 bits. To use the maximum card capacity, the software reads the status of these bits. It is related to the maximum word size.

The three lower-order bits are used for host recovery. For error recovery the software must know how many bytes have been sent to that instant by determining whether any bytes are still in the queue. Some versions include the byte waiting to-be-sent in the transmission register. Bit 2 of the status register shows the current situation.

Bit	Function	
7	1	Interrupts are level triggered
	0	Interrupts are edge triggered
6-4	00h	Accepts maximum 16 bit words
	01h	Accepts maximum 8 bit words
	02h	Accepts maximum 32 bit words
	03h-07h	Reserved for future expansion
3	Reserved	
2	<i>Host recovery: conveyor/ transmitted bytes, including the queue</i>	
	0	In forward direction the byte in the conveyor of the transmitter does not affect the FIFO Full bit.
	1	In forward direction the byte in the conveyor of the transmitter affects the FIFO Full bit.
1-0	<i>Host recovery: unsent byte left in the queue.</i>	
	00	Completed word
	01	1 valid byte
	10	2 valid bytes
	11	3 valid bytes

Table 2.12. Configuration register A

The output of the parallel port is only of 8 bits, but it may occur that a 16- or 32-bit word must be used. In such a case the transmission is executed in parts. Bit 0 and bit 1 give indication of the number of valid bytes left in the queue.

ECP's Configuration Register B (cnfgB)

Likewise configuration register A, the configuration register B is accessible only when the ECP port is in configuration mode. CnfgA may be accessible also on address base +401h.

The configuration register B is accessible for reading and writing. Some ports are conFigd by the software which gives the possibility to set up the number of interruption and the resource for the DMA by the register (Table 2.13).

Bit 7 of cnfgB determines whether the outgoing data must be compressed using the Run Length Encoding (RLE). When it is set the data is transmitted to the peripheral without compression. Bit 6 returns the status of the interruption line. It can be used also for checking if there is a conflict with another device.

Bits 5 - 3 show the status of the port's assigned interruption; bits 2 - 0 give the status of the DMA channel assigned. These fields are accessible for reading and writing.

Bit	Function	
7	1	Output data compression using RLE.
	0	Without data compression.
6	Interrupt status – shows the output current status for interruption request.	
5-3	<i>Selects or displays the status of interrupt request line.</i>	
	000	Interrupt selected via jumper.
	001	IRQ 7
	010	IRQ 9
	011	IRQ 10
	100	IRQ 11
	101	IRQ 14
	110	IRQ 15
	111	IRQ 5
2-0	<i>Selects or displays the status of the DMA channel used by the printer's card.</i>	
	000	Uses 8-bit DMA channel selected with a jumper.
	001	DMA Channel 1
	010	DMA Channel 2
	011	DMA Channel 3
	100	Uses 16-bit DMA channel selected with a jumper.
	101	DMA Channel 5
	110	DMA Channel 6
	111	DMA Channel 7

Table 2.13. Configuration register B

2.3. Interface IEEE 488

IEEE 488 is a standard for a universal interface bus, designed for interconnection between computers and devices like voltmeters, logic analyzers and other similar instruments (Fig. 2.14). This standard was developed by Hewlett Packard, so sometimes it is named 'HPIB'. The IEEE 488 defines how controllers and instruments communicate with each other. A device connected to the interface can perform one or several functions [12].

1. Controller. After switching-on the electric supply only one device functions as a system controller. It takes the role of an active controller and provides commands to the interface bus for sending data at random moments of time only from one device. The system controller can pass the active control to another device that is able to perform these functions.

After passing the control to another device the system controller is able to reset the bus in initial status and re-take the active control.

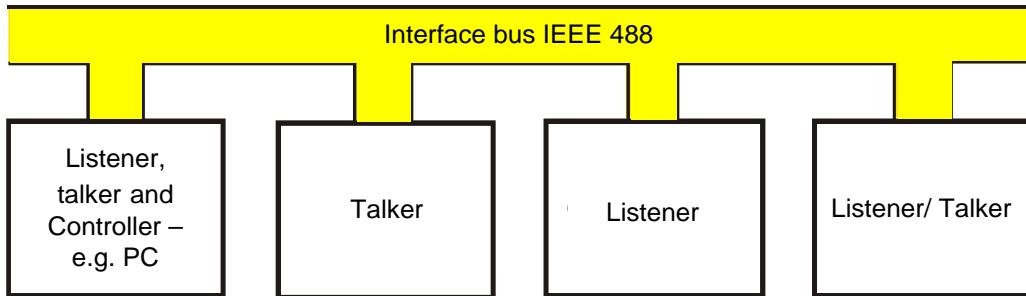


Fig 2.14. Interface bus IEEE 488

2. Talker. A device that is able to send data along the bus to other devices. At a random moment of time only one device can be activated as a talker.

3. Listener. It is the device that receives information from the Talker. Printers and other performance devices function as listeners. One or several devices connected to the bus can be initiated as listeners.

Each device connected on the bus has its specific address so that the active controller can initiate it for data transmission.

The interface bus IEEE488 consists of an eight-bit bi-directional data line, three lines for controlling the data transmission and five lines for general control signals with the help of which the information status on the data lines is indicated.

The described signals are of low active level.

Three signals are used for data control via interface and commands.

1. DAV - Data Valid. It indicates valid data on the data line.

2. NRFD - Not Ready For Data. The signal has a high potential when all listeners are ready for data acceptance (logic “AND”). The device that is not ready to accept the data generates NRFD line low and forbids the source to send the next data byte.

3. NDAC - Not Data Accepted. In high it shows that all listeners accepted the data sent by the talker.

Before data is being sent along the lines the talker must wait for NRFD to become in high logic level. This means all devices connected to the bus are ready for data acceptance. Passing of DAV to low level means data ready for reading. The high logic level of NDAC signifies that the last device has finished reading of data. HPIB handshake timing is shown in Fig.2.15.

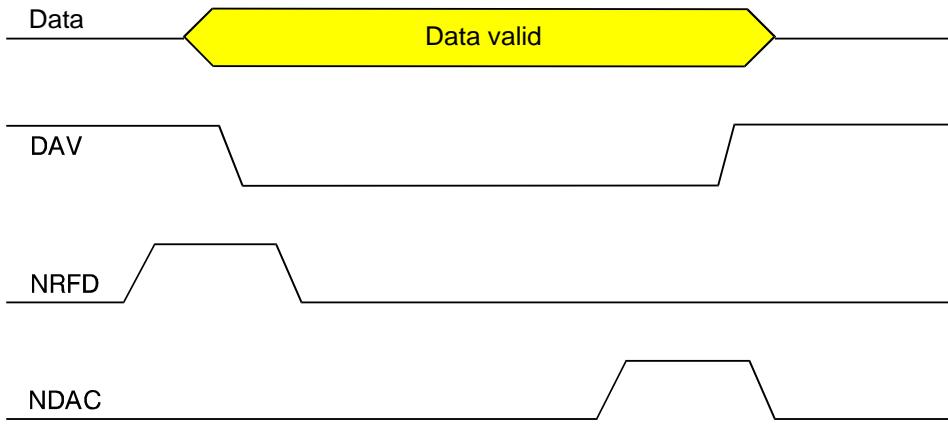


Fig 2.15. Timing diagrams of DAV, NRFD and NDAC signals

The five control lines determine and set the type of data that is sent along the eight data lines by signifying whether it is data or commands (the address is sent as a part of the command – the five least significant bits).

1. ATN (“Attention”). This line is controlled by the active controller and serves to signify the type of data on the data line. ATN low indicates command, ATN high indicates data.
2. IFC (“Interface Clear”). The system controller uses this line for bus initializing or re-taking the control.
3. SRQ (“Service Request”). A device which requests the controller’s attention activates this line low.
4. EOI (“End or Identify”). This line low indicates the end of data block sent along data line.
5. REN (“Remote Enable”). This line is controlled by the system controller and it enables the remote control of devices.

2.4. Serial Interface

The parallel format of data transmission between two devices is effective only at short distances. When a peer is far apart from a peer the cost of data transfer cables and buffering devices is raised significantly.

In contrast to the parallel interface the serial interface does not have a separate line for each parallel-transmitted bit. The information is transferred bit by bit at a time along one line.

At the start of transmission the data source (sender) loads the information to the shift register with a parallel record and then triggers the scheme of the clock generator and the counter (Fig. 2.16). At every clock pulse the data is shifted one position to the right and goes to the interface line. The receiver includes also a shift generator with a serial-in and parallel-out, a counter and a logic scheme which are controlled by the clock pulses. After the required amount of synchronization characters is registered by the counter the latter activates the data transfer from the shift register to the receiver’s buffer. This is the so called **synchronous** method of data transmission – the clock pulses are transmitted together with the data.

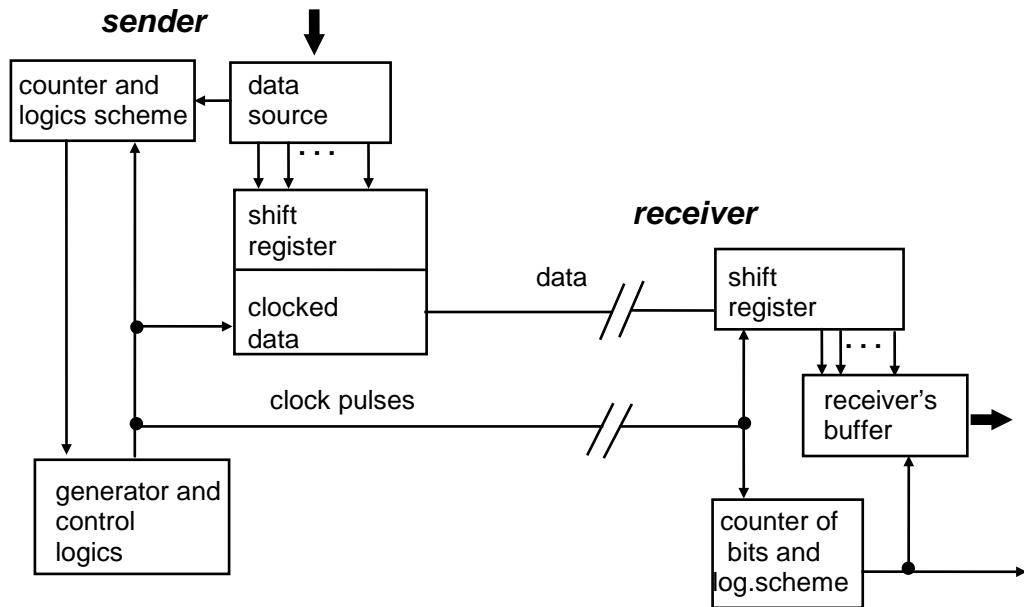


Fig 2.16. Structural scheme of serial synchronous transmission

If the clock pulse line is excluded, the system turns into a modified variant named **asynchronous** data transmission (Fig. 2.17).

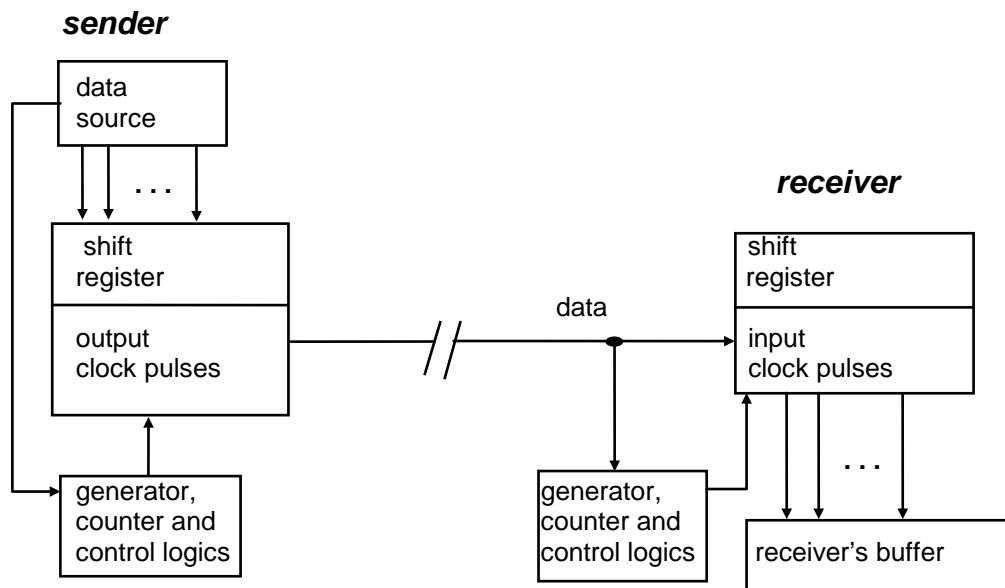


Fig 2.17. Structural scheme of serial asynchronous transmission

For implementing asynchronous serial transmission it is necessary to observe the following conditions:

- both generators must have equal frequencies;
- the receiver must be informed about the transmission start;
- there must be due time for initial setting of the receiver between two transmissions.

The first condition is fulfilled if the frequencies of transmission correspond to the standardized values which are specified as rates of transmission. The rate is measured in bauds (bits per second, baud rate).

The transmission rates are standardized. Depending on the performance of the devices, the rate of transmission is presented as follows [10]:

- 50, 75, 110, 150, 300, 600 bit/s – for teletypes and other slow electro-mechanical devices;
- 1200, 1800, 2400 bit/s – for medium-rate devices;
- 3600, 4800, 9600, 19200 bit/s – for high rate transmission.

The completion of the two above mentioned conditions for asynchronous transmission requires a standardized data transmission protocol. A standard data transmission protocol usually has a start bit followed by 7 or 8 data bits, one control bit and two stop bits.

Asynchronous Serial Transmission

At asynchronous transmission the data values are of the habitual logic “0” or “1”. It is expedient to assume that the data are transmitted in their real values.

In the beginning one start bit informs the receiver that the transmission is commenced. After that the data bit is sent at every clock pulse period in a sequence from the least significant to most significant bits. A parity bit is added if it is required to carry out a parity check in the process of transmission, and its value is determined by the data bits values. The parity check is sent after the most significant bit.

A time interval follows each transmission in order to reset the receiver in the initial state prior to the new-coming successor. This time is the stop interval that is implemented in different systems either as 1; 1,5 or 2 stop bits having the value of logic “1” (Fig. 2.18).

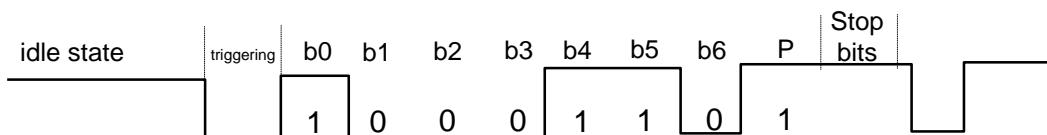


Fig 2.18. Timing diagram of asynchronous serial transmission

At asynchronous data transmission the events occur in the following sequence:

- idle state – logic “1”;
- one start bit with the value of logic “0”;
- 7 or 8 data bits starting form the least significant;
- one parity bit;
- 1; 1,5 or 2 stop bits with the value of logic “1” corresponding to the idle state;
- next start bit with the value of logic “0”.

The block of characters shown in Fig. 2.18 consists of seven data bits 0110001 or together with the parity bit it becomes 10110001. After the transmission of the eight bits, two stop bits follow, and immediately after the stop interval the new-coming transmission starts.

Most frequently the stop interval is composed of one bit. Two bits are used for the electro-mechanical teletypes as they need more time for the initial reset. In that case the total number of bits in the character block is 10, i.e. 7 are data bits, 1 parity bit and 2 stop bits. At transmission rate

of 1200 baud rate (bits/sec), the maximum number of data characters sent per second is 120.

2.5. Standards for Serial Transmission (RS232, V24, RS423)

Most widespread interface standards are based on the involvement of a special voltage source positioned at the side of the sender [10]. The logic signals have the following meanings:

- 1 – negative voltage;
- 0 – positive voltage.

Standard RS232 defines less the signal levels than the numbers of connector pinouts used for signal transmission between devices. Modem connections (modulator-demodulator) on a telephone line are also specified.

Standard V24 was developed by the Consultative Committee for International Telegraphy and Telephony (CCITT). It is analogous to the American RS 232 used in the European industry.

According to RS232 the logic “1” is defined as the voltage from -3 to -25V, and the logic “0” within the range of +3 to +25V. The standard assumes the usage of a 25-pin connector, each pin-out being relevantly defined.

RS232 is not designed for communication at long distances. The maximum distance apart is 20 meters. This standard involves the usage of modems which allow data transmission at long distances on telephone channels of general or specific purposes. Usually the transmission is in the range of 300 to 9200 baud rate.

The main signal circuits used in RS232 are [10, 12]:

- TxD (Transmitted Data) – data is transmitted;
- RxD (Received Data) – data is received;
- RTS (Request to Send) – request for transmission from one device to another;
- CTS (Clear to Send) – signal to the device indicating that the receiver is ready to accept the transmission;
- DSR (Data Set Ready) – the device is ready for transmission;
- GND (Signal Ground) – signal for the ground;
- CD (Carrier Detected) – indicates detection of the carrying frequency of the modem – detector;
- DTR (Data Terminal Ready) – signal to the device for setting up the connection;
- RI (Ring Indicator) – used at operating on commutating line as a call sign.

The simplest scheme of transmission between two devices is shown in Fig.2.19. The output circuit of the sender, as defined by the standard, must be within the range of +/-12 V and its input signal - standard logic levels.

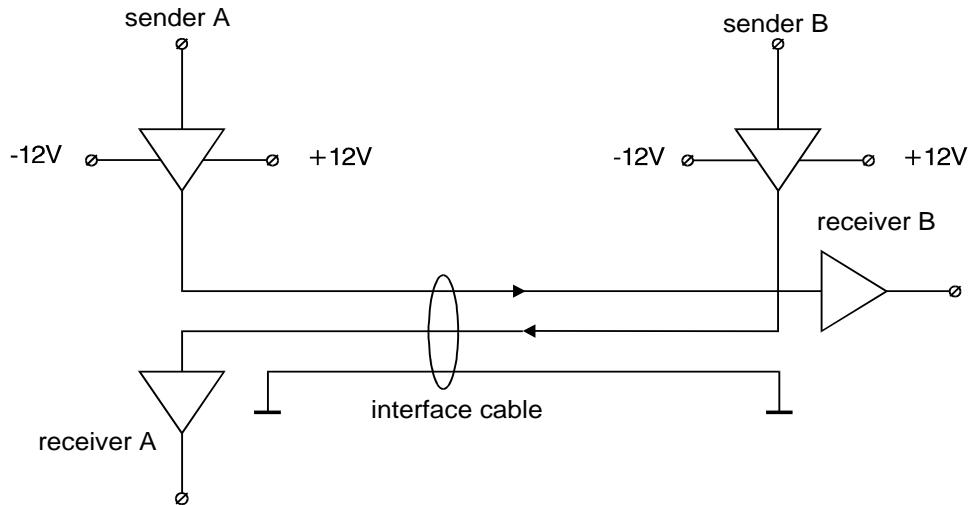


Fig 2.19. Connecting diagram at serial transmission

Standard RS423 uses lower voltage on the transmission line (e.g. -5V) and allows high rates up to 100Kbps at a distance of 10 m. If far apart, for instance 1300 m, the baud rate is 1 Kbps. Compared to RS232, the standard RS423 allows servicing of several devices (max.10) via a single transmission line.

2.6. Current Loop Interface

The interface line contains two conductors which form a circuit including a current loop generator and a receiver (Fig. 2.20).

When the current exceeds 17mA, logic "1" is transmitted; when it is below 2mA, the logic level is "0".

This interface is widely incorporated in industrial systems as it ensures a reliable connection at long distances without a modem.

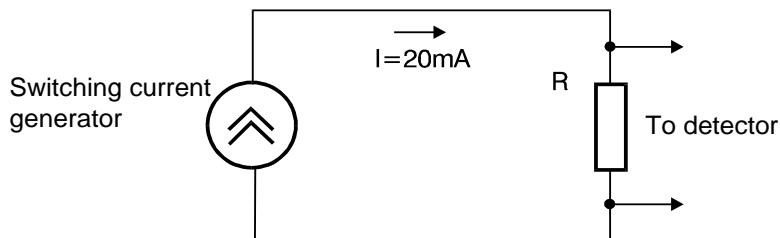


Fig 2.20. Representation of a current loop interface

Fig. 2.21 shows the simplified line diagram of a current loop transmission. It implements also a galvanic disconnection between devices.

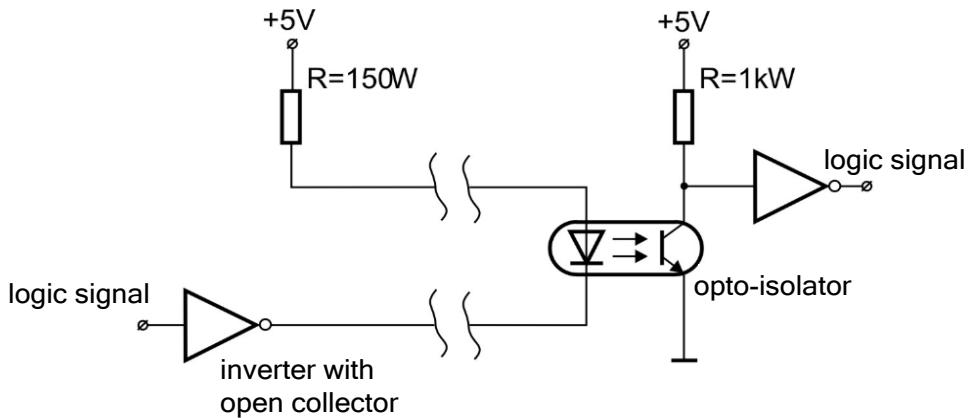


Fig 2.21. Simplified line diagram of a current loop transmission

2.7. Standards for Differential Input and Output (RS-422 and RS-485)

Fig. 2.22 shows the connection of two devices in accordance with standard RS-422.

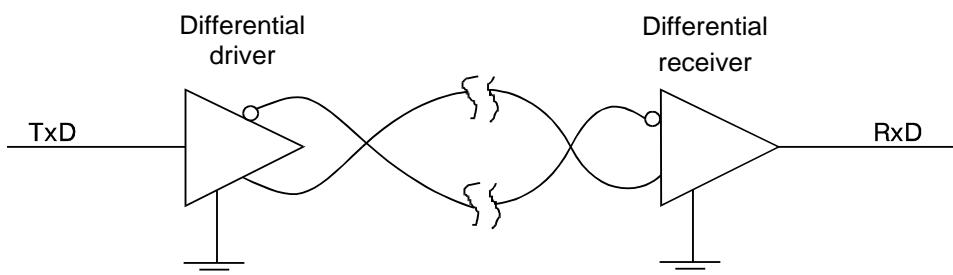


Fig 2.22. Connection diagram acc.to standard RS-422

This standard is characterized by the highest noise-resistance, the longest connection lines and the highest transmission speed compared to the above described standards for serial transmission [10, 12]. It specifies drivers with differential outputs and receivers with differential inputs as MC3486 and MC3487 of Motorola or 9636 and 9637 of Texas Instruments. Each signal is transmitted over a separate twisted pair of conductors. Thus the sinphase disturbances are reduced. The connection lines' length may reach up to 1300 m, and the transmission speed - up to 10 Mbps.

Standard RS-485 has the same features as RS-422 but it is applied to asynchronous data transmission, i.e. an additional twisted pair is added for transmitting the synchronizing clock pulse. The approach is reliable also for data transfer in different technological environments. It is widely applied by Siemens for communication between Simatik industrial controllers.

2.8. Universal Serial Bus (USB)

In 1995 Compaq, Digital, IBM, Intel, Microsoft, NEC and Northern Telecom companies united to make a decision concerning the development of an improved interface. The same year a forum for USB was commenced and in 1996 the new release of USB interface was announced. It eliminated the disadvantages of both serial and parallel interfaces. In order to achieve higher efficiency of communication, the USB was designed with a data transfer rates of 12 Mbps (and an alternative for low speed of 1,5 Mbps). The companies developed a strictly defined cabling system with one special cable that covered all types of connections and allowed the possibility more peripheral devices (up to 127) to be connected to one PC port. Furthermore, Plug and Play support was incorporated so that each connection could be self-controlled. While working, the system allows inclusion of new devices (hot plug) and their activation without re-loading the operation system [33].

On May, 27th 2000, an improved version USB 2.0 of the standard was released by a group of designers from Compaq, Hewlett-Packard, Intel, Microsoft, NEC and Philips. The key modification concerns the higher efficiency in result of the increased speed from 12 Mbps to 480 Mbps. The new system is completely compatible with the old version and includes all its protocols. The devices involved in the transmission negotiate up to the highest common speed and use it for the data transfer. Cables and connectors remain unchanged.

Environment

The USB divides the serial hardware into two classes: hubs and functions. The USB hub provides the sockets for plugging the functions (devices). The USB function is a device intended to execute a certain operation. According to USB designers the function is any device that can be plugged to the computer incl. keyboard, mouse, modem, printer, plotter, scanner, etc.

The USB sooner performs the function of a bus that allows a set of peripherals to be connected to one socket of a computer using one and the same signals to all devices, rather than being a peer-to-peer communication. The data passes through the bus in the form of packets that all devices accept. The computer assigns a concrete device by adding a specific address to the packet, so the device becomes the only one that can read the addressed packet.

For connecting more than one device to a computer it must be provided with due sockets. Modern computers have 2 – 10 ports. There is a possibility of adding a multi-port hub as well as to implement a hub-to-hub connection that will increase the number of hubs.

The USB technology represents a hierarchical system with hubs connected to hubs, which are connected to hubs, which are connected to hubs, etc. That allows the hub to have lots of connections simulating a tree structure. The conceptual model of a USB system is shown in Fig.2.23.

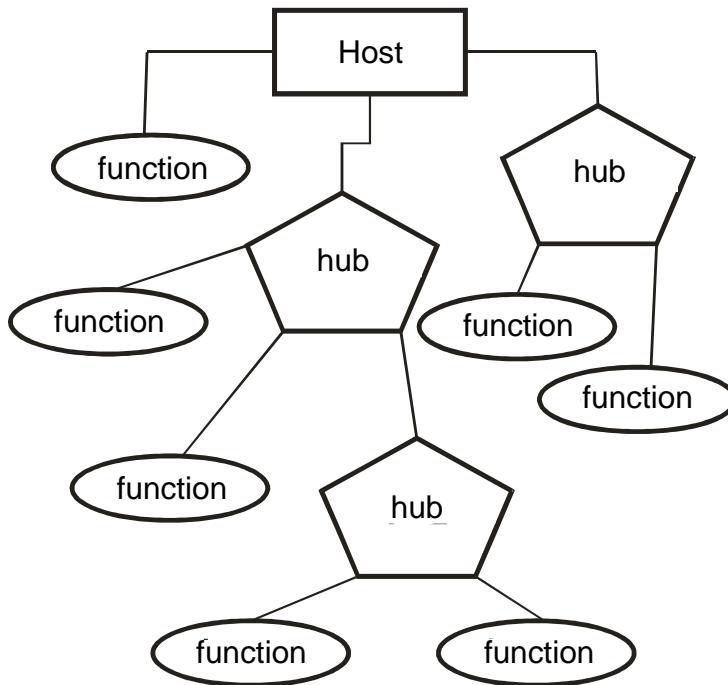


Fig. 2.23. Hierarchical USB connection diagram

Host, hub, function

A computer acts as a main hub for the USB system, and as such, it is implied as 'host'. The circuit in the computer that controls the integrated hub and the rest of the USB system is called a 'bus controller'. A USB system has only one bus controller.

The USB 2.0 standard specifies three types of speeds for devices: Low speed (1,5 Mbps), Full speed (12 Mbps) and High speed (480 Mbps).

USB 2.0 is compatible with USB 1.1 standard, i.e. all USB 1.1 devices work with USB 2.0 devices and vice versa.

Except for speed limitations, it is of no importance to the USB system which device is connected to whichever hub or how many levels below in the hierarchical structure a device is connected. The USB systems require all devices to be connected in the right way according to the rule 'a device connected to a hub'. Then the USB software performs all the rest. The software builds the USB protocol which in fact is the most sophisticated part of the designing process. In contrast, the hardware itself is simple but it does not work without protocol.

On the other hand, the cablings set limits for the distance at which a device can be positioned. The maximum length of a USB cable is 5 m. But the distance can be extended by means of hubs.

As part of the Plug and Play process, when a computer is started, the USB controller begins to scan all devices. It queries each of them about the purpose of identification, and then builds a map which locates them by hub and number of its port. This identification becomes part of the address in the packet. When the USB driver sends data to a port, it routes the data to the respective device, to the hub and the port number.

The USB requires the availability of special software support. Any device with a USB connector is provided with firm's software for controlling

the USB bus. A computer needs to have relevant software to make the USB system function. The operation system must know how to send adequate signals to the corresponding USB ports. All Windows versions, starting with Windows 98, support USB. Windows 95 and Windows NT do not support USB. Each USB device connected to a computer asks for the installation of software in the process of hardware connection. The device driver generates commands or data packets. As a whole, the USB driver acts as a service for deliveries ensuring a channel for the data flow.

Connectors

A USB system includes 4 kinds of connectors: two kinds of installation sockets and two kinds of cable plugs. Each socket and plug is completed in two variants A and B (Fig. 2.24).

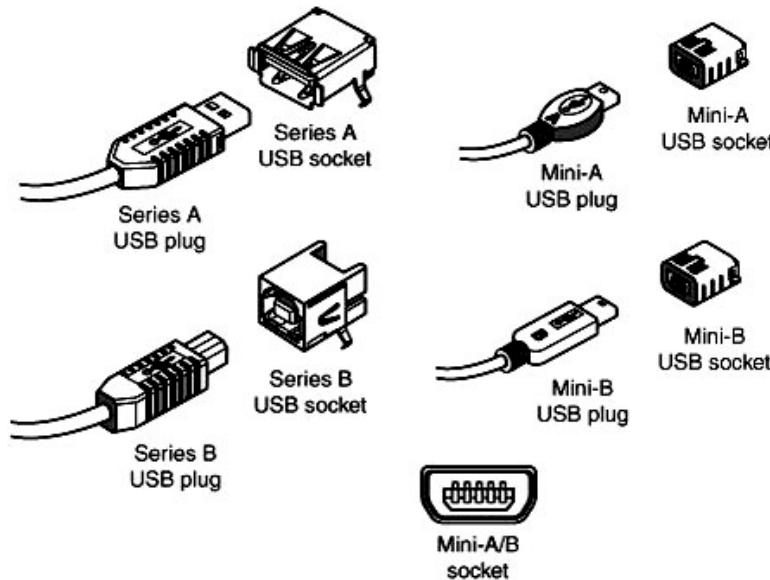


Fig. 2.24. Connectors of USB interface

The hubs are provided with sockets of type A (on a computer, as well) which are the basic external plugs for the USB ports. They can be wide and thin USB slots positioned on the rear/front side of a computer.

In order to adapt devices to cables the USB standard allows a different type of plug and socket for a device, unlike a hub. If there is no cable in the device, the latter uses a USB 'B socket' which is a small opening with an almost square shape suitable for plugs of type B.

The motivation of making connectors more complex aims to avoid confusion. All USB cables have at one end a connector of type 'A', and at the other end – of type 'B'. One end is inserted in the hub, the other – in the device.

Cables

The USB system uses two types of cables: for 'Low speed', and 'Full and High speed' connections. Both cable types have a special type of 4-wire conductor. Two wires are intended for data transmission with differential signals. The other two USB wires are for rated voltage of $U_{cc} =$

5V and ground which allows the power supply of external devices. The two data wires form a twisted pair while the power wires may be of another type.

The difference between Low and High speed cables is that the capacitance of the high speed cable is measured to cover relevant requirements, while the standard does not require a twisted pair for low speed cable.

Each conductor in the USB cable is of specific colour. Data wires form a twisted pair of white ('-Data') and green ('+Data') colour. The Power 5V-wire is red. The Ground-wire is black (Table 2.14).

Signal	Colour
+Data	Green
-Data	White
+U _{cc}	Red
Ground	Black

Table 2.14. USB signals

Two computers cannot be connected by USB. The standard defines only one USB controller in the entire system. Physically, the cabling prevents from such a mistake. The exception, when a cable has a built-in bridge, the latter allows communication between two computers. The bridge is the active part of a circuit that transforms the signals.

2.9. Interface IEEE 1394

The IEEE 1394 interface is also known by the brand name of 'FireWire'. It is a serial interface designed for devices which require continuous guaranteed bandwidth, like CDs, tape devices, multimedia periphery (digital cameras, digital video and digital TV). In the beginning it was thought to be a general purpose interface convenient to replace the serial ports but with higher transmission rate [33].

The FireWire standard is a system for peer-to-peer communication in contrast to USB standard. The FireWire device is plugged in directly to the computer port. In order to connect more devices the computer should have more sockets, or the FireWire devices should be provided with input and output ports which allow serial communication. Unlike USB, this standard does not use hubs in its network.

According to the FireWire interface topology a single port device is so-called 'leaf', a 2 port device is 'transition' and 3 port device is 'branch' or 'hub'. Transition devices and hubs operate as repeaters, re-generating the signal to the next node. Each FireWire system has one main node. It is the root over which the system is organized. The connection sequence can reach up to 16 hops (device to device). If more, the signals delay is increased and goes beyond the standard requirements. When lots of devices should be connected it requires the usage of branches for creating parallel paths.

The current 1394a standard allows a cable length of up to 4,5 m. In case of 16 serial connections the maximum distance between the first and the last device is 72 m.

Each FireWire cable consists of two active connections for a complete duplex. Connectors at the cable ends are the same, so cabling is easily implemented. The software takes care of all details of the communication. Standard 1394a also allows the usage of minimized jacks which are convenient for limited space.

The FireWire protocol uses 64-bit addressing:

- 6 bits for identifying the devices;
- 10 bits for network identifiers;
- 48 bits for an address in the memory.

Signals

For noise reduction FireWire uses two pairs of differential signals. One pair transmits the actual data. The second pair called 'strob line' presents the status of the data pair, so that only one of the pairs changes its polarity at each clock pulse. For instance, if the data line transmits two successive bits of equal value, the strobbing line converts the polarity in order to clock the sequence between them. If the sequence of two bits with different values of the data line is logic '1' followed by logic '0' or vice versa, the strobbing line would not change its polarity. The summing of the data line with the strobbing line represents precisely the clock signal of the sender system, allowing a clear understanding between the sender and receiver device.

FireWire operates as a bi-directional channel for data sending and receiving by means of two pairs of wires. When one pair sends the data the other acts as a strobe signal. At data receipt the pair used for strobe becomes a data sending line, while the other one transmits the strobe signal. In other words, when the device passes from sending to receiving mode it converts the pairs of wires used for data and strobe signals.

Configuration

FireWire allows connecting lots of devices and uses an address system so that the signals sent through the common channel can be identified only by the matching target device. Devices, which are connected, can communicate independently without the inclusion of a computer. Each device communicates at its own speed. A FireWire connection scans the speeds so as to choose the most appropriate one for the transmission. Certainly, the low-speed devices pass through the signals of the high-speed devices, but the vice versa is not possible. Therefore it is necessary to take measures in the process of system design so that all devices can operate at optimum speed.

FireWire eliminates the necessity of assigning a unique identifier to the device since the configuration process is automatic. When a device is added to the system, or it is started, the system detects it and the automatic process of configuration begins. Each device determines its position within the system, i.e. root, branch, transition or leaf. After the hierarchy is built up the FireWire devices determine their unique identification numbers depending on their location in the hierarchy and

send them to the root node. When a new device is added the change starts up an initial configuration and all devices are re-addressed.

Arbitration

FireWire sends the data in packets. Each data block is preceded by a heading that determines the data flow pipelines and their priority. Each device in the FireWire base system, which shares the connection, has the chance to send one packet within the so-called arbitration period that is also named a ‘fair interval’. Different devices have the possibility to equally use the bus during that time. If no device has started data transmission while the interval is passing, then all devices wait longer and extend the time for an arbitration reset. After that a new fair interval begins and all devices can send one more data packet. This cycle continues until data transmission in the system ends.

In order to manage the devices which need a continuous data flow for playing of video or sound signals in real time mode, the FireWire uses a special isochronous (simultaneous) mode which determines the necessary time of transmission. An interval of 125 micro-seconds is generated for the other devices within which each device takes its turn according to its priority, leaving a short delay between the packet transfer.

The next FireWire 1394b release has a new arbitration system called BOSS (Bus Owner/Supervisor/Selector). The last device, called BOSS, which confirms the receipt of the packet, also gets the specific rights for the system control. A BOSS device takes the full control over the tree and can select the next node as a BOSS.

Connectors

The presence of only one small 6-pin connector for all possible connections with identical plugs at both ends is typical of the initial FireWire release. All ports are equivalent, and the outputs are arranged in two parallel rows at the inner side of the connector. The asymmetrical D-shape of the active end prevents it from a wrong plugging-in (Fig. 2.25).

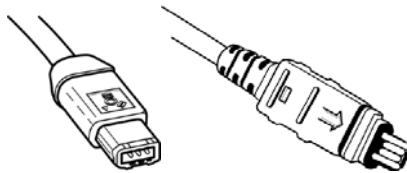


Fig 2.25. FireWire connectors

The amendment of standard 1394a added a minimized connector. To keep it compact the two power pin-outs were removed. The new jack is convenient for small electronic devices like camcorders, digital cameras, etc.

The 1394b standard adds two more connectors to the FireWire stock. Beta jack is designed for systems which use only the new 1394b signal system and are incompatible with standard’s earlier versions. The bilingual jack is intended for devices which use the new and the old FireWire standard (Fig. 2.26).

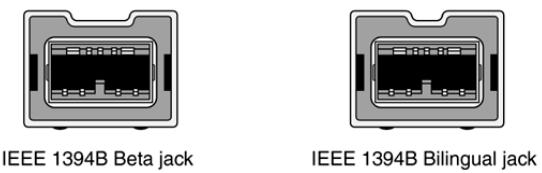


Fig 2. 26. 1394b FireWire jacks

Cables

In the current format the FireWire uses a special design. Two versions are allowed – one with 4 signal wires and the other with 4 signal wires and 2 power supply lines. In both versions the data passes through shielded twisted pairs with rated impedance of 110Ω . The second version provides power supply of 8 to 33 V with 1,5 A for supplying the peripheral devices.

The 1394b standard allows two more connection forms with optical fibers – glass and plastic as well as a standard twisted pair category 5. The maximum length of the plastic optical fiber connection is 50 m; for the glass optical fiber it is 100 m. The transmission rates of these connections are 100 or 200 Mbps. Cable category 5 allows a length of up to 100 m but only at low speed of max 10 Mbps.

All FireWire cables are twisted pairs. Signals which appear on pin 1 and pin 2 at one end of the cable cross on pin 3 and pin 4 at the other end. This allows all FireWire ports to be provided with cabling in one and the same way. The same connectors can be used at both cable ends and concordance as for USB is not required.

2.10. Infrared Data Association (IrDA)

Basic parameters

The originally designed IrDa standard is so configd that it can replace the serial cables and make data transmission technologically easy and inexpensive. The first IrDA version uses asynchronous communication in the same format and standard data transfer rates, as the serial port does.

To keep the consumed power low and to prevent from interference the devices which are placed in one room, the IrDA has a short-range of data exchange, i.e. about 1 meter.

Besides the standard serial port, the IrDA can also replace the parallel interface with a specified transfer rate up to 4 Mbps. The data exchange uses a packet-based synchronization which requires a special controller. The latter tracks and manages the data flow between the computer bus and the communication buffers.

Gradually the IrDA systems were divided into slow and fast systems. The faster standard is so designed that it is fully compatible with the older version. Fast devices can communicate with the slow ones at

lower rates, but the slow devices cannot use high rates unless they are improved.

IrDA standard defines not only the hardware aspect but the data format used in the systems, as well. Six standards have been published which cover these aspects of IrDA communication. The hardware itself determines the so-called physical layer. IrDA defines the protocol for link access, named IrLAP, as well as the protocol for link management – IrMP which specifies the data format used for handshake (negotiation) and communication support. All IrDA systems must adhere to the standards. For easier integration of the system into modern computers a Transport Protocol is defined as well as an optional extension “Plug and Use”. The IrCOMM group of standards specifies the standard way for emulation of conventional computer's serial and parallel ports [34].

IrDA specifications are based on data exchange in the close infrared spectrum with wave length between 750 nm and 900 nm. The frequency range (wave length) has been narrowed aiming to ensure protection from parasite signals coming from external light sources.

The infrared ports are characterized by the following parameters (see Table 2.15).

Baud rate	9,6 bit/sec ÷ 115,2 kbit/sec
Wave length	850 ÷ 900 nm
Maximum pulse length	1,41 μ s
Maximal bit length	3/16 of the bit time of RS 232

Table 2.15. Parameters of IrDA ports

The angle of infrared rays' spread spectrum can be max. $\pm 30^\circ$, which determines a comparatively small sight cone of the wireless devices that use the standard (Fig.2.27).

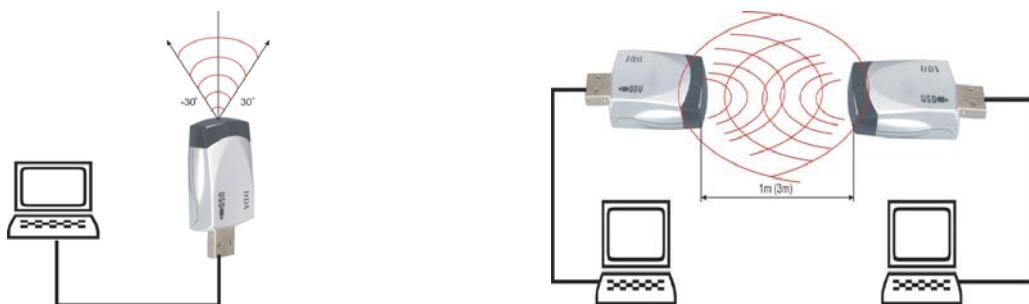


Fig 2. 27. Angle of IrDA rays spread spectrum

Fig 2. 28. IrDA communication between two devices

Two devices, which use an infrared channel for data transfer, are able to communicate at a maximum distance of 1m to 3 m depending on the type of the device (Fig.2.28).

Transmission rates

IrDA specification allows the usage of transmission rates well known for the serial port, i.e. from 2400 to 115200 bps. All these rates use the default modulation circuit Return to Zero Inverted (RZI). The high speed version adds three more rates, i.e. 576Kbps; 1,152 Mbps and 4,0 Mbps which are based on the Pulse Position Modulation (PPM) circuit.

Irrespective of the speed which devices can support, the IrDA users at the first place establish a connection at speed of 9600 bps by means of the access protocol. After the connection is set up the devices switch over and start using the rate suitable for both of them.

Pulse width

The infrared cell of IrDA transmitter sends the data in pulses each of them a tiny part of the main clock period or bit. The comparatively big distance between the pulses makes them more easily distinguishable for the optical receiver [33].

The length of each infrared pulse is at least 1,41 µs for transmission rates up to 115200 bps. The information IrDA pulse must be at least 3/16 of the transmission time of the information bit, in contrast to RS 232 standard. Yet, a pulse length slightly over 10% is also acceptable. For instance, every bit at 9600 bps has a duration of 104,2 µs (it is 1s divided into 9600). The typical infrared pulse for such speed is 3/16 of that period or 19,53 µs.

At higher rates the minimum pulse length lessens to 296,2 ns at 576 Kbps, and to 115 ns at 4 Mbps. The interconnection between rates and pulse lengths is shown in Table 2.16.

Rate	Modulation	Pulse length
2,4Kbps	RZI	78,13 µs
9,6Kbps	RZI	19,53 µs
19,2Kbps	RZI	9,77 µs
38,4Kbps	RZI	4,88 µs
57,6Kbps	RZI	3,26 µs
115,2Kbps	RZI	1,63 µs
0,576Mbps	RZI	434,0 ns
1,152Mbps	RZI	217,0 ns
4,0Mbps	4PPM, single pulse	125 ns
4,0Mbps	4PPM, double pulse	250,0 ns

Table 2.16. IrDA pulse rates and lengths

Modulation

Depending on the rate at which the connection is functioning, the system can use one of the two modulation forms. At rates lower than 4 Mbps the system uses RZI (Return to Zero Inverted) modulation.

At rates of 4 Mbps the IrDA system passes to Pulse Position Modulation (PPM) which uses four separate pulse positions and is designated with 4PPM.

RTZ – Return-to-Zero

For keeping the right synchronization most of the digital signal lines operate by switching every bit upwards and downwards at every time interval. For transmitting the logic “1” the line switches at high level and then again gets a zero potential before the time interval is completed. The encoding is known as “Return-to-Zero”.

RZI - Return-to-Zero Inverted

The only difference from the RTZ method is the inversion, which means that the pulse from the voltage encodes digital “zero” instead of digital “one”. The line’s potential is constant with the exception when it is interrupted by the data pulse.

PPM - Pulse Position Modulation

It uses the pulse interim position in the clock interval to show a discrete value. For instance, one clock period length can be divided into four equal segments. The pulse appears only in one segment and encodes a definite value as shown in Fig. 2.29.

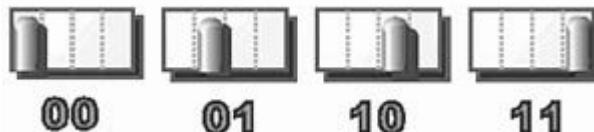


Fig 2.29. Pulse positions for the four valid values of 4PPM

IrDA requires data transmission only in 8-bit format. In the context of the serial interface terms, the data format at IrDA is determined in the following way: one start bit, 8 data bits, one stop bit, i.e. 10 bits in total for one symbol. Regardless of the modulation used, all bytes are transmitted starting with the least significant bit.

Format

IrDA systems do not operate on level bits or bytes transmission, but use packets of bytes, the so-called frames. One frame consists of 5 to 2010 bytes. A frame includes: address information, data, and error correction. The frame format is determined by the link access protocol.

Interference suppression

High-speed systems automatically stop low-speed devices which operate in the same environment. Thus they avoid the affects of

interference. Transmission of low-speed systems is stopped by high-speed systems which send a special Serial Infrared Interaction Pulse (SIP) at intervals not longer than half a second. SIP has a duration of 1,6 μ s followed by a pause of 7,1 μ s. The parameters are the same as those of a packet starting pulse. Meeting a SIP, the low-speed system thinks that it a starting pulse and automatically begins to track of low-speed data, suppressing its own transmission for half a second. Before getting the chance to start the transmission of its own data, another SIP puts to sleep the low-speed system for another half a second.

While in mobile devices the infrared ports are commonly incorporated in their cases, PCs use the so-called infrared adapters which are switched via serial, parallel or USB ports (Fig.2.30).

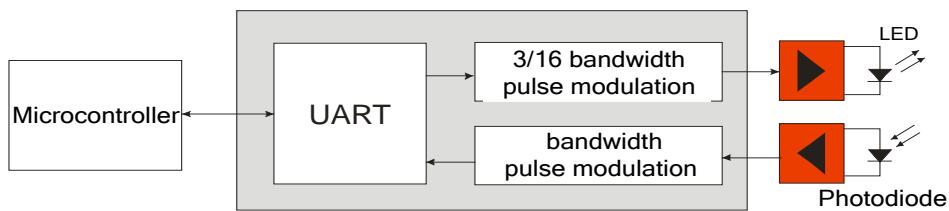


Fig 2.30. IrDA adapter for connection with a PC

In transmission mode the data from a computer goes through parallel channels to Universal Asynchronous Receiver/Transmitter (UART). There, it is multiplexed and sent by bits to a modulator that forms the pulse length. After the modulator the data is sent to an infrared diode. The active level is logic "0" (Fig.2.31).

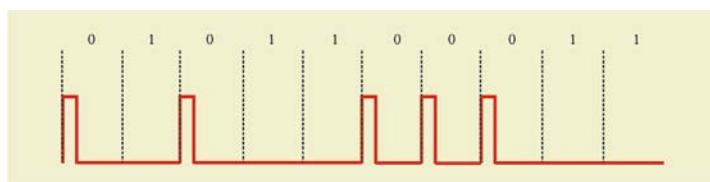


Fig 2.31. Data format transmitted by the IrDA adapter

The received data goes to a light sensitive element (photodiode or phototransistor) which re-sends it to a de-modulator. Then it is transferred to UART which de-multiplexes and sends it to the computer over parallel channels.

The data exchange protocol contains one start bit, eight data bits and one stop bit. The control bit is omitted. In normal state the port is in logic "1".

Data transmission through the infrared port is analogous to the data transfer via serial port and that is why the introduced interface R232 can be duly applied.

Application

The infrared communication is used mostly in mobile device. It saves costs for additional cable equipping. Besides, wireless technologies

provide the possibility of designing network architectures of several users (Fig.2.32).

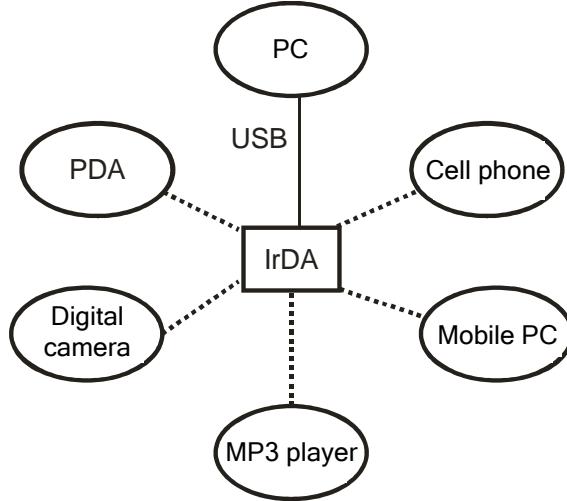


Fig 2.32. IrDA applications

IrDA is used for connecting PCs to different peripherals or other PCs.

The wide application of infrared communication imposed the creation of IrDA. Owing to this standard many operation systems contain drivers for infrared port.

2.11. Bluetooth

Bluetooth is a wireless packet-based communication radio system which allows sharing of data by many devices in a small network. This interface can be connected with a USB cable to a computer and it can share logic levels with IrDA. Bluetooth not only processes the data as the conventional serial port but it can transfer more than 60 RS 232 connections [35].

Theoretically, the Bluetooth systems operate transparently to the maximum extent. Devices can be connected without user's special assistance, providing they are in the range of communication.

In practice however, things are more complicated. Systems with Bluetooth must adapt to the variety of devices and data types. It is necessary to maintain a constant contact with each device. When a device appears or switches off, or goes out of the range, the system must detect the event. Besides, it is necessary to oversee the communication of devices to avoid simultaneous transmission and disturbances.

Structure and basic parameters

Bluetooth is a hardware module performing the functions of a radio transmitter and receiver under a driver's control (Fig. 2.33).



Fig 2.33. Bluetooth controller and transceiver

According to the specifications the transceiver operates within the range of 2400 ÷ 2483,5 MHz that is license-free in most countries. The distance at which two devices can be set apart is within 10 ÷ 100 m (Fig.2.34.). In return, several Bluetooth devices can be connected in a network through walls or several floors in a building, and without the need of direct visibility or external antenna. The channel width is 723,2 Kbps for devices operating in asynchronous mode and 433,9 Kbps for devices in synchronous mode.

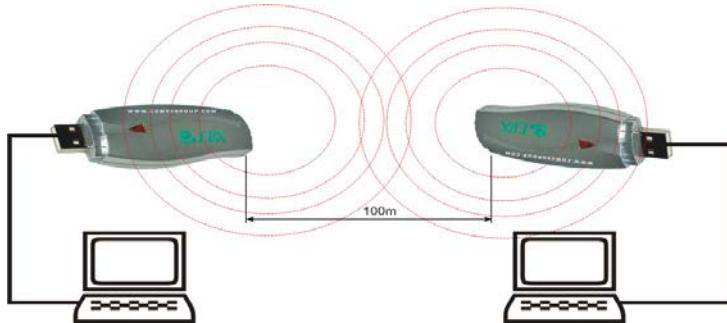


Fig 2.34. Bluetooth interaction between two devices

Data exchange can be implemented via 3 audio channels, each of them supporting a 64 Kbps synchronous transfer. The Bluetooth device enables the transmission of a composite signal of data bytes and audio data.

Bluetooth has another distinctive feature compared to other technologies and namely, different Bluetooth devices establish connections with each other automatically, immediately after discovering the transceiver's range. The software takes care of setting up the connection and communication.

Principle of operation

One of the Bluetooth advantages is that any device coming into the range is able to communicate with one device or a set of devices which support the technology, and they are not required to actively communicate.

In the context of the Bluetooth terminology, the device that exchanges information with other devices is a **master** and the device with which the master actively communicates is a **slave** (Fig.2.35a). The maximum number of slaves is 7. However, there can be an unrestricted number of inactive slaves, which after an established link and synchronization with the master device, do not communicate with it, but

passively wait for their turn to perform data transfers. Such type of communication is called **piconet** (Fig.2.35b). In the frame of a piconet communication there is only one master device, but whenever necessary, the connected slave can change its status to a master device forming its piconet structure. This type of a complex composite structure is known as **scatternet** (Fig.2.35c), where any device can be a master and a slave at one and the same time, depending on the specific situation and the position of the device within the architecture.

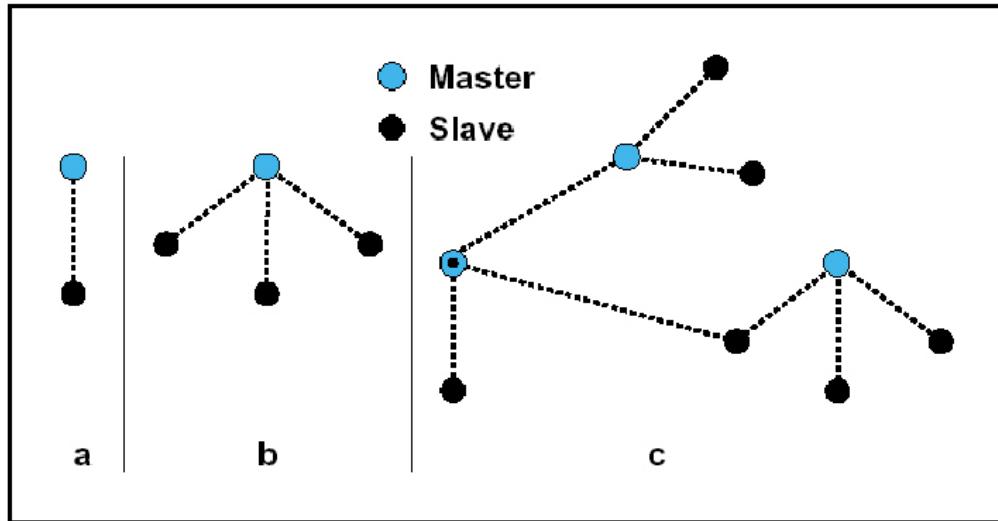


Fig 2.35. Bluetooth types of connections

Each device in the system has its name to avoid a duplicated position and other undesirable faults. It communicates at different frequencies with other devices via communication channels, known as **hopping channels**, characterized with a specific **hopping** parameter. Hopping is the periodical change of frequency defined with a **hopping sequence** parameter (sequence of frequency changes). According to Bluetooth specifications a hopping parameter defines 10 variants in which the frequency changes by 1 600 hops/sec.

At the start, a Bluetooth device scans the hopping channels looking for other devices. This is the so-called **Device Discovery** technology, and, depending on which of the below described three modes the devices are discovered, the communication takes place or not.

1. Discovery mode. Devices are ready to accept the connection setting up procedures. They exchange service information setting up all parameters specific for the connection between them.

2. Limited discoverable mode. Devices acknowledge the connection providing some conditions were observed (e.g. limited period of time). Devices can be discovered also by other participants in the linking process but do not allow setting up of some connection parameters, respectively data receipt/transmission.

3. Non-discoverable mode. Devices do not accept new requests.

Devices in the first or second above mentioned modes can be available also in **connectable** or **non-connectable mode**.

In the next stage of discovering, the names of all accessible Bluetooth devices are read. According to Bluetooth specifications each device is assigned not only with a unique network address, but on user level it operates with its own name that is up to 248 bytes long and not necessarily unique within the Bluetooth network.

When the link is established, the services of participating devices are automatically united. This is implemented by the **Service Discovery Protocol (SDP)**.

Security and Data Protection

Data protection technology contained in the Bluetooth protocol is defined by three modes:

1. Security mode 1 (non-secure). The device is not entitled to activate security tools.

2. Security mode 2 (service level enforced security). The device cannot activate security tools unless it is connected to a trusted device which triggers them depending on the type and requirements of the services used.

3. Security mode 3 (link level enforced security). The security tool is activated since establishing the link, but if a device does not respond to the requirements it will not be able to communicate.

Security modes 2 and 3 might be used in a partnership that adds to raising the level of protection. The basis for the highest protection level in Security mode 3 is the *Link key* or *Bond*. The link key is generated in the process of establishing the link between two devices and it is used for data identification and encrypting. In spite of all recognized and specific protection methods used by the Bluetooth security system, there is a possibility of data traffic eavesdropping and deciphering. This is a serious disadvantage of the wireless radio communications.

The applications of Bluetooth interface are analogous to those of the IrDA.

Benefits and disadvantages

On one hand, the Bluetooth technology has some disadvantages which limit its share of application, compared to the interfaces which are currently used, e.g.:

- Comparatively high production cost (to the moment);
- Lack of reliable support on operation system level (the problem can be solved by writing of drivers);
- Lack of manufacturers' interest to offer devices with Bluetooth interface;
- Unsolved problems related to vulnerability of data transfer between devices (it is easily intercepted and deciphered remotely);
- Comparatively low speed of data transfer;
- Limited broadcasting range.

On the other hand, the Bluetooth technology is featured by benefits which may help its further successful development, e.g.

- Well-built architecture;
- Continually lowering cost of the hardware module (single chip solution);
- Support granted by the consortium of over 2000 members including companies like IBM, Intel, Nokia, Ericson, Toshiba, 3COM, Lucent, Microsoft, etc.

Software aspect

As a new interface Bluetooth processes the raw source data which it further transmits. It collects the information bits in packets with an incorporated error controller. Then it unites the sequence of data packets in a conjunction (logic "AND"). It multiplexes the link so that several serial flows of data are simultaneously identified as one Bluetooth link [33].

Bluetooth packs the data by cutting the serial data flow into tiny pieces each assigned with an address and optional information for error correction. The link is represented by the sequence of packets starting with a single data flow that is recovered eventually to the exact copy of the data flow.

The Bluetooth standard supports two types of connections – synchronous and asynchronous. Synchronous connection is the typical voice information analogous to the sound of a phone conversation, while asynchronous connection is the typical computer information.

The synchronous data depends on the time interval and it is transmitted once, without the necessity to be confirmed. If synchronous packets are lost during the transmission they are lost for ever. Synchronous connections between Bluetooth devices ensure a full duplex channel with effective transmission rate of 64 Kbps in each direction. In fact, the synchronous connection is a standard digital phone channel with 8 bits and a sampling rate of 8 KHz. The Bluetooth standard allows three synchronous connections to be shared simultaneously by two devices that is equivalent to three phone conversations in a real time mode. All connections start asynchronously, as the commands can be sent only in asynchronous packets. After the link is established the master and slave device negotiate for switching to synchronous connection for the voice data transfer, or for asynchronous data transfer.

Each piconet has one master and several slaves. It shares one communication channel with all devices which operate in the same frequency and clocking. The channel is shared by one or more connections - asynchronous and/or synchronous.

To oversee the competition among lots of connections in a channel, the Bluetooth uses a time-division multiplexing. Each individual packet in the established connection gets its time for the data transfer. The Bluetooth standard divides the communication channel into time slots of 625 μ s. The tiniest single packet uses one slot and has a spare time left. Nevertheless, Bluetooth allows the packets to be delayed up to 5 time slots.

Within the Bluetooth systems each packet specifies the frequency hopping, i.e. the frequency switching after each packet. The minimum hopping length corresponds to one slot though it can last up to 5 time slots in order to match with the longer packets.

Time-division duplex in the Bluetooth systems operates by assigning odd time-slots to the master device and even slots to the slave. The master can start transmitting only in an odd slot. If the packet takes an even number of time slots (two or four), the slave is not able to start until the next even slot has come. In fact, more even slots are used than odd ones, despite of the availability of short packet transmissions.

Hardware aspect

Bluetooth hardware ensures the connection for controlling the process and the data packet formation. Although the connections are wireless, the hardware system is a set of electrical circuits for implementing correct data transmission and receipt [33].

As all radio systems, Bluetooth starts with a carrying frequency and modulates it according to the data. However this interface does not use a fixed carrying frequency but it sooner hops to different frequencies more than a thousand times per second. Each device supports a clock generator which determines the occurrence of each bit in the serial data flow. Bluetooth skillfully combines the elements necessary for constructing a wireless communication network.

Clock generators

Each Bluetooth device has its internal clock generator that controls the communication process. It operates independently almost at the required speed.

For effective demodulation of signals, the clock generators of the master and slave devices must be synchronized. While the master sets the frequency for all slaves with which it communicates, the slaves determine the exact clock frequency depending on the data packet. The heading part of every packet contains a pre-defined template of several clocks which the slaves can use to achieve synchronization. Bluetooth systems do not modify the frequency of slaves' clock generator. Instead, they store the frequency differences between the master's and slaves' clock generators and use these differences for compensating the master's clock generator.

When another master device takes the control during a communication session, every slave adapts the stored differences, in order to set up the frequency coordination precisely.

Topology

Bluetooth is designed to support plenty of connection types. The main connection is peer-to-peer where one device operates as a master and the other device - as a slave. In the piconet configuration one master is able to communicate with maximum seven slaves. Besides, other slave devices may follow the signal of the master so that they become ready for communication without sending an active signal. Such passive slaves are defined as being in a 'parking' state.

The master in a piconet connection determines which devices are able to communicate (i.e. which devices are active and which are parked). Several piconets can be united in a scatternet, i.e. the master of one

piconet communicates with the master or slave of another piconet connection.

Frequencies

Bluetooth operates in radio frequencies for industrial, scientific, medical, etc. purposes. The range is within the high frequencies and in most countries it is unlicensed for low-power devices. Table 2.17 gives examples for frequencies and channels used in some regions [33].

A Bluetooth system does not operate only in one frequency but uses all the possible ones and forms channels that vary up to 1 600 times per second. Regretfully, Bluetooth cannot get use of the entire 2,4 GHz range. The wireless network standard IEEE 802.11 uses the same frequency range and the interferences between the two systems are unavoidable.

Region	Frequency range	Channels used
North America	2,400÷2,4835GHz	79
Europe (except Spain and France)	2,400÷2,4835GHz	79
Spain	2,445÷2,475GHz	23
France	2,4465÷2,4835GHz	23
Japan	2,471÷2,497GHz	23

Table 2.17. Bluetooth frequencies and transmission channels

Power

Bluetooth specification defines three classes of equipment based on the transmission power (Table 2.18).

Power class	Maximum transmission power	Minimum transmission power
1	100 mW	1 mW
2	2,5 mW	0,25 mW
3	1 mW	Not specified

Table 2.18. Transmission power of different power classes

As for any radio system, the higher power covers a wider range. On the other hand the higher output power means higher power consumption that immediately reflects on the power supply source (battery).

Components

Bluetooth architecture is composed of three parts: radio, link control and support, which provide the control of communication and interface with the terminal device. These are functional partitions which are

integrated in most Bluetooth devices. As for a PC, their location is on the periphery controller that is installed like any other additional hardware on the computer.

The radio part includes the Bluetooth hardware. It determines the device power and frequency range, broadcasts the carrying frequency via the antenna after modulation and amplification. The high frequencies, used by the Bluetooth system, have a short wave length which allows the antenna to be tiny and incorporated in the case of many mobile devices.

The part responsible for the link control is the base of the Bluetooth system. It includes various control and management protocols for setting up and supporting the wireless communication. It searches and identifies new devices willing to join the piconet connection, follows up the transition of frequencies and controls the working state of the device.

The supporting part provides the actual interface between the terminal device logics and the Bluetooth connection. It adapts the signals (electrical and logic) of the terminal device to those of the Bluetooth system. For instance, on the Bluetooth card of a PC the supporting part adapts the signals from the parallel bus of the PCI connection into the form of serial packets used by the Bluetooth. It also checks the input data from the wireless connection for errors and requests re-transmission, if necessary.

3. DISPLAYS

3.1. Classification of Displays

The analysis of modern state of electronic devices development for information displays shows that practically, it is very difficult to classify them in view of their diverse properties and functional capabilities. (Fig. 3.1). [20]

The classification should be based on the principle of grouping the displays according to: type of displayed information, indicator type, method of constructing images, dimension of displayed information, intellect level, etc.

Depending on the type of information displayed on the screen, displays are divided into alphanumeric (character) and graphical. Alphanumeric displays are designed to operate with information consisting of letters, numbers and accompanying symbols. Displays are organized into complexes within data transmission systems. In display complexes, in addition to the functions common for all displays (display, storage, operating input and editing of information), data transmission along the link channel is added. Alphanumeric displays have a very wide range of applications which covers all spheres of using problem-oriented complexes.

The modern stage of development of input and output units for graphic information is characterized by intensive development of graphic displays. They are designed for operating interaction between man and computer in dialog graphic mode (drawings, charts, diagrams, etc.) in various systems for computer aided design: computer aided research, management of enterprises, computer aided manufacture, robots, training process, etc.

All graphic displays can show characters and accompanying symbols, and some character displays contain equipment for plotting simple graphic images. Therefore it is very difficult to draw a clear boundary line between these two classes of displays.

Since from classification according to the type of displayed information, it is easy to go to classification according to the method of constructing images, most of displays are divided into two major classes.

Vector displays with arbitrary coordinate scanning. The image is formed from the totality of graphic elements (points, vectors, circumferences, arcs, characters, etc.). They are specified on the screen through coordinates.

Raster displays. The image is created by changing the brightness of the displayed spot at a certain moment of time during its uniform motion along a constant trajectory.

Two other fundamental groups of displays are formed on the basis of physical principles.

Displays based on cathode ray tubes (CRT): colour and black-and-white, storing, multi-ray, projection CRT, etc.

Displays based on flat panels.

The CRT is the main component of the display. CRT advantages guarantee leadership of displays in indicator equipment.

Although the latest flat-panel indicators were designed implementing recent technological achievements, an ideal CRT does not exist yet.

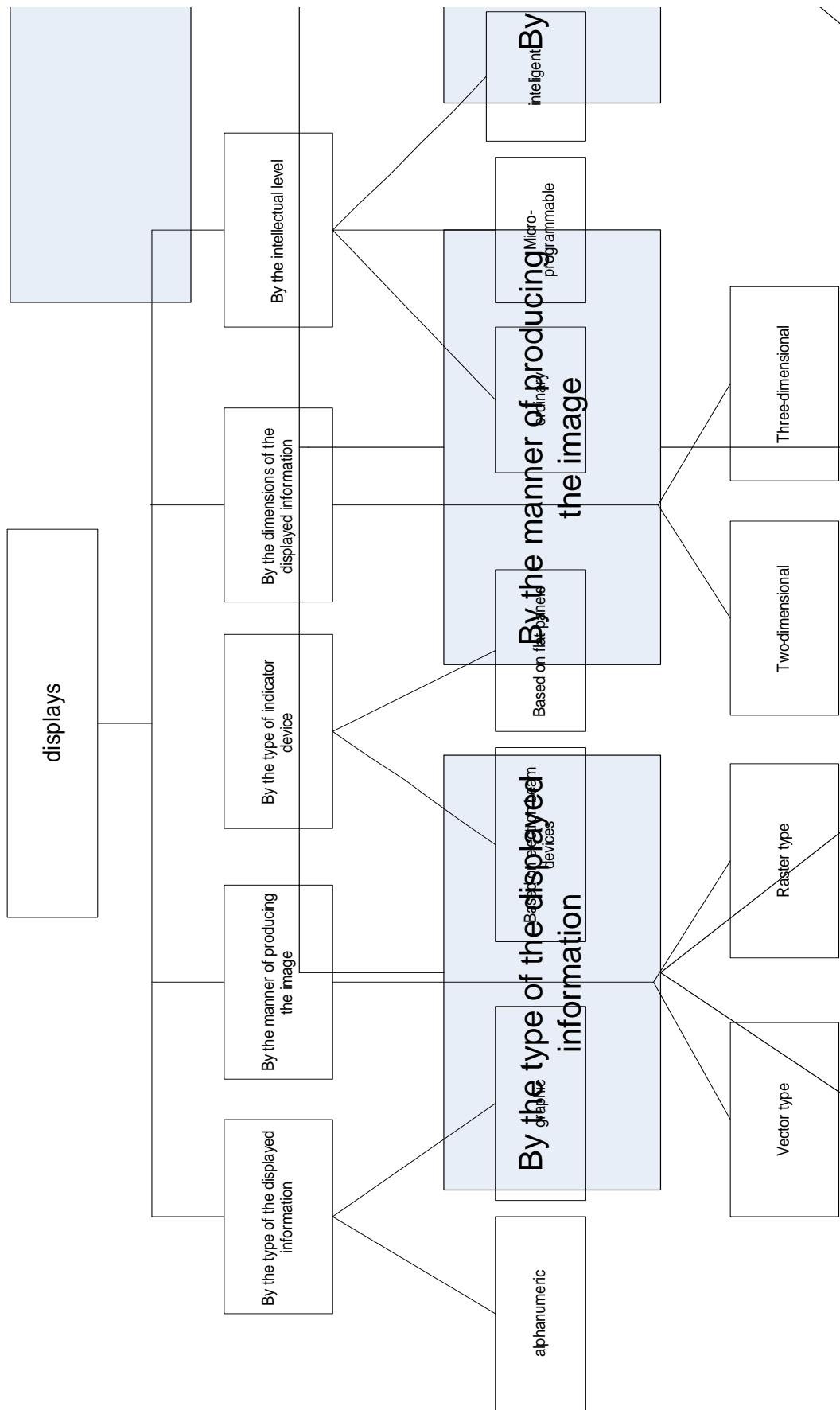


Fig.3.1 Classification of displays

The advantages of flat-panel indicators are obvious: small overall dimensions, small mass and low power consumption; lack of analog nodes for the scanning of high-voltage power sources; high quality of image – lack of flicker, blinking, flash and distortion; high stability; brightness and contrast. All these characteristics appear to be important and defining for choosing a display. Therefore users prefer displays with flat-panel indicators.

According to the intellect level (a certain set of system functions which are performed in the display itself), displays are divided into: ordinary, micro-programmable and intelligent. In ordinary displays a limited number of functions: display, storage, editing, input and output of data are performed by hardware. Micro-programmable displays allow flexible change and extension of the set of functions and also not complicated data preprocessing. As a rule, intelligent displays contain: a processor, main and read-only-memory, external memory, input-output devices and allow a wide range of functions concerning data processing and display. Modern graphic displays are an example of this type.

3.2. Video block

Principle of information display on video monitors (based on CRT).

The video display generated by a certain type of monitor appears to be the chief criterion of its class. It is evaluated by the number of dots (elements, pixels) displayed on the screen of the video monitor, by the number of colours which can be used simultaneously, by the manner of information display on the screen. [22]

Video monitors included in operating configurations of personal computers, work on the same principle as television sets. The only difference is the lack of an antenna unit for receiving television programs and the lack of an audio unit for providing audio information.

The principal part of the video monitor is the cathode ray tube (kinescope). It comprises a screen at one end and an electron gun at the other end. A simplified construction of a cathode ray tube is shown in Fig. 3.2.

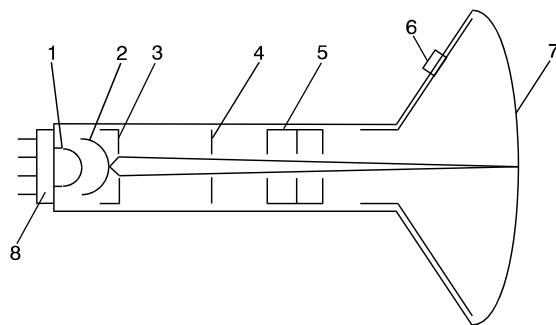


Fig 3.2. Cathode ray tube construction:

- 1 – filament heater;
- 2 - cathode;
- 3 – control electrode;
- 4 – accelerating electrode;
- 5 – focusing electrode;
- 6 - anode;
- 7 - screen;
- 8 - base.

The cathode, which emits the electrons necessary for the operation of the tube, contains a filament heater (1). It is similar to the filaments in incandescent lamps and is heated by electric current. The filament is coated with a layer of insulating fireproof material. The heat provided by the filament is transmitted through this layer to a metal tube which encloses it tightly. Finally, the surface of the metal tube is coated with an oxide layer producing electrons and this is the cathode itself (2). The emitting layer has a small surface so that the emitted electrons can be formed into a narrow stream.

The electrons emitted by the cathode are formed into a beam and are directed at high velocity to the screen of the cathode ray tube. The inner surface of the screen is coated with a thin layer of semi-translucent substance called phosphor. When this layer is 'bombarded' by electrons it fluoresces for a short time.

It is the fluorescing spots 'bombarded' by the electron beam that produce the desired image. However, to obtain a satisfactory image, the electron beam must be accelerated sufficiently because the more strongly electrons hit the screen, the brighter it becomes. Furthermore, the brightness of a certain spot on the screen can vary by sending different number of electrons. The electrons in the electron beam are regulated and accelerated by placing various electrodes along their path.

The electrons emitted by the cathode would not reach far if they met air molecules along their path. Therefore, it is necessary to evacuate the tube and to ensure free movement for the electrons.

The control electrode (grid electrode) (3) is a cylinder enclosing the cathode. A negative potential, with respect to the cathode, is supplied to the control electrode. It hinders the motion of the electrons emitted by the cathode. If this negative voltage is sufficient, the control electrode repels most of the emitted electrons back to the cathode and lets pass only a very small part of them. Conversely, if the control electrode is not very negative, a large part of the electrons pass through it and go further. In this way varying the potential between the control electrode and cathode, the number of electrons let pass to the screen varies and hence, the brightness of the spot which they hit varies as well.

The accelerating electrode (4) is positioned after the control electrode in the CRT. It is a metal disk with a hole in the center. Positive potential with respect to the cathode, of the order of several hundreds of volts, is supplied to this electrode. Owing to this positive potential, the electrons allowed to pass by the control electrode are attracted strongly and as a result their velocity increases sharply. So they travel through the hole of the accelerating electrode to the screen (7). Along their path they are further accelerated by the focusing electrode (5). The latter is also supplied with positive potential with respect to the cathode and amplitude of several hundreds of volts. The electric field between the accelerating and focusing electrodes changes the trajectories of the electrons, directing them towards the CRT axis and thus focusing them. In this way the electrons are directed further to the screen as a very narrow beam.

The part of the cathode ray tube that we have dealt with so far, which produces and forms the electron beam is called electron gun. The electrons leaving the electron gun have a very high speed exceeding 60000 km/s. As we have mentioned above, when the electrons strike the screen, they form a

luminescent spot. At the same time these electrons have to be taken out of the cathode ray tube because otherwise they will accumulate in the screen zone and will repel the other coming new electrons. To this end, the inner side of the CRT between the focusing electrode and the screen is coated with a conducting graphite layer called anode (6). It is supplied with very high positive, as regard the cathode, potential (over 10000V) and as a result its field attracts the electrons which have impinged upon the screen and after that 'leads them away' to the current source. Moreover, on the inner side of the screen, over the luminescent layer, a very thin transparent layer of aluminium is applied through which the electrons from the electron gun can pass easily. This layer facilitates their leading away from the screen. However, its main purpose is to enhance the image brightness, reflecting this part of the light beam to the viewer which would be lost irretrievably if it were directed towards the interior of the cathode ray tube.

The terminals of all electric parts, except the anode, are located on the base (8) of the CRT. Since very high voltage is applied on the anode, it is necessary to separate it from the other electrodes and therefore its terminal is located in the conical part of the flask.

In short, the electrons start from the cathode, travel through the holes of the control, accelerating and focusing electrodes, head for the screen and rebounding from it, they return to the current source through the anode. Figuratively explained, this motion of electrons is like an electron pencil with which luminescent images can be drawn on the screen. To accomplish this, it is necessary to move the electron beam to all spots of the screen. How is the electron beam deflected?

The electric current passing through a conductor produces a circular magnetic field around the conductor. The conductor is the center of the magnetic field. An electric current is a flow of electrons. A circular magnetic field is produced around the electron beam generated by the electron beam. What will the electron beam behavior be if it is placed with its own magnetic field in another magnetic field? Fig 3.3 shows the side of the electron gun placed between the poles of a horseshoe permanent magnet. The small black dot in the CRT center is an electron beam moving towards the screen. The two magnetic fields (the field of the permanent magnet –the parallel magnetic lines of force and the circular field of the moving electrons) are in continuous interaction. In the upper part their directions are opposite and forces of attraction act. In their lower part the fields have the same directions and forces

of repulsion exist there. As a result of the interaction of the two fields, the moving electrons are deflected upwards. If the poles of the magnet are changed, the electrons will be deflected downwards. If the magnet is rotated to 90°, so that its magnetic lines of force become vertical, the beam will be deflected to the left or to the right. And finally, if we want the electron beam to move describing a random continuous line, both the direction

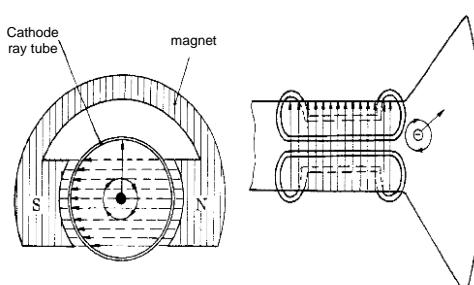


Fig. 3.3.
Electron beam
deflection
in a magnetic
field.

Fig. 3.4.
Horizontal
deflector
coils.

and intensity of the external magnetic field should be varied continuously. Therefore, two sets of coils are placed at the end of the electron gun. The first set produces a vertical magnetic field (Fig.3.4) and by varying its intensity and direction the electron beam starts deflecting to the left or to the right. The second set produces a horizontal magnetic field which can deflect the electron beam upwards or downwards. When alternating voltage is applied to this set, a bright spot moving along a vertical line will be seen on the screen. If the frequency of this voltage is increased, the spot will move so fast that its successive positions will not be seen as separate. Only a vertical line will be shown on the screen.

Television raster. Video signal.

The electron beam can be deflected horizontally and vertically by means of two sets of coils. However, it also can be made to draw a given image on the screen. [22]

Let a potential with the shape shown in Fig 3.5 be applied to the coils causing horizontal deflection. In this case the bright spot on the screen will move at constant velocity from left to right, drawing a horizontal line. When it reaches the right end, it will return very fast to the left end and then its motion to the right will start again and so on.

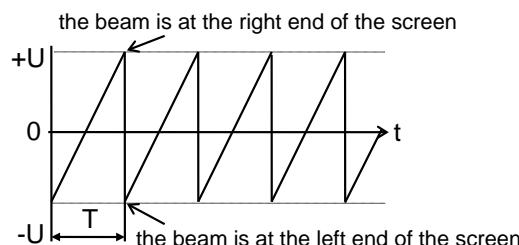


Fig. 3.5. Shape of the potential for horizontal scanning

return of the spot to the left end and then everything repeats many times. Due to its shape this potential is called saw-tooth. Time T , called saw-tooth potential period, is the time during which the electron beam scans a line on the screen.

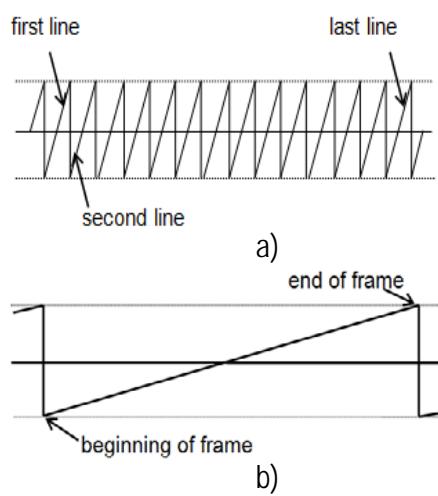


Fig. 3.6. Potentials for horizontal (a) and vertical (b) scanning

This motion of the spot is similar to the movement of the eyes when reading one and the same line in a book. The slow transfer from negative potential $-U$ to positive potential $+U$ controls the movement of the spot from left to right. After that the instantaneous change of potential from $+U$ to $-U$ controls the fast

Now, continuing to apply saw-tooth potential to the coils causing horizontal deflection, let's apply voltage which slowly moves the beam from top to bottom to the coils which cause vertical deflection. In this case, when the beam scans a line from left to right, it returns again to the left corner of the screen but lower and starts scanning a second line, then a third one and so on.

Now the beam motion is similar to the movement of the eyes when reading a book one line after

another. After reaching the bottom right corner, the beam should be able to move to the upper left corner. Therefore, the vertical deflection voltage (VD) should also be saw-tooth, but to vary much more slowly than the horizontal deflection voltage (HD) – Fig. 3.6.

By means of these voltages called respectively horizontal and vertical scanning voltages, the beam is made to cover all elements on the screen and in this way the so-called television raster is obtained (Fig. 3.7).

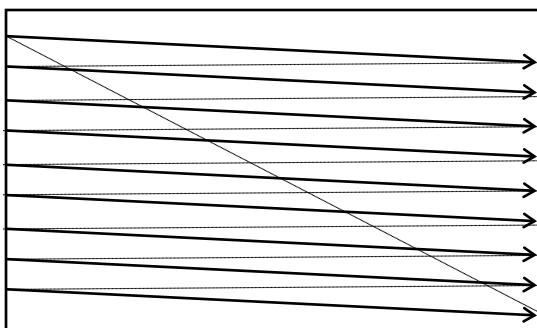


Fig. 3.7. Television raster

The horizontal scanning (horizontal timebase) is also called line scanning, and the vertical scanning (vertical timebase) – frame scanning. The number of frames in the raster must be large enough so the viewer's eye can see the picture as a whole without discerning separate lines. For some monitors this number is 625.

In the raster obtained, the brightness of all elements is the same. However, in order to display an image (text, graphic or picture) on the screen, there must be a combination of dark and bright spots. This combination is provided under the control of the personal computer and especially of its block called video monitor controller. From the main memory of the computer it receives the information which must be displayed on the screen and on this basis produces the signals controlling the operation of the video monitor. These signals are passed from the controller to the video monitor and determine successively the brightness of all elements taking part in a certain image.

To obtain a durable image on the screen, the controller of the video monitor generates the brightness signal determining in succession the brightness of all spots of the transmitted image. It must produce four more types of signals.

1. Signals for horizontal scanning synchronization (line sync pulses - LSP), which fix the end of each line of the image.
2. Signals for vertical scanning synchronization (frame sync pulses – FSP), which fix the end of each frame of the image.
3. Signals for blanking the electron beam during its return in horizontal direction (line blanking pulses - LBP).
4. Signals for blanking the electron beam during its return in vertical direction (vertical blanking pulses – VBP).

The same types of signals are also used in television for television broadcast.

From these four types of signals plus the brightness signal, the so-called complex video signal is formed (further it will be called simply video signal). It is received by the video monitor controller by adding the brightness signal to the sync pulses (SP) and the blanking pulses (BP), as it is shown in Fig. 3.8.

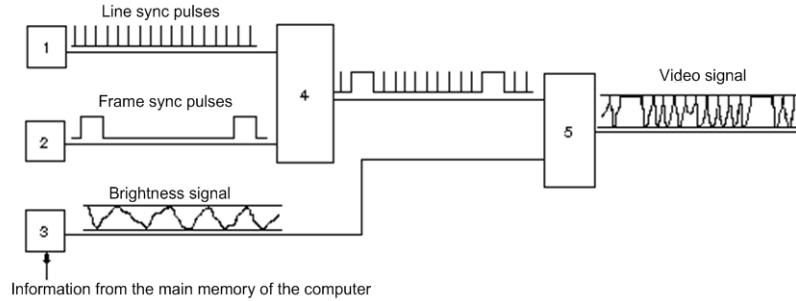


Fig. 3.8. Block diagram for obtaining a video signal in personal computer:
 1 – LSP generator; 2 – FSP generator; 3 – video generator;
 4 – SP mixer; 5 – video signal mixer.

The LSP generator (1) produces SP which set the return travel of the beam during line scanning. Similarly, the FSP generator (2) produces SP, which set the return travel of the beam during frame scanning. In order to be able to distinguish these two types of SP, they should have different length. LSP and FSP are summed up in the SP mixer (4) at whose output a total signal is produced containing both line and frame SP. The video generator (3) receives the information which must be displayed on the screen from the main memory and depending on it produces the respective brightness signals. These signals go to the video signal mixer (5) where they are added to the SP. So at the mixer output, a video signal containing all signals necessary for the control of the video monitor (or TV set) is formed.

The video signal shape and components are shown in Fig. 3.9.

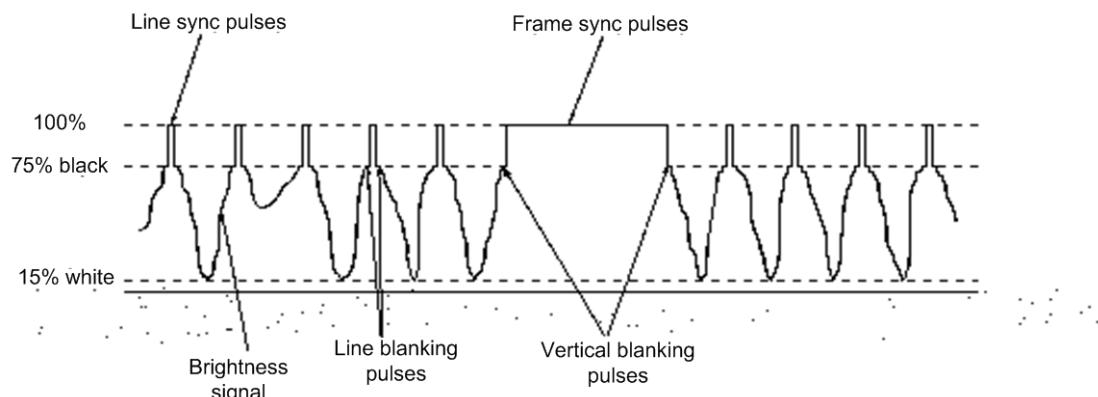


Fig. 3.9. Simplified shape of the video signal

SP amplitude is much larger than the amplitude of the brightness signal which can be 75 % of the amplitude of the whole signal at the most. A signal with amplitude over 75 % than the maximum corresponds to the least brightness (black level), whilst a signal with an amplitude below 15 % than the maximum corresponds to the greatest brightness (white level). All the remaining brightness grades correspond to values within the region between these two boundary values of the signal.

LBP are in the region immediately before and after LSP. Their amplitude is 75% of the maximum, i.e. corresponds to black level and therefore they blank the beam in a section of the line end and in a section at

the beginning of the next line. Similarly, VBP blank the beam in a section of the frame end and in a section at the beginning of the next frame. As a result of the BP action at the four sides of the television raster a band is ‘cut’ which remains invisible to the viewer.

So, if the potential between the cathode and the control electrode of the CRT is varied, the number of electrons in the electron beam is also varied. This leads to changing the brightness of the spot on the screen where the beam strikes. In this way, through controlling the brightness of all spots of the raster in compliance with the video signal received from the computer, the desired image is obtained. This is done most easily if, after the respective amplification, the video signal is sent to the CRT cathode. When the cathode is given the part of the video signal with larger amplitude, the control electrode becomes more negative with respect to the cathode and lets pass fewer electrons (i.e. provides little brightness). Vice versa – when the sent signal has smaller amplitude, the brightness will be greater. The video signal is sent to the cathode in its complex form, i.e. together with SP. They are not an obstacle because they have large amplitude and set small brightness. Moreover, this is very convenient, because during the return travel, both in lines and frames, the beam is invisible.

The video signal has a wide frequency band. The maximum participating frequency increases when the frame frequency rises. The excessive increase in frequency, however, results in difficulties in the video signal transmission. On the other hand, for better visualization of the image, i.e., the beam motion cannot be perceived (having in mind the inertness of human sight), the frame frequency should not be less than 30 Hz. In order to avoid this controversy, instead of the progressive scanning that we have already dealt with above (Figs. 3.5, 3.6 and 3.7), where lines are transmitted in succession, in television the so called interlaced scanning is employed where all odd lines are transmitted first and then all even lines. So in interlaced scanning each frame consists of two images transmitted in succession and they are called fields. Displaying the fields one after the other happens so fast that the eye sees them on the screen as a whole picture. As it was mentioned above, in some video monitors each frame consists of 625 lines (one line is transmitted for $64\mu s$), i.e. each field contains 312,5 lines. The frame frequency is 25 Hz. Owing to the use of interlaced scanning, an illusion is created about doubling the frame frequency, because 50 fields per second are transmitted. This improves significantly the image quality, without increasing the maximum frequency of the video signal. For the given SP parameters, the video signal frequency band has a width of up to 6MHz.

Block diagram of a black-and-white video monitor (monochromatic video monitor).

A block diagram of a black-and-white video monitor is shown in Fig. 3.10. The video signal from the respective controller of the personal computer enters the video monitor input. It is amplified by video amplifier (VA) (1) up to the amplitude required for controlling the beam brightness and then is passed to the cathode of the CRT. This amplitude can be varied by the contrast control: when the signal has larger amplitude, the difference in brightness between light and dark spots is greater, i.e. the image is more contrasting.[22]

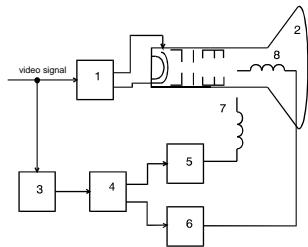


Fig. 3.10. Block diagram of a black-and-white video monitor:

1 – video amplifier; 2 - CRT; 3 – amplitude limiter; 4 - selector; 5 – line time-base oscillator; 6 – frame time-base oscillator; 7 - horizontal deflector coils; 8 – vertical deflector coils.

Constant negative voltage determining the average screen brightness which can be varied with the brightness control is applied to the control electrode of the CRT.

The video signal is passed both to the CRT cathode and to the amplitude limiter (3) whose function is to separate the line and frame sync pulses from the complex video signal by passing them further (Fig. 3.11). Then the so obtained SP enter the selector (4), where they are separated into LSP and FSP (Fig. 3.12). LSP are passed to the time-base oscillator for line scanning (5). It produces saw-tooth potential for the horizontal deflector coils. Similarly, the FSP are passed to the time-base oscillator for frame scanning, (6), which produces saw-tooth potential for the vertical deflector coils.



Fig. 3.11. Function of the amplitude limiter

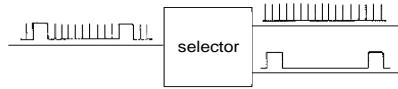


Fig. 3.12. Separation of line and frame SP

Block diagram of an alphanumeric video controller for black-and-white display.

The internal architecture of the controller is shown in Fig. 3.13. The characters displayed on the screen (their ASCII codes) are stored in video RAM. They are saved there by the microprocessor of the personal computer. The character generator is ROM memory where the shape tables of all characters displayed are saved. The addressing to these tables is as follows: the ASCII code of the character is used as the most significant part of the address and sets the shift of the respective shape table from the beginning of the character generator. The least significant part of the address is set by the CRT controller and specifies which line of the shape table is displayed at the moment on the screen. [18]

The CRT controller fulfills the following functions:

- selects the codes from video RAM sequentially and they are passed to the character generator and serve as a base address of the shape table of the respective character;
- selects sequentially the lines of the shape tables saved in the character generator which enter the shift register input;
- produces line and frame sync pulses and also the clock cycle for the shift register.

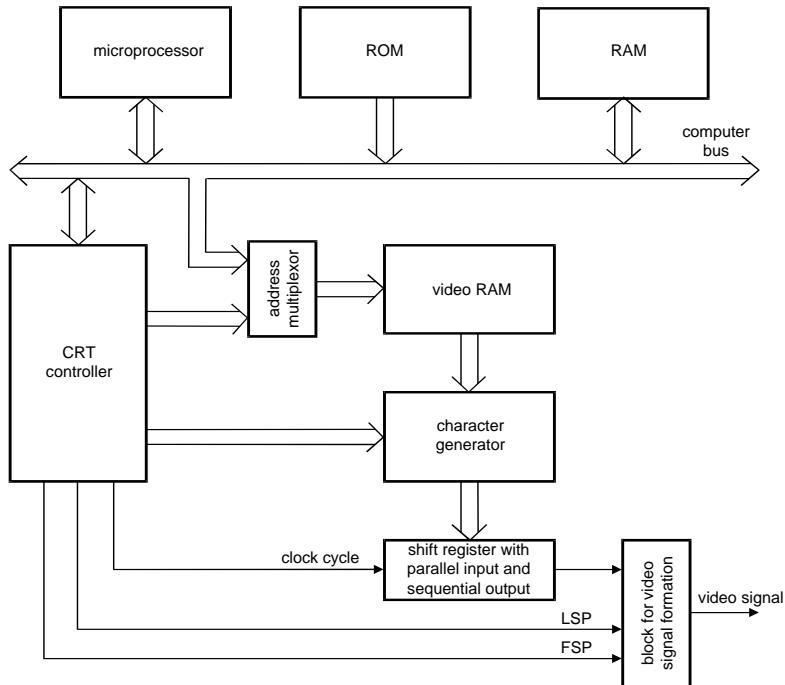


Fig. 3.13. Block diagram of an alphanumeric video controller for black-and-white display

The shift register converts the data from the character generator into serial format. So the brightness signal for the CRT is obtained. In the block for video signal formation, the brightness signal, LSP and FSP are mixed to form the complex video signal. The address multiplexor serves for preventing conflicts when there is simultaneous access to video RAM of the microprocessor and controller of the CRT.

The sequence of displaying a line of characters on the screen contains the eight steps given below:

1. The CRT controller addresses the ASCII code of the first character in the line in video RAM. It is used as the most significant part of the character generator address.
 2. The CRT controller sets the least significant part of the character generator address so that it points to the first line of the character shape table.
 3. The data from the character generator are saved in the shift register and sent in a serial format to the CRT.
 4. The CRT controller addresses the ASCII code of the next character in the line in video RAM.
 5. Steps 3 and 4 are repeated until the last character in the line is reached.
 6. The CRT controller addresses the ASCII code of the first character in the line in video RAM and sets the least significant part of the character generator address so that it points to the next line in the shape table of the character.
 7. Steps 3, 4, 5 and 6 are repeated until the last line in the shape table is reached.
 8. The end .
- The described procedure is repeated in displaying all lines.

3.3. Colour video monitor. The principle of producing colour images. Colour CRT.

Block diagram of colour monitor

The principle of producing colour images.

Colour video monitors have an entirely similar purpose as black-and-white video monitors. On the screen they display the information received from the personal computer: alphanumeric texts, graphics or pictures from electronic games but owing to the usage of colour CRT and the respective controlling signals, a colour image is produced.[22]

Fig. 3.15. Colour perception when watching a colour triade from a large distance

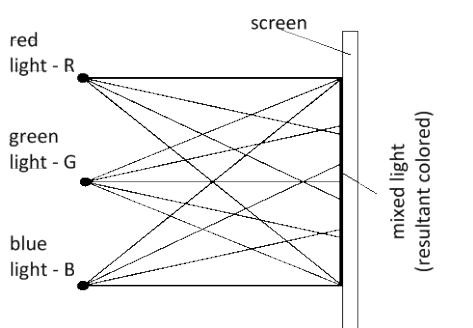


Fig. 3.14. Producing different colours by mixing the three primary colours

light, purple colours are produced, and from blue and green light – green-blue colours. The combination of three lights consisting of 33% R, 33% G and 33% B, creates the impression of white colourless light.

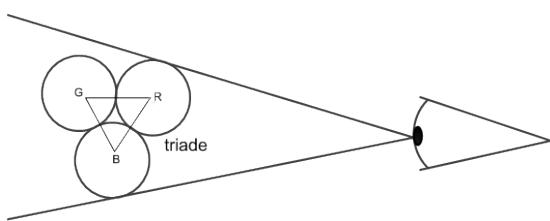


Fig. 3.15. Colour perception when watching a colour triade from a large distance

corresponding to the respective proportion (Fig. 3.15).

The colour image produced in television is based on an optical principle, i.e. all colours, white including, can be produced by mixing in certain proportions of light in the three primary colours: R - red, G – green and B - blue - Fig. 3.14. For example, when mixing red and green light, depending on their proportion, red (100% R, 0% G), orange (75% R, 25% G), yellow (50% R, 50% G) or green (0% R, 100% G) colour can be obtained.

Similarly, when mixing red and blue

In colour video monitors an elementary light source on the screen is the combination of three adjacent dots, respectively having red, green and blue colour, called triade. The mixed light of the triade seen from sufficiently great distance creates the impression that only one dot is lighted with a resultant colour

Colour CRT.

The appearance and operating principle of colour CRT are similar to the monochromatic CRT, but their constructions differ. The inner surface of the monochromatic CRT screen is coated with a thin phosphor layer which emits monochromatic light (green, blue or white) when an electron beam strikes it. In colour CRT the screen is covered by dots of three different in chemical composition phosphors which, when hit by an electron beam radiate, respectively red, green or blue light. [22]

Two types of colour CRT are mainly used:

- shadowmask kinescope ;
- chromatron .

In the shadowmask kinescope the screen is covered with three types of round phosphor dots which, when bombarded by electrons radiate, respectively red, green or blue light. In the first line the dots are arranged in the sequence: red, green, blue, red, green, blue, etc. In the second, third and following lines, the phosphor dots are arranged in the same way, but they are shifted in relation to one another so that they form an equilateral triangle – the above-mentioned triade (Fig. 3. 16).

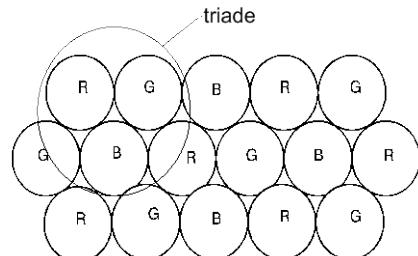


Figure 3.16. Positioning of phosphor dots on shadowmask kinescope

The dots arranged in this way are coated with a thin aluminium foil as in the monochromatic CRT. There are three electron guns (red, green and blue) in the shadowmask kinescope which have the same construction as in the monochromatic CRT.

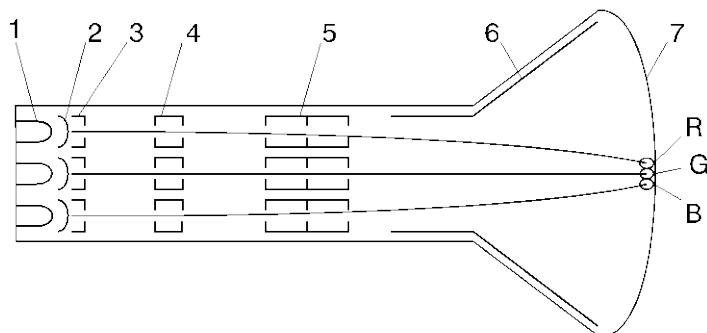


Fig 3.17. Construction of shadowmask kinescope:

- 1 – filament heaters; 2 - cathodes;
- 3 – control electrodes;
- 4 – accelerating electrodes; 5 – focusing electrodes; 6 - anode; 7 – screen.

Each of them directs its electron beam respectively to the red, green and blue phosphor dots (Fig. 3.17). The three identical electron guns are positioned along a circumference, shifted at 120° to one another. The blue gun is at the top, and the green and red ones are positioned symmetrically below it. A mask - 0.15 mm thick steel sheet in which circular holes are drilled at equal distances (a hole for each triade), is placed at a distance of 10÷15mm from the screen, so that each electron gun can direct its electron beam only to the phosphor dots with the corresponding colour. When the three electron beams pass through a hole in the mask, their paths cross and then the red beam lands on the red phosphor dot (R), the green

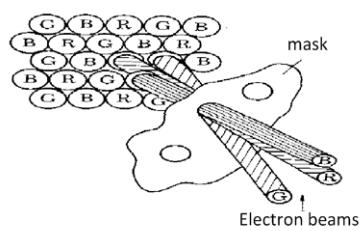


Fig. 3.18. Electron beams pass through the mask holes to the kinescope screen

cross and then the red beam lands on the red phosphor dot (R), the green

beam – on the green dot (G) and the blue beam – on the blue dot (B) of the triade (Fig. 3.18). If you look from the red gun only the R phosphor dots are seen through the mask holes. The G and B dots are not seen because they are overshadowed by the metal mask (therefore it is known as a shadowmask). Similarly, only the G dots are seen from the green gun, and only B dots – from the blue gun.

The display of colour images on the screen requires the simultaneous deflection of the three beams in horizontal and vertical direction. The beams are deflected as in black-and-white video monitors by sequential scanning. The essential thing here is that the three electron beams must always fall in a common hole in the mask regardless of its position (in the middle or at the end of the mask).

As a result of the action of the mask, the screen is lighted in red only when the red gun works. The red electron beam deflected in lines and frames causes the red phosphor dots, i.e. each third dot in triades, to be lighted sequentially. Similarly, the image is green, when only the green electron beam works, exciting the blue phosphor dots.

For producing other colours except the primary ones, two or three dots in each triade are excited simultaneously. Depending on the brightness ratio of dots having primary colours, all colours of the visible spectrum, white light and combinations of these colours can be produced.

The control of the electron beams in the three guns is similar to the electron beam control in the black-and-white kinescope. By varying the voltages across the cathodes and the respective control electrodes, the intensity of the three beams is controlled and hence the brightness of the red, green and blue dots in the triade and thus the resultant colour of the triade.

During the horizontal and vertical scanning (timebase) part of the electrons hit the mask (between the holes). As a result, the shadowmask kinescope has low efficiency which leads to the need to apply higher voltage at its anode compared to the black-and-white kinescope. Besides, the mask is heated very much by the intercepted electrons and as a result it expands and this leads to its displacement in relation to the screen. This can result in the beam (for a certain colour) falling on an adjacent phosphor dot, i.e. wrong colour. Therefore, the mask holes are smaller than the phosphor dots.

The three gun system is mounted in the narrow part (the neck) of the kinescope and therefore it is wider than the one of the black-and-white kinescope which has one electron gun.

A deflection system is used to deflect the beams in horizontal and vertical direction which is similar to the deflection system of the black-and-white kinescopes in construction and operating principle. For the deflection system of the three-beam mask kinescope, however, more powerful electric sources are necessary for supplying saw-tooth currents for horizontal and vertical deflection. This is required due to the higher anode voltage and the larger diameter of the narrow part

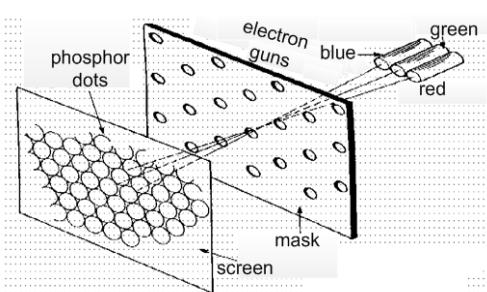


Fig 3.19. Kinescope for video monitor

(the neck) of the kinescope.

In kinescopes used in video monitors, the three electron guns are positioned in a horizontal line instead of in the vertices of an equilateral (Fig. 3.19). [18] Thus three electron beams crossing in the mask plane, are more easily produced. This construction allows through special make of the focusing and deflecting systems to avoid the use of a system for dynamic convergence of beams.

The major disadvantages of kinescopes with triangular (delta) positioning of the guns are the complicated system for dynamic convergence of beams and low luminous efficiency of the screen related to the low mask transparency. These disadvantages are removed to a considerable degree in the new electron device – colour kinescope with planar positioning of guns and a slot shadow mask (stripe kinescope).

The chromatrone is a colour CRT which looks like a mask kinescope but differs in the screen and the mask is replaced by a grid consisting of parallel vertical filaments. On the inner side the chromatrone screen is coated with red, green and blue phosphor in the form of several hundreds of thin vertical stripes (Fig. 3.20).

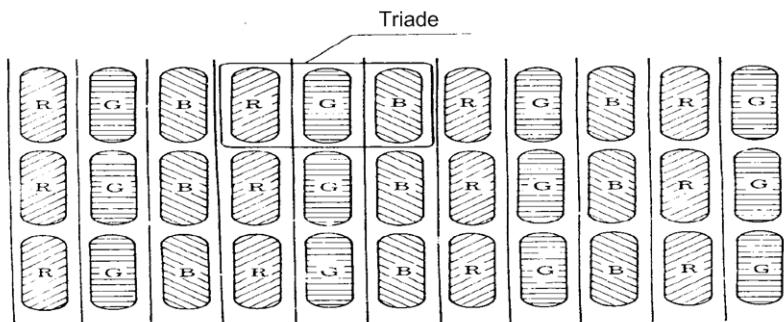


Fig 3.20. Positioning of triades in chromatrone

The phosphor layer is covered by a thin aluminium foil which has the role of a mirror which reflects light. At a certain distance from the chromatrone screen is the above-mentioned grid made of thin metal filaments, parallel to phosphor stripes. Whereas in the mask kinescope the guns are positioned along a circumference at 120° to one another, in the chromatrone the triple gun is horizontally placed in a plane. Furthermore, the phosphor stripe width, the distance between the grid and screen are selected so that the three beams cross in the middle between the grid filaments and fall on the three phosphor stripes.

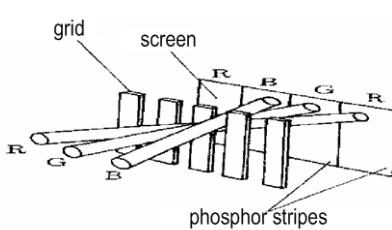


Fig. 3.21. Excitation of the dots with the three primary colours by the chromatrone phosphors

strips, so that each beam must fall only onto its corresponding phosphor stripe (Fig. 3.21). The grid has a lower potential than the anode potential and therefore an electronic lens is created which additionally focuses the electron beams. Thus the dots on the screen created by the electron beams have an elliptical shape whose longer axis is parallel to phosphor stripes. Whereat, the focused dot has smaller dimensions than the phosphor stripe. The potential

difference between the screen and grid produces additional acceleration of electrons which ensures greater dot brightness.

The chromatrone has greater efficiency because only 15 to 20% of electrons get lost in the grid filaments. The deflective system of the chromatrone is similar to the deflective system of the mask kinescope. The chromatrone, however has a greater sensitivity to deflection and therefore the needed power for deflection in horizontal and vertical direction is smaller compared to the mask kinescope. The chromatrone has a shortcoming, namely, that strong acoustic vibrations cause vibrations in the grid filaments which distort the image.

Block diagram of a video monitor.

A block diagram of a video monitor is shown in Fig. 3.22. In order to produce the desired colour image, from the colour video monitor controller, placed in the computer, the following signals controlling the operation of this video monitor are sent. [22]

1. LSP control the line scanning. These SP determine the end of each line and respectively, the beginning of the next line.
2. FSP control the frame scanning .
3. Three brightness signals: R, G and B, determine respectively the presence or absence of red, green and blue colours in all triades synchronized with line and frame scanning.

These three signals have the following binary digits:

0 means that the respective colour is missing;

1 means that the respective colour is available.

Depending on these digits the following combinations exist for the absence or presence of the three colours:

R	G	B	colour
0	0	0	black
0	0	1	blue
0	1	0	green
0	1	1	green-blue
1	0	0	red
1	0	1	purple
1	1	0	yellow
1	1	1	white

The LSP are passed to the horizontal time-base oscillator (4), which produces saw-tooth voltage for the horizontal scanning (time-base). Similarly, FSP are passed to the vertical time-base oscillator (5), which produces the saw-tooth voltage for vertical scanning (time-base).

R, G and B signals are passed to the video amplifiers (1), (2) and (3), which amplify them to levels required for controlling the electron beam intensity. The amplified signals are passed to the control electrodes of the three guns and as a result they stop or transmit respectively the red, green and blue beam. In this way the lighting and blanking of each dot in the triade is controlled.

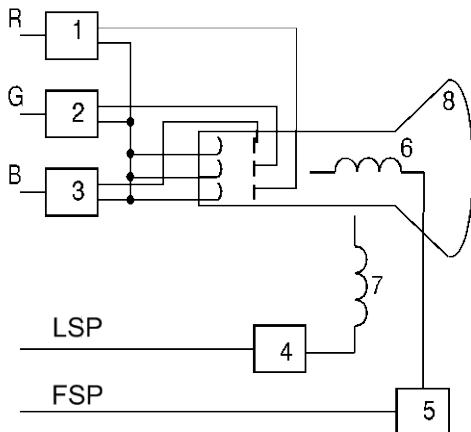


Fig. 3.22. Block diagram of colour video monitor:

1 – video amplifier for red colour; 2 – video amplifier for green colour; 3 – video amplifier for blue colour; 4 – horizontal time-base oscillator; 5 – vertical time-base oscillator; 6 – vertical deflecting coil; 7 – horizontal deflecting coil; 8 – colour CRT.

The colour video monitor appearance is similar to the appearance of the black-and-white video monitor and has the same control elements.

3.4. Vector displays

Vector displays are devices for outputting graphic images from the computer by CRT, whereat the image is built by a controllable coordinate motion of the electron beam in correspondence with the vector presentation of the graphic object. [1] Unlike raster displays, in the vector displays the beam is positioned at a dot on the screen with arbitrary set coordinates. The beam motion is controlled by an electromagnetic deflection system such that its projection on the screen within a command, is a straight line (vector). The beam brightness control does not depend on its motion. In this way, by executing commands for vector motion and turning on/off of the beam, the whole graphic image of the processed object is built sequentially.

A vector display (Fig. 3.23) is controlled by a display processor which processes the commands on the display file and sets the coordinates of the next beam motion and the intensity value in digital form. The coordinates are converted from digital quantities to analogue ones and thus the new value of the deflecting electromagnetic field is set.

Vector displays have high resolution which is the main advantage over raster displays but they are more expensive electronic devices. Therefore, vector displays with conventional CRT are rarely manufactured. The most widespread vector displays are based on memory CRT. Since critical oscillation frequency does not exist in them, the time for image building is not critical. Therefore lower requirements for fast operation are set to the electronic nodes for electron beam motion control. Therefore, vector displays can be manufactured using more reliable and less expensive elements and at the same time much greater resolution is reached (4096×4096 dots).

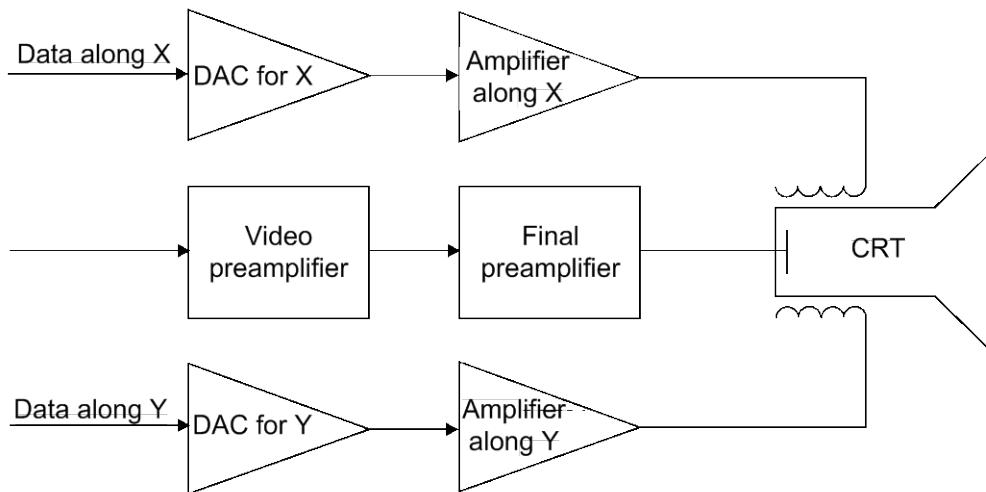


Fig. 3.23. Vector display control

3.5. Graphic terminals and stations

Graphic terminals and stations are autonomous operating technical complexes of a graphic system. The minimum technical complex for a graphic operator place comprises a graphic display, interactive means and alphanumeric and/or functional keyboard. When this complex operates remotely from the main computer, it is called graphic terminal. It is obligatory that it has a microprocessor control with internal software which as a minimum interprets the display file, serves the interactive protocol and makes the interface exchange with the main computer. If the local functions concerning graphic information processing in the graphic terminal are enhanced, it becomes ‘intelligent’. When the possibility to control additional input-output graphic devices: plotters, screen copiers, digitizers, etc., is added to this, a graphic station is formed. [1]

A typical structural diagram of a graphic terminal is shown in Fig. 3.24. Its major functions are:

- maintenance of computer exchange protocol;
- display file maintenance;
- performing basic graphic operations on the display file;
- frame memory maintenance;
- graphic monitor control;
- control of keyboard and interactive means.

According to the type of monitor used, terminals are monochromatic and colour. The display of colour causes multiple increase in frame memory (video RAM) and the volume of basic graphic processing. The greatest effect on hardware and hence on the price of a terminal (similarly of a graphic stations) has the resolution or the complexity of the image displayed on the screen.

Graphic terminals and stations are classified into several (three or more) classes depending on the following criteria:

- resolution;
- speed of graphic image generation;
- number of colours displayed on the screen;

- development of local graphic operations;
- availability of additional graphic input-output devices;
- possibility to work with three-dimensional graphic.

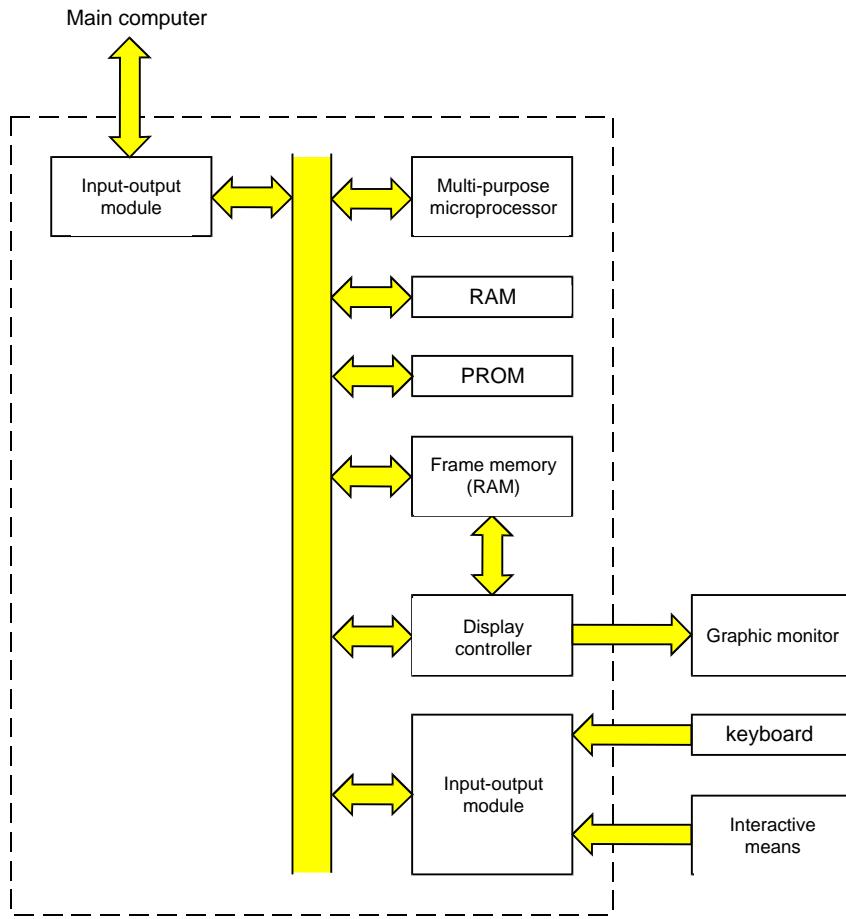


Fig 3.24. Structural diagram of a graphic terminal

The first criterion is decisive in the classification and has a direct or indirect effect on the remaining five criteria. The resolution is reflected in two secondary criteria: frame memory volume and video signal frequency.

Fig. 3.24 shows the specific functional blocks in the structure of a raster graphic terminal. The frame memory is in the first place. The whole information to build a whole frame of the image set by the display file is stored in it. Unlike the memory in which the display file (in vector form) is stored, the frame memory contains the raster equivalent of the image. For each dot from the raster image from one to several bits of information are stored. In monochromatic displays minimum one bit is necessary, while in colour displays the minimum number is three bits.

The frame memory as an independent functional block, in the general case, has two exchange channels. One of them is used by the graphic processor to load it, and the second serves for periodic frame information retrieval with a view of building and updating the image on the screen.

The frame memory is a main store device (RAM) with capacity, corresponding to the terminal resolution and the number of bits for each dot (pixel). There are terminals whose frame memory capacity is larger than the

required for building an image, with a view of more effective information processing in real time.

The display controller is a programmable control device which has already been designed by an integrated circuit. It consists of two functional parts. One of the parts synchronizes the scanning of the monitor and the frame memory, depending on the operational mode of the whole display tract, and the second part processes the information from the frame memory into a video signal (for each colour). The synchronization aims at periodically producing horizontal and vertical pulses for synchronous control of the raster monitor scanning. Simultaneously it is also necessary to retrieve information synchronously from the frame memory. Since in monitors with high resolution, the video signal frequency is tens of megahertz, and the access time to RAM corresponds to the retrieval frequency of 5÷10MHz, temporary buffering and parallel-serial converting of video information is necessary. The information from the frame memory is retrieved in portions of 8, 16 or 32 bits and then it is loaded in a shift register clocked with the video signal frequency. In this way, the capabilities of modern RAM integrated circuits join the necessity for a high-frequency video signal for generating a frame with a great amount of information.

Frame memory loading is an essential process in the graphic terminal. It is performed either by a general purpose microprocessor or by a special-purpose graphic processor. In the first case comparatively low computing velocity is achieved. Therefore, the tendency is to develop more and more perfect and faster graphic processors. The major function of the graphic processor is vector-to-raster conversion. The speed of building the new image directly depends on the time for fulfilling the function. It takes a long time to fulfill the function because the algorithms solve the problem not in relation to a group of dots in the raster which can be simultaneously loaded but in relation to a certain vector. The information about the dots which is loaded in the frame memory is located at different addresses and for each dot a separate record cycle in RAM, i.e. for loading a new image are needed as many addressings of the graphic processor to the frame memory as many active dots are in the raster image. The time for calculating the next active dot in the interpolating algorithm, when designing the graphic controller by a large integrated circuit, is reduced to a value many times smaller than the value of the access time to RAM.

Remember that along with loading the new image in the frame memory, the old image has to be erased. The algorithms for both operations can be equivalent and the time required for carrying them out are commensurable. This operation sets additional requirements to the computing velocity of the graphic processor.

Some technical embodiments overcome the reduced computing velocity caused by the necessity to erase the frame memory, by making it double or even triple. As a result, the operations of erasing, recording and displaying in the separate RAM modules become time compatible. Such construction leads to essential rise in cost of the device due to the relatively high price of a frame module with dimension $1024 \times 1024 \times 8$ (8 bits per colour). Therefore, along with building perfect in terms of image dynamics graphic terminals, by employing doubled frame buffers, efforts are made to enhance

the computing velocity of graphic processors and reducing the access time to RAM.

Other fundamental functions of the graphic terminal are the typical graphic processing of information in most intelligent graphic devices:

- scaling;
- translation;
- rotation;
- character generation;
- filling areas with colour shades;
- area hatching;
- moving a cursor as a result of the function of an interactive means.

3.6. Displays based on flat panels

When producing monitors for modern computers and for the new generation of television sets *cathode ray tubes* (CRT) are more and more rarely used. Displays with thin-film transistors have been preferred lately, owing to their advantages and reasonable price. These displays are known as *TFT LCD* (*Thin Film Transistor Liquid Crystal Display*). Colour displays based on thin-film transistors oust also displays with monochromatic liquid crystal indicator elements. An example of this are the displays of modern mobile telecommunication apparatuses. While monochromatic liquid crystal indicators made with a thermotropic liquid crystal or with liotropic liquid crystals only transmit or scatter the light passing through them, *TFT LCD* are light sources whose brightness can be adjusted depending on the illumination of the environment where they operate, or on the operator working with the display. These qualities determine their high consumer significance and the expanding applications in all fields of electronics.

The electro-optical properties of the liquid crystal placed in an electric field are used in *LCD*. Unlike other displays, *LCD* use light from external light source. Liquid crystals transmit or do not transmit light depending on the applied voltage and on temperature. Through controlling these two parameters within certain limits, bright and contrasting images are created.

LCD can be considered as a light key which blocks or transmits light. The image in *LCD* is formed by applying an electric field which changes the chemical properties of each liquid-crystal cell on the display, controlling the light-absorbing abilities of the pixels. These cells change the image formed by the light background on the screen and set by the controller. Although the output screen is coloured, liquid-crystal cells are monochromatic, and colour is added afterwards by a filtering process.

Operating principle.

In thin film transistor liquid crystal displays *TFT LCD* the liquid crystal substance is placed between two thin glass plates. Light passes through crystals according to the direction in which their molecules are oriented. Polarization filters control the light passing through them (Fig. 3.25). When

voltage is applied, the molecules of the crystal are aligned so that light falls perpendicularly on the polarization filter (at an angle of 90°). The voltage forces liquid crystals of each pixel to operate like a camera diaphragm – thus the filter either stops or transmits the light falling on it. The low power consumption, low weight and size, high efficiency, low level of electromagnetic radiation account for the high effectiveness of *TFT LCD*. [37]

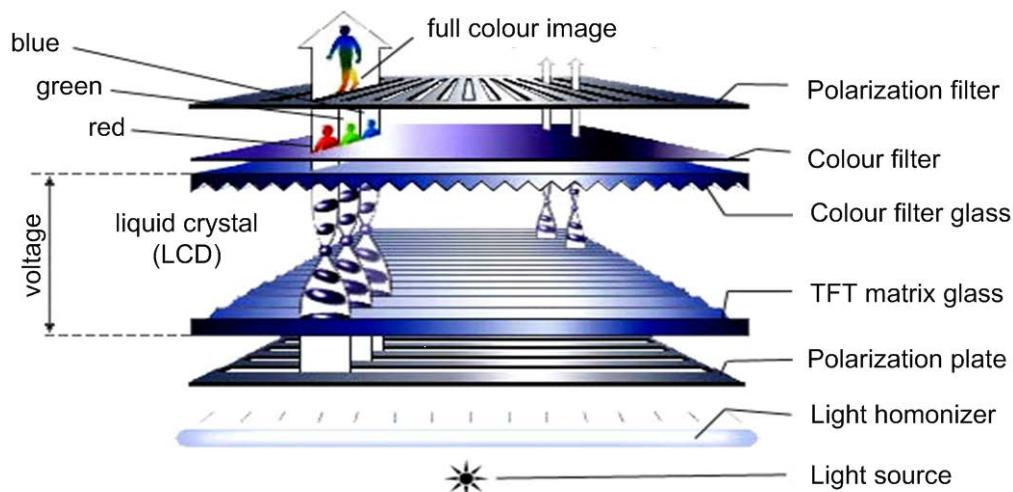


Fig. 3.25. Construction of TFT LCD

In order to explain the *LCD* operation, the light beam path has to be traced – from the light source to the viewer (Fig. 3.26).

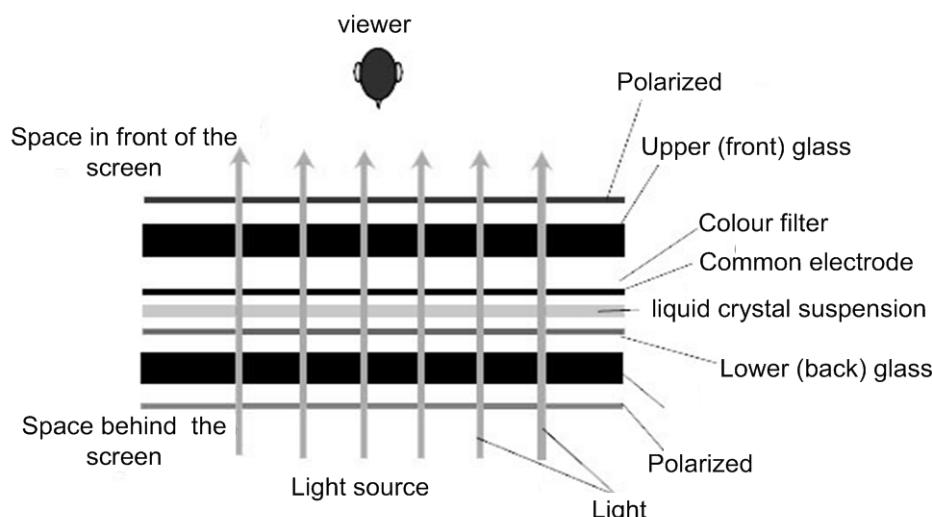


Fig. 3.26. Operating principle of TFT LCD

The light source is located just behind the *LCD*. From this source the light beam passes through an optical polarizer and enters the *TFT* matrix. Then the light beam passes through the *TFT* matrix and determines if a certain pixel is 'on' or 'off'. If the pixel is 'on', the liquid-crystal cell is electrically activated and the molecules in the liquid set out in one and the same direction. This allows light to pass without being changed. If the pixel is 'off', the electric field changes and the molecules disperse inside the liquid (Fig. 3.27), which automatically reduces the light passing through the pixel. In

colour displays, after the light passes through the *TFT* matrix, it passes also through a coloured filter (usually made of glass). This filter blocks the length of all light waves, except those within the range of the given pixel.

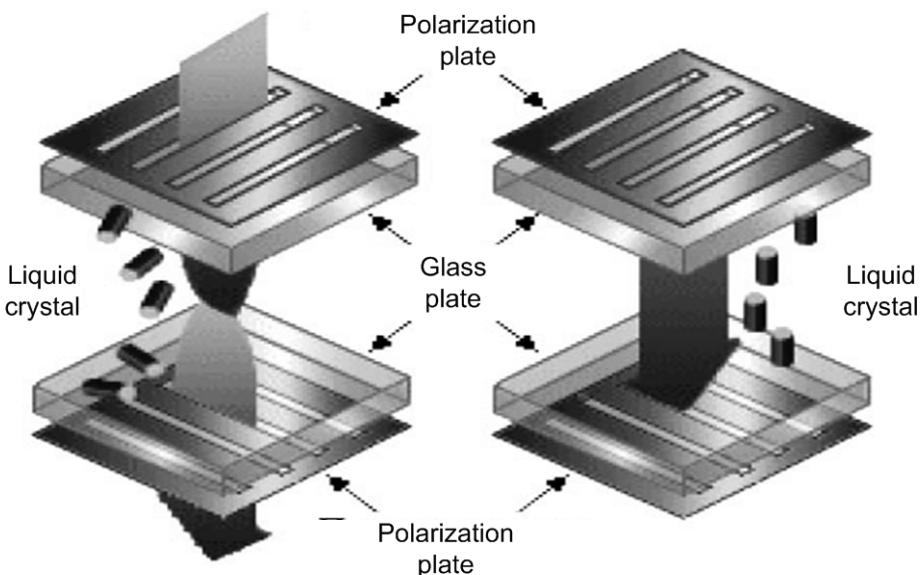


Fig. 3.27. Polarizing matrix of TFT LCD

The coloured filter is a liquid crystal panel in which red, green and blue elements positioned in a definite order, form the colour matrix (*RGB* matrix). The coloured filter is placed next to the top or bottom glass panel of the display. The three colour elements are activated autonomously, and their joint action results in a large number of colour shades (over 65 million colours). In the typical *RGB* display, the coloured filter makes a whole with the upper glass which is microscopically coloured and expresses each pixel – red, green or blue. The space on the filter between the coloured pixels remains black and thus contrast is achieved. Each beam of light passing through the coloured filter passes through another polarizer so that a clear image is produced. In this way the image contrast is enhanced and the luster is neutralized.

Types of matrices.

TN TFT or TH+Film TFT

The most widespread type of digital panel is based on the technology called *TN TFT* or *TH+Film TFT* (Twisted Nematic + Film). The term „+Film” means an additional external film which allows the habitual viewing angle of 90° (45° on each of both sides) to be increased up to 140°. *TN TFT* is the first officially presented LCD technology and it is still popular in medium and low panels of portable and desk computers. This is due to the low manufacturing cost of these panels. LCD panels, based on *TN TFT* matrix have some deficiencies. The most serious one is that black colour looks dark grey on the older panels and that means weak contrast. Throughout the years this technology has been improved and modern *TN* panels offer greater depth in displaying dark shades from grey to black. The second deficiency is that when a transistor ‘dies’, it leaves behind a light ‘dead’ pixel area. It appears on the screen and is much more noticeable than the one with black colour.

Super-TFT or IPS (In-Plane Switching)

This matrix has been designed to improve some of the shortcomings of TN TFT technology. This technology was developed by Hitachi. IPS allows the visual range angle to be expanded almost up to 170°, by employing a more precise method of control in arranging liquid crystals, which is the major contribution to this technology. In spite of that, contrast relations remain at the same level as in TN TFT technology and the response time is even longer. The positive side of this method is that ‘dying’ (burnt) pixels automatically become black unlike the white ones in TN TFT panels. This technology however has a chief disadvantage: when working in this system more energy is needed and it takes more time to switch the transistors. This results in an increase in monitor response time.

Multi-Domain Vertical Alignment (MVA)

This matrix is an heir to the previous VA technologies It was developed by Fujitsu and it is promising in terms of overcoming the main disadvantages of LCD panels. Its advantages are in the capability to improve the visual range angle and colour display. All colour elements divided into cells and zones, are positioned on the panel. They are formed on the lifted parts of the inner structure of the filters. The purpose of this design is to enable liquid crystals to move in direction opposite to their adjacent crystals. This allows the viewer to see one and the same degree of shadowing and colour quality regardless of the viewing angle. MVA has a great potential. One of its principal advantages is shortened response time. The complexity of the panel adds to the cost of LCD. Now the market share of MVA LCD technology is small but it is constantly increasing. At the moment MVA is the most advanced technical solution of LCD.

3.7. Other display technologies

PDP - Plasma display panel.

This technology is not entirely new. The study of plasma displays started in 1960 in the USA. Noble gas (e.g. argon) which is closed in certain space, is used in them. High voltage (several hundreds of volts) is applied by using electrodes. Thus the gas is transformed into the state of plasma. One of the important advantage of plasma displays over LCD and CRT, is the capacity to display a wider colour range. Another quality of plasmas is their large visual range angle, particularly compared to LCD technology. Brightness is another strength of plasmas. Among the shortcomings is the large size of pixels – in practice it is a problem to obtain size less than 0,5÷0,6 mm. [36]

Monitors with Organic Light Emitting Diodes (OLED)

According to experts, this is the technology which at a certain moment of its development will overtake TFT technology and will dominate. It will be employed not only in monitors and television sets, but also in mobile phones, digital cameras, etc. The advantage of this technology is that lamps are not necessary to illuminate the matrix at the back. Each pixel emits light controlled

by OLEDs which are separate pixels. The light brightness is determined by the current passed to the diode. [36]

LEP (Light Emission Plastic).

LEP technology was developed by the British company Cambridge Display Tehnology (CDT) about 5 years ago. It is one of the steps for gradual improvement of light emitting polymers. It is expected to become popular in a few years replacing the now existing mass TFT.

Field Emission Display (FED).

FED technology to some extent is similar to the processes in ordinary CRT monitors, since in both types of displays phosphor is used, whose particles fluoresce under the effect of an electron beam. This technology can be explained most easily as a combination of kinescopes and liquid crystals.

3.8. Advertising displays

The most important moment in developing advertising displays is to determine the way of displaying information and control. The image is built by luminescent elements arranged in a rectangular matrix with m lines and n columns. When the information is alphanumeric, each character is described by a matrix. Graphically it is presented in a so called shape table. In order to display a certain character, its address in the shape table is determined and the respective matrix is visualized.

Static and dynamic methods for controlling advertising displays are known in practice.

In static control a controlling element (trigger) corresponds to each indicator. Hence, for the control of the whole matrix, $m \times n$ controlling elements are necessary. The main deficiency of this approach to control is the high power consumption and high hardware and manufacturing costs.

Dynamic control involves raster principle of building the image. Its essence is based on the 'integrating' ability of human eye. When there is a change in the intensity of a light source with specified frequency, higher than the frequency critical for human eye, a viewer perceives this as stable without detecting any flicker.

By the raster principle, the whole work field is scanned sequentially until a complete frame is formed. For practical reasons, line scanning from left to right and from the top downwards is used. In advertising displays whole lines and columns can be selected without scanning the elements one by one.

It is necessary the refresh rate (frequency) to be higher than 45Hz to avoid a noticeable decrease in luminescent dot brightness. The effect of display element afterglow is also made use of here.

For better perception of long texts, it is necessary information to be moved horizontally from right to left. This allows greater amount of information to be displayed on comparatively small-size displays.

Light-emitting diode displays.

Fig. 3.28 shows a generalized block diagram of a light-emitting diode display with computer control for displaying alphanumeric and graphic information.

Relatively, the diagram is divided into three blocks according to their purpose: light-emitting diode display block, control and synchronization block and information input block.

The light-emitting diode display block is designed for information output in a suitable form. This block consists of a light-emitting diode matrix and two control blocks, by lines and by columns, respectively. Usually they are made up of triggers, registers or other store elements, connected to commutating elements (transistors) which supply current suitable for quality image.

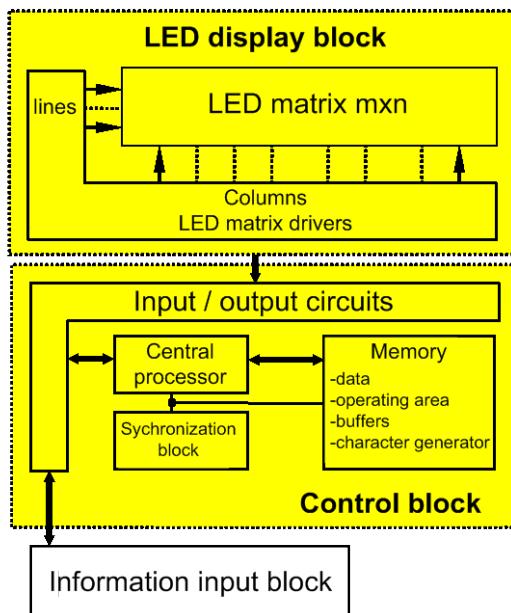


Fig. 3.28. Summarized block diagram of a LED display

information about the first column from the video memory, to record it in the buffer register and then employing the decoder, to select first column from the light-emitting diode matrix as well. In the same way the computer gets information about the second column, then about the third and so on until the whole matrix is entirely scanned. The frame formation rate depends on the matrix digit capacity. In longer display boards the time for building the image is longer leading to reduction in information refresh frequency (rate), which however should not be lower than 45Hz. Otherwise, the human eye can see blinking.

The light-emitting diode display has the following quality parameters:

- contrast – the operating state brightness versus non-operating state brightness ratio of an element;
- display dimensions – number of lines and columns;
- speed of display shift – number of renewed characters per unit of time (it is normal from 6 to 12 characters per second);
- type characteristics of light-emitting diode indicators.

Each frame can be formed by scanning in lines or by scanning in columns.

When building the image by columns (Fig. 3.29), it is necessary the single-chip computer to read the

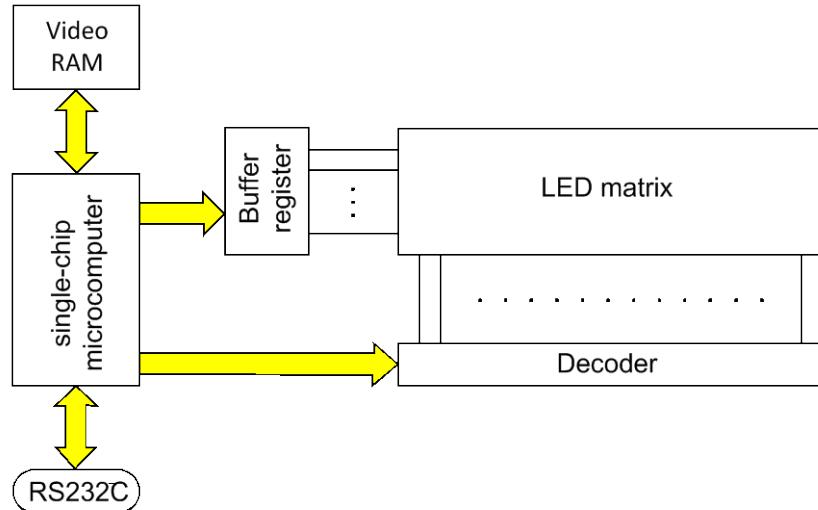


Fig 3.29. Frame formation by column scanning

When building large display boards, the line scanning method is adopted. (Fig. 3.30). Several shift registers are used in the diagram. These registers have serial input and parallel output. Line information is recorded in them controlled by a clock cycle. The visualization of the information recorded in the respective line is controlled by a decoder. When its first output is activated, the light-emitting diodes from the first line emit light with the content in the shift registers. After updating the information in the shift registers, the second line of the decoder is activated. The decoder determines the state of the next line in the display. In this way all display lines are scanned dynamically. By using shift registers instead of the traditional approach with buffers, a more economical solution is found. Fewer elements and simplified manufacturing facility is used.

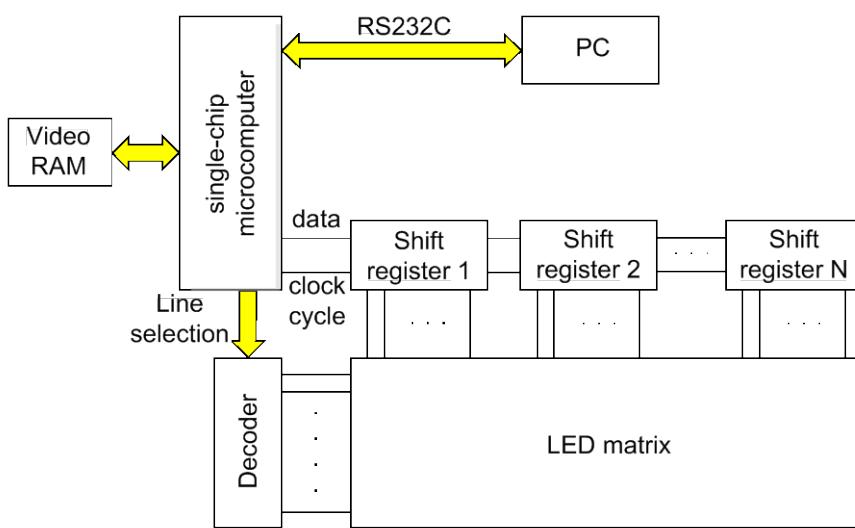


Fig. 3.30. Dynamic control of a light-emitting diode display by line scanning

Practically, the number of shift registers and the decoder digit capacity are crucial for determining the size of the dashboard light. The advertising file is prepared by a personal computer and by RS232 interface is sent to the device.

The single-chip computer receives the information from the personal computer and loads it in video RAM of the display. The data are transformed and recorded as shape tables. The approach to forming tables is traditional and it is based on matrix presentation of characters (Fig. 3.31). Coding is performed in lines. Binary (hexadecimal) code of luminescent dots forming the character corresponds to each line. The character matrix dimension is determined by the display board. The single-chip computer reads from the video memory and loads the shift registers with the content of the respective line from the display.

Fig. 3.32 shows an algorithm which illustrates the sequence and relations between the above-listed steps and stages in the dynamic control of light-emitting diode displays. Different advertising effects are included in it as well.

The figure consists of two parts. On the left is a 10x10 grid representing the shape table of character 'A'. The grid contains 'O' characters at specific positions to form the letter. On the right is a table showing the binary (hexadecimal) code for each row of the shape table.

		O	O	O					
	O	O			O	O			
O	O					O	O		
O	O					O	O		
O	O	O	O	O	O	O	O		
O	O					O	O		
O	O					O	O		
O	O					O	O		
O	O					O	O		

0011 1000	38
0110 1100	6C
1100 0110	C6
1100 0110	C6
1111 1110	FE
1100 0110	C6
0000 0000	00

Fig 3.31. Shape table of A character

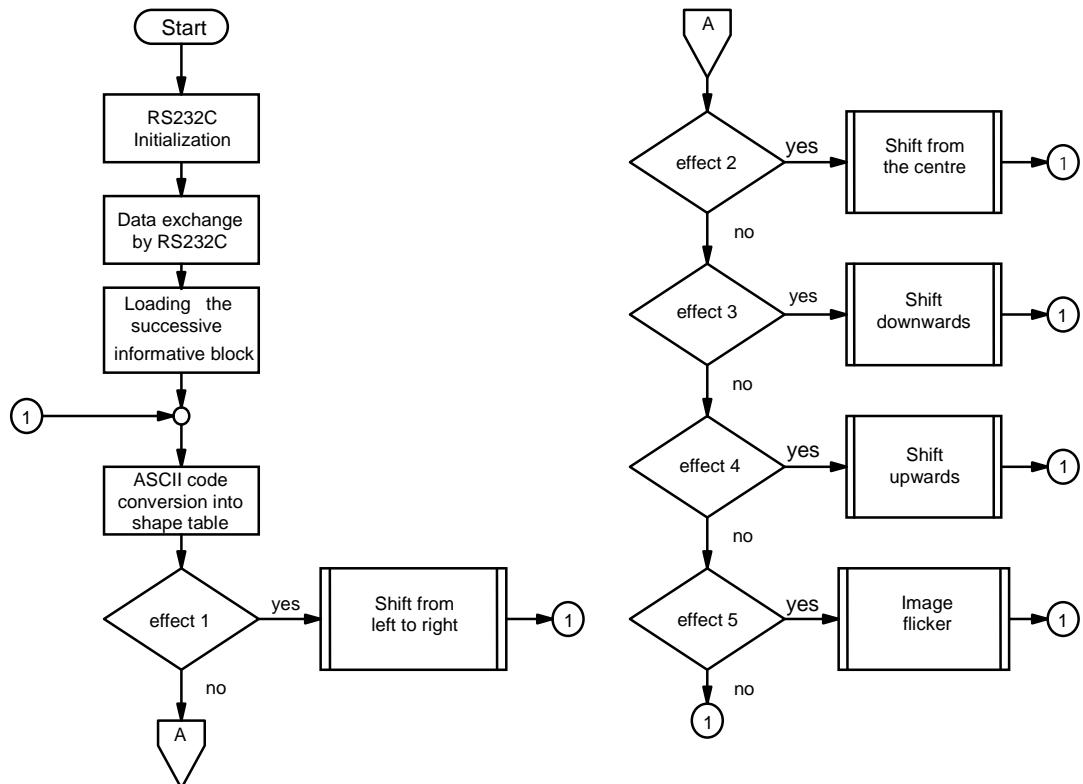


Fig 3.32. Algorithm of controlling a light-emitting diode display

Displays based on incandescent filament lamps.

Incandescent filament lamps are most frequently used for large-area displays. Taking into account that these indicators have long turning on and off time, high power consumption and inductivity, static method of control is suitable. When the device is designed for displaying alphanumeric information, the suitable matrix for writing characters is 5×7 . This economical way allows text visualization using either capital letters or small letters. The static control requires a memory element for each light dot. The most simplified approach is to synthesize a shift register with a serial input and parallel outputs for each line. According to the chosen matrix 5×7 a configuration of 7 shift registers is formed. The number of store elements in the registers is determined by the display length. The incandescent filament lamps have to give light in the yellow spectrum, in order to be seen in the daytime. For example, if they are chosen to have voltage of 48V and power rate of 10 W, they will give satisfactory light in the yellow spectrum using power supply of 36V. To simplify the control circuit, the lamps should be connected directly as collector resistors of the controlling transistors, as it is shown in Fig. 3.33. A PC is employed for controlling the whole advertising unit. It allows advertising information to be loaded by a text editor, writing in Cyrillic and Latin alphabets, easy correction and inserting additional information.

The Centronics interface of PC is suitable for transmitting information from the computer to the shift register. Data lines D0:D6 are connected to the shift register inputs. By STR signal a shift clock cycle is generated. The AUTO LF output controls the reset of the lamps (registers). The algorithm by which the device is controlled is shown in Fig. 3.34. It allows an advertising file to be formed, from which information is read and sent in 7-bit columns. For each 5 columns a column with zero content is generated which is necessary to separate two adjacent letters (the matrix is 5×7). Naturally, after each column there is time delay which forms the shift rate. At the end of the algorithm the file end is checked. In practice, the whole program is a cycle which displays the advertising file until it is not deliberately interrupted by the operator.

When producing software, character generators are also synthesized for the symbols with which the messages have to be sent by the 5×7 matrix. Besides Cyrillic and Latin letters, various other characters can be synthesized as telephone, arrows, numbers and other symbols. Huge equipment of this type is used for displaying information on stadiums, city squares, etc. [24]

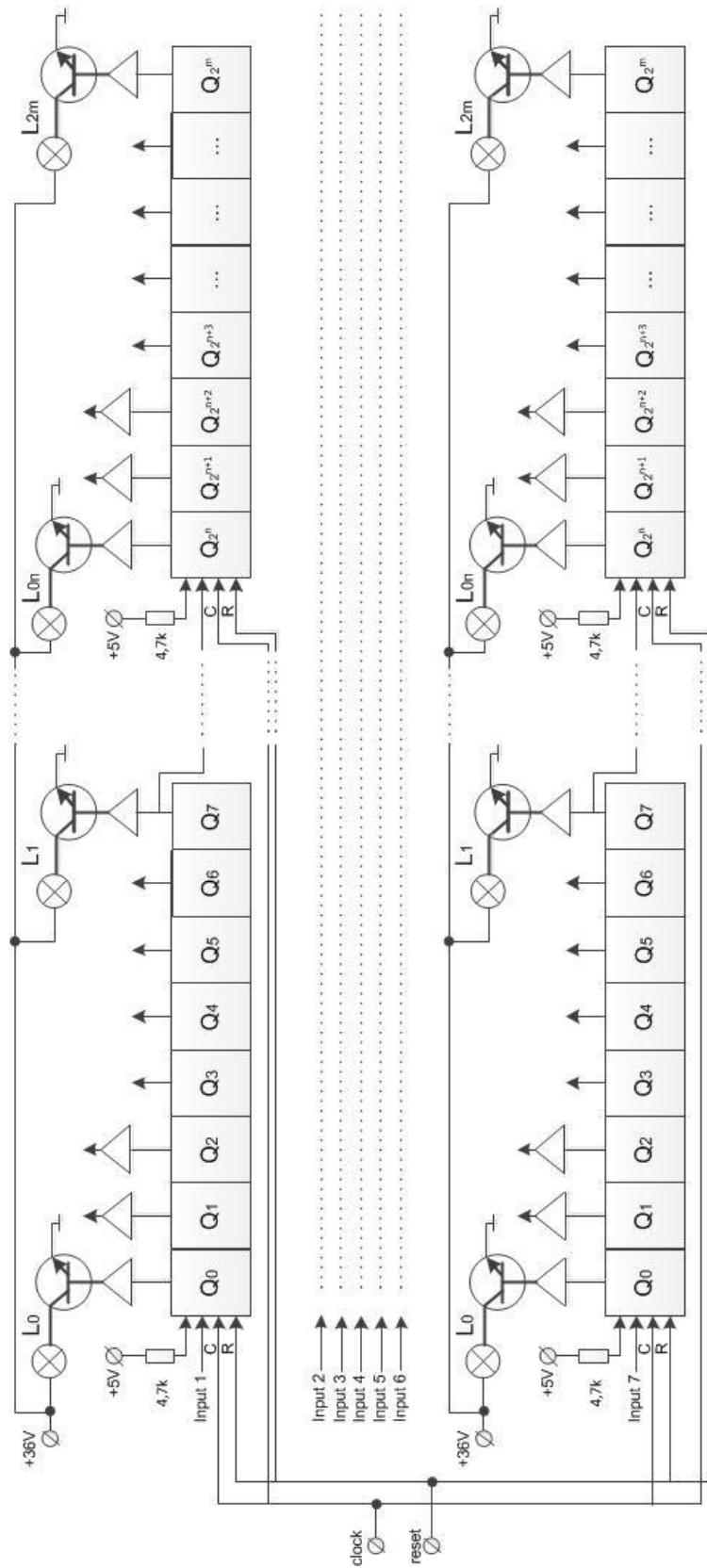


Fig 3.33. Displays based on incandescent filament lamps

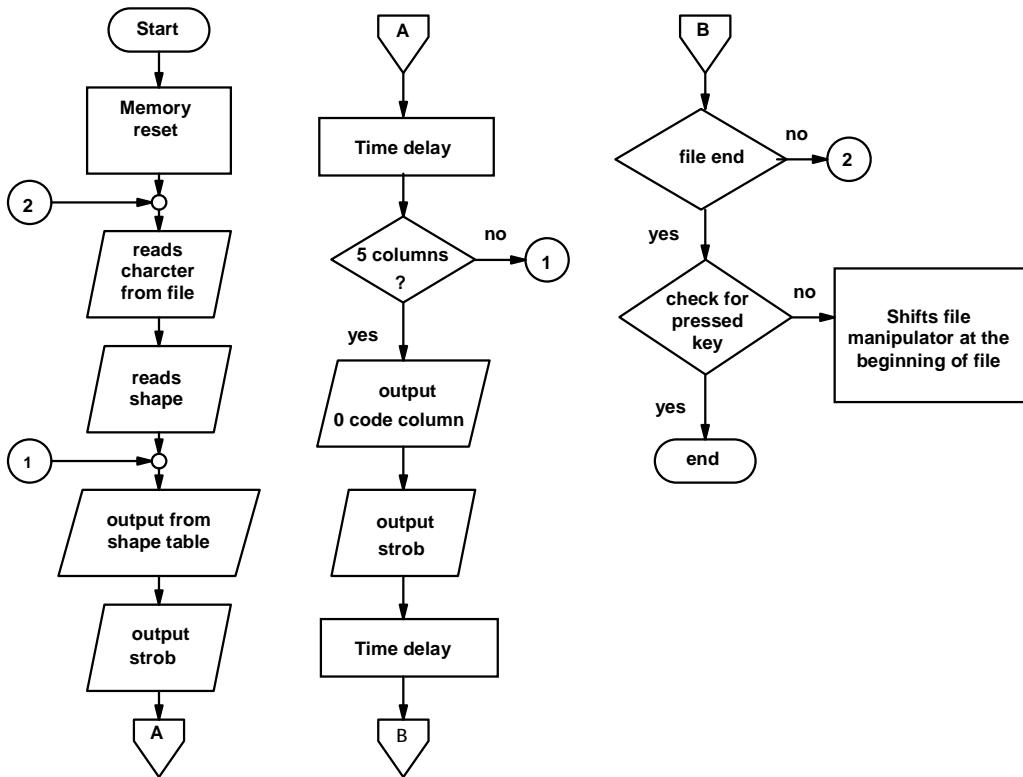


Fig. 3.34. Algorithm for light display control

3.9. Multimedia projectors

A projector is a device which is used for projecting images created in a computer (Fig.3.35). It is usually connected to the monitor port but models exist with a USB connection, analogue video port, DVI port, memory card port on which presentation can be recorded. There are even models with a wireless connection. The projector is most frequently used for presentations, training, film projections, advertisements, etc. [39]



Fig. 3.35. Multimedia projector

The oldest technology used in multimedia projectors is based on CRT (Cathode Ray Tubes). The construction and operating principle is almost the same as the cathode ray tube of ordinary monitors. Three emitters of red, green and blue colours are used, controlled by RGB signals. This type of devices have minimum maintenance (unlike modern projectors which use expensive lamps).

LCD projectors use LCD light gate. It consists of a small LCD display positioned in front of a powerful lamp. Light passes through it and it acts as a filter for its beam. The image is focused through an optical system and then it is projected on the opposite surface. This is the simplest and widespread system affordable for home and business applications. Its greatest problem is the pixelizing effect which is reduced to minimum in the latest models.

DLP (Digital Light Processing) projectors are based on Texas Instruments' technology. They use one, two or three microlight emissions, called DMD (digital micromirror devices). Projectors which use only a single DMD, send each of the red, green and blue images one after the other synchronized by a colour wheel, as it is shown in Fig. 3.36. Single and double DMD versions use rotating colour wheels synchronized with the mirror refresh for colour image modulation. The major problem in single and double DMD variants is the visible 'rainbow' which people see when they move their eyes. The three-DMD systems do not have this problem. They employ three circuits operating in parallel, one for each colour – red, green and blue.

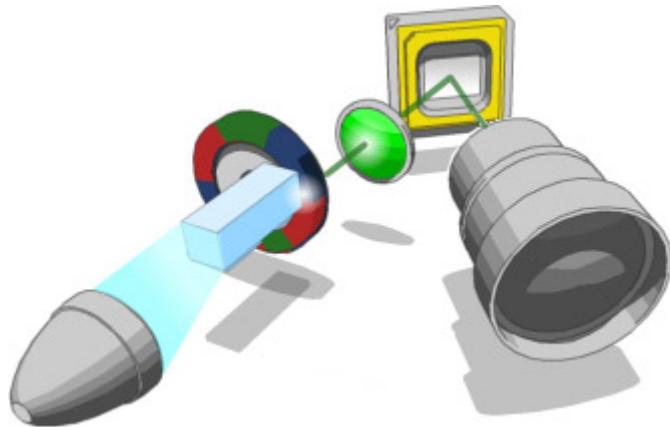


Fig 3.36. DLP projector with a single rotating wheel

LCoS - Liquid crystal on silicon projectors are based on DLP technology but they use liquid crystals instead of separate mirrors. The liquid crystals are placed directly on the silicon chip surface, soldered with an aluminized film which is extremely reflecting. LCoS technology can produce images with much higher resolution than the resolution of liquid crystal and plasma technologies.

One of the basic differences between various digital projectors is the number of panels. LCD and LCoS projectors almost always use one chip for each red, blue and green component of the image, although the company Phillips manufactured a single LCoS panel which provides the three primary colours. In three-panel projectors each panel is illuminated by the respective colour of light. Therefore the projector must have a colour-separating filter which separates the three colours from the white one and directs them to their respective panels.

Some DLP projectors also have three panels (3 DMD). This approach is applied mostly to systems for home usage. Professional devices use only one panel. In single DMD projectors the red, green and blue information appears on DMD in quick sequence, so that the eye merges the components together in a complete colour image.

Another element in the digital projector is the projector lens which focuses the image from the panel on the projection screen. To achieve this with the required precision, a large number of lens elements is needed. (Fig. 3.37).



Fig. 3.37. Projector lenses

Most projector lenses have an aligning focus and magnification. In modern devices the lens can move upwards and downwards, to the left and right. This moves the image without damaging and distorting the picture. The focus, magnification and displacement of the lens are performed by a small motor under the control of the central processor.

An example of a generalized block diagram for control is shown in Fig. 3.38. [39]

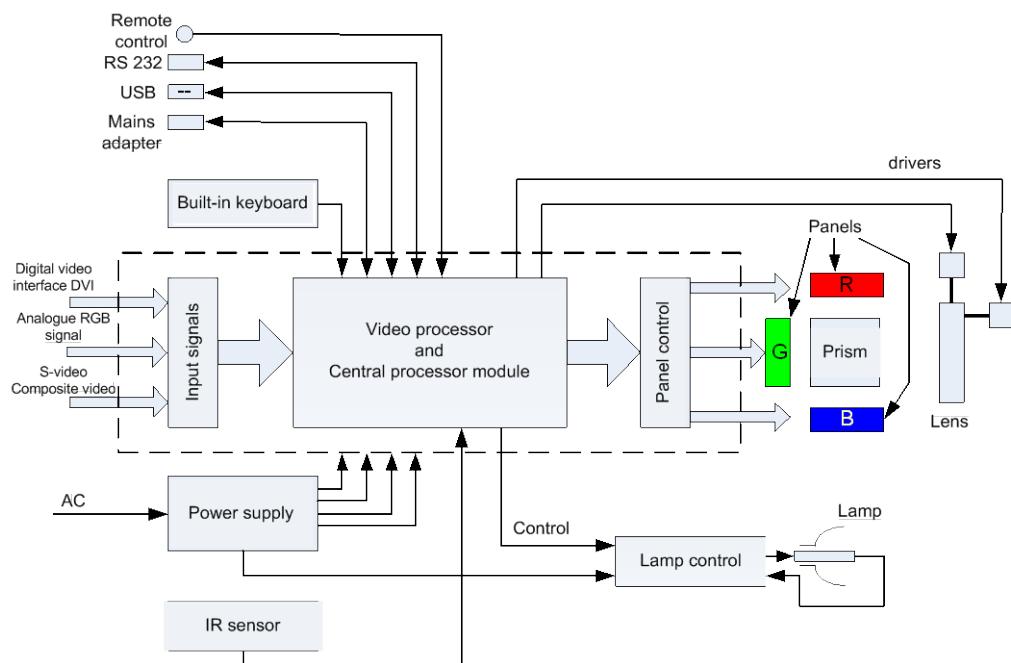


Fig. 3.38. Generalized block diagram for projector control

Important parameters when choosing a projector.

Brightness is a parameter showing the illumination of the image produced by a projector. The measuring unit is ANSI Lumens (ANSI - American National Standards Institute). It measures the screen illumination by the following formula:

$$\text{ANSI Lumens} = (L_1 + L_2 + L_3 + L_4 + L_5 + L_6 + L_7 + L_8 + L_9)/9$$

A screen with a standard coating is divided into 9 identical rectangles (3x3) and an entirely white image is projected on it. The illumination in the centre of each rectangle is measured in Lux. Then the arithmetic mean value of the nine rectangles is found and multiplied by the screen surface area which is given in square meters. The measuring result shows the difference in illumination in the various parts of the screen. It is normal the illumination in the screen centre to be higher than in the periphery, but in order to obtain a quality image, its whole surface must be evenly illuminated. The closer the L₁/L₅ ratio to one is, the better the image quality is. In technical specifications a parameter is given showing brightness uniformity, which is often ignored though it is important. It expresses the ratio of the most illuminated areas on the screen to the least illuminated ones. The value of this parameter must be over 80%, and values below 70% give low-quality image.

What brightness is required? The illumination provided by the most popular luminescent office lighting is of the order of 200÷250 Lux. To be projected a good quality image, the illumination on the screen should be at least by 50 Lux more, i.e. around 300 Lux. If the dimensions of the displayed image are 2 x 1,5 m, then its surface area will be 3 m² (300 Lux x 3 m² = 900 ANSI Lumens – minimum required illumination which the projector must provide). Some additional factors are also important and have to be taken into consideration:

1. The projector lamp loses some of its brightness after some period of operation. The life of the lamp is specified in terms of working hours and shows how long the lamp can work until it reaches 50% of its initial brightness. This means that before replacing the lamp, the projector operates with nearly twice lower brightness than its initial brightness.

2. Modern offices have large windows, and they are quite bright in the daytime. Daylight provides considerably better illumination than artificial lighting. Surface illumination supplied directly by the sun is of the order of 95 000 Lux, and at illumination angle of 50° - 60 000 Lux. This is the reason which does not allow images to be displayed by multimedia projector in the open (the most powerful projectors have brightness of 12÷15 000 ANSI Lumens).

Resolution is defined as the number of pixels which can be addressed along the horizontal and vertical direction on the screen, e.g. 800 x 600 or 1024 x 768.

The 4:3 format (the ratio of the width to the height of the image) is the most widely used format in computers and standard television. The manufactured models comply with the standards listed below and their relative share, expressed in percentage, versus the total production is given:

- SVGA (800 x 600) < 18%;
- XGA (1024 x 768) > 70%;
- SXGA (1280 x 1024) - only 2 or 3 models (non-standard format);
- SXGA+ (1400 x 1050) > 12%;
- UXGA (1600 x 1200) – only 2 or 3 models;
- QXGA (2048 x 1546) – only one model.

The 16:9 format is a wide-screen video format. Most DVD films, all materials in HDTV (High Definition Television – digital television with high resolution) are shown in this format. Specialized projectors for home video

and those for professional video halls also have the same format of display. They are manufactured in the following modifications:

- WVGA (854 x 480) – 15%;
- WSVGA (1024 x 576) – 8 %;
- WXGA (1280 x 720) – 68%;
- HD (1920 x 1080) – 8%;
- others: 960 x 540; 1366 x 768; 2048 x 1080; 4096 x 2160 – in several models.

Other parameters which should be taken into account when choosing a projector are: compatibility with different data standards and formats, interface capacities, quality of optics, additional presentation functions, network integration, equipment completion, reliability, etc.

4. INFORMATION INPUT DEVICES

4.1. Interactive Manipulators

Interactive manipulators are devices by means of which the display cursor is controlled. The **joystick** has the most simple construction of all. It operates on the principle of potentiometric function modeling. In its configuration it comprises two potentiometers, for X and Y coordinates, respectively (Fig. 4.1). [1] The cursor is controlled in two ways. In the first way the potentiometric signal is converted by means of an analogue-to-digital converter. Thus absolute values of the cursor coordinates are obtained.

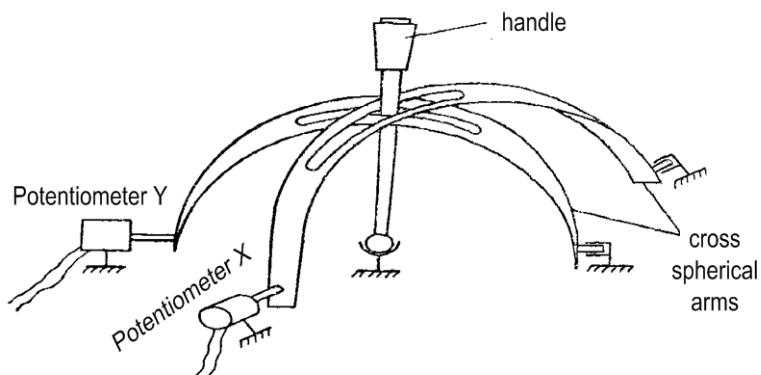


Fig. 4.1. Joystick

The second method is based on a RC (resistor-capacitor) circuit in which the potentiometer of the joystick for the respective coordinate takes part as variable resistance. In most devices the second method is used because it is less expensive.

The joystick cannot meet great demands for precise positioning. It is usually used in computer games, in devices for rough positioning as well as in more professional devices where instead of potentiometers, incremental transducers of angle-code type are employed.

Light Pen

The light pen looks like a pen with which the operator points at the object or dot on the screen which is of interest. When the electron beam of the video-monitor builds the object (once at each frame), the light-sensitive sensor in the pen sends an electric signal. This signal switches a trigger which either generates interruption or it is read by the program control of the video-controller. Since the video-controller controls the monitor scanning, it can determine the coordinates of the pointed dot. [18]

The light-sensitive sensor is a photodetector (photodiode or phototransistor) mounted in the case of the pen (Fig. 4.2). The light reaches it through an optical fiber. It is also possible to employ a photomultiplier but it works too slowly.

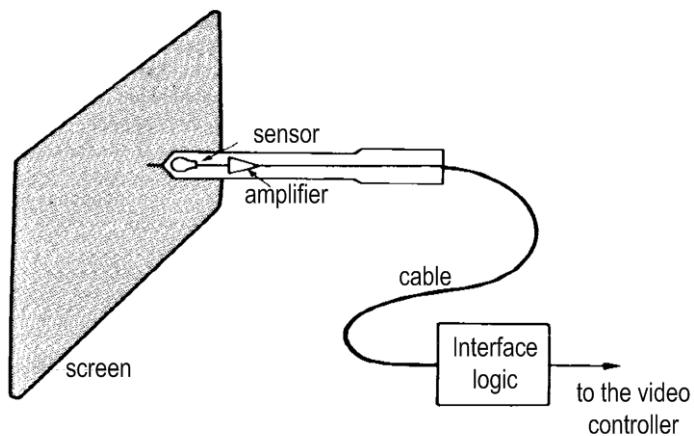


Fig. 4.2. Light pen

The light pen has to respond to the light from the screen and therefore it is necessary its sensor to be sensitive to the wavelength emitted by the phosphor of the CRT. It is required that the signal from the pen should be a little delayed for the correct operation of the device.

Nowadays, the light pen is rarely included in the display subsystem because it reduces the ergonomic qualities. When it is used, it is necessary the operator (the user of the system) to be very close to the display which is not recommended.

Optical Mouse

Interactive manipulators of this type are optical transducers of linear displacement. The transducer consists of two parts: a box (mouse) and a pad (Fig. 4.3).

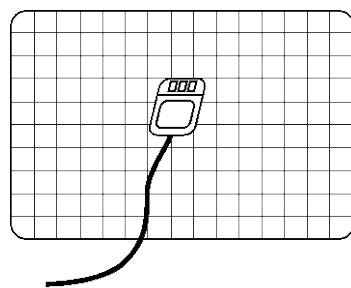


Fig. 4.3. Optical mouse

As a whole, the mouse is a movable detector means in which an optical-electronic transducer is placed. It slides on a mirror surface on which a raster grid in two colours is plotted. The colours are necessary for the recognition of the displacement and direction.

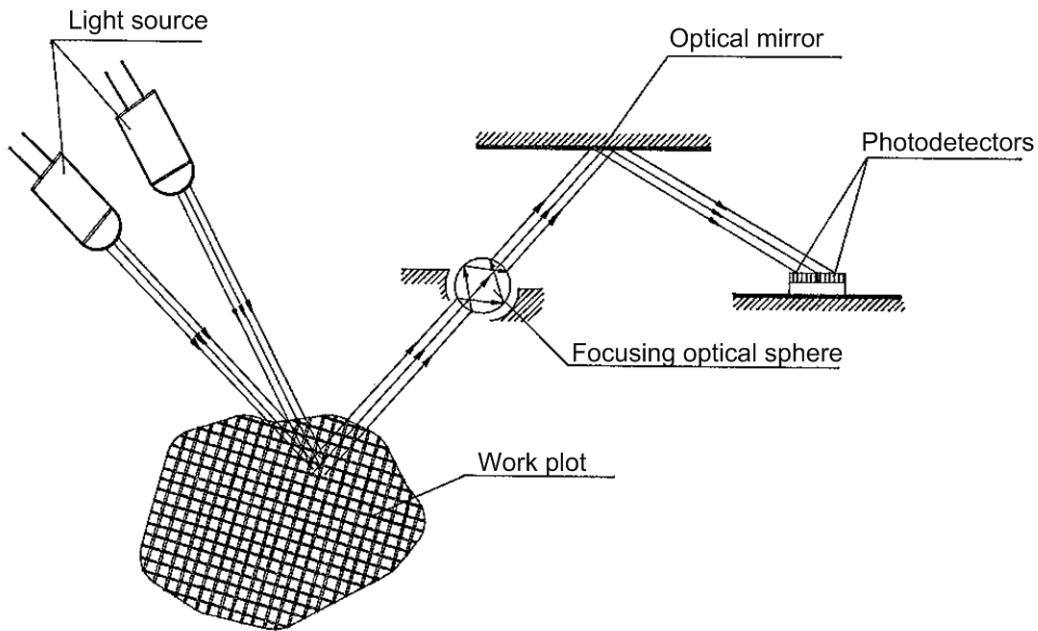


Fig. 4.4. Construction of the optical mouse

The optical system of the interactive manipulator (Fig. 4.4) incorporates the following elements:

1. Work plot. This is a metal surface on which two kinds of lines are drawn—horizontal and vertical (Fig. 4.5).

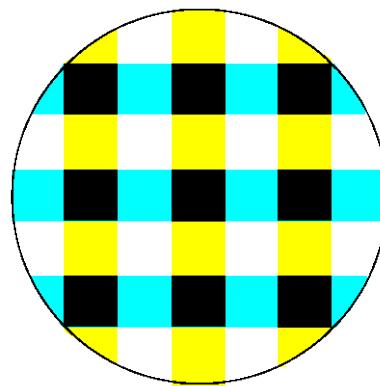


Fig. 4.5. Work plot of an optical mouse

The horizontal lines are coloured in red while the vertical ones – in green. The red lines reflect the red light and absorb the green light, and the green ones reflect the green light and absorb the red one. The pitch between the lines is constant and it is a metal reflection or something else with sufficient contrast. To produce symmetric signals in the output circuit, it is necessary the line width to be approximately equal to the pitch (interval) between them.

2. Two sources of monochromatic light. Their beams are shifted one towards the other at a distance equal to half the line width on the work plot.
3. Optical sphere.

4. Reflective optical mirror.
5. Photodetectors. They are positioned so that each of them can detect the reflected light beam emitted by the respective light source.

The light sources, work plot, optical sphere, reflective mirror and the photodetectors as a whole form an incremental photoelectric transducer of linear displacement.

The optical mouse contains two pairs of light sources positioned perpendicularly to each other. One of them emits monochromatic light with a spectrum determined by the colour of the horizontal lines, and the other one emits with a spectrum determined by the colour of the vertical lines. Light emitting diodes are used as light sources. They throw light on a relatively small spot on the work plot, which has a diameter approximately equal to the coloured line pitch.

The light signal reflected by the plot is focused by means of an optical sphere. Then it is reflected by an optical mirror and falls exactly on the photodetectors which are incorporated in matrices.

When the mouse is moved, each matrix element converts the reflected light lines into a sinusoidal signal which, with the help of a comparator is formed as a rectangular pulse sequence (Fig. 4.12). The signals are electrically dephased to one another at 90° . The phase difference between them is determined by the location of the two light sources for the respective group of lines (green or red).

The direction of motion of the mouse along the respective coordinate on the work plot is determined by the advance of one of the signals, and the magnitude of displacement is set by the number of pulses.

Modern Optical Mice

Microsoft and other companies implement optical technology for converting motion. These mice do not have mechanical parts and do not need a pad.

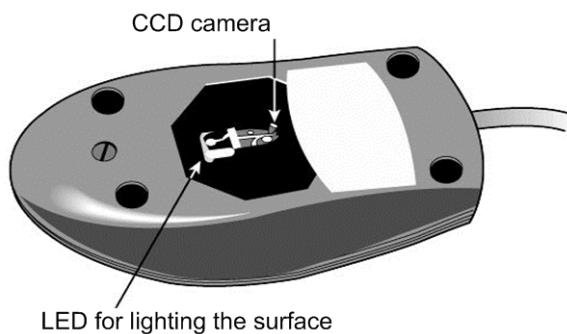


Fig. 4.6 Optical mouse – underside view

They work on almost any surface (Fig. 4.6). This is achieved by improving the optical sensor. Charge Couple Device (CCD) matrices are technologically included. They are sensors of a video camera. They specify the displacement, following the surface on which the mouse moves. A light emitting diode or laser diode is used to illuminate the surface. [33]

Owing to recent optical technologies and their affordable prices, the optical mouse is a good choice from the variety of models. All optical mice of this type have resolution not less than 400 dpi (dots per inch) and at least one sensor. The general view of the optical mouse is shown in Fig. 4.7.

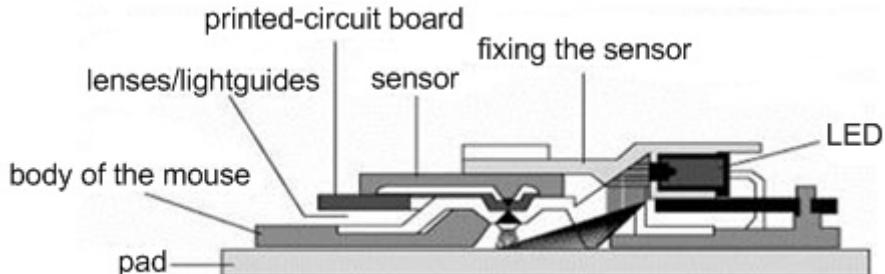


Fig. 4.7. Components of the optical mouse

The light emitting diode emits light which illuminates the surface under the mouse. The light is reflected by small microscopic imperfections on the surface. Lenses collect the reflected light and focus it on the sensor. The obtained image is a set of black and white spots reflecting the surface structure. The sensor sequentially takes pictures of the surface with the motion of the mouse. The pictures which the sensor takes are a lot (about 1500 pictures per second) and they are sufficient so that two successive pictures overlap regardless of the speed of the mouse motion (Fig. 4.8). The images are sent to the optical navigation processor to be processed.

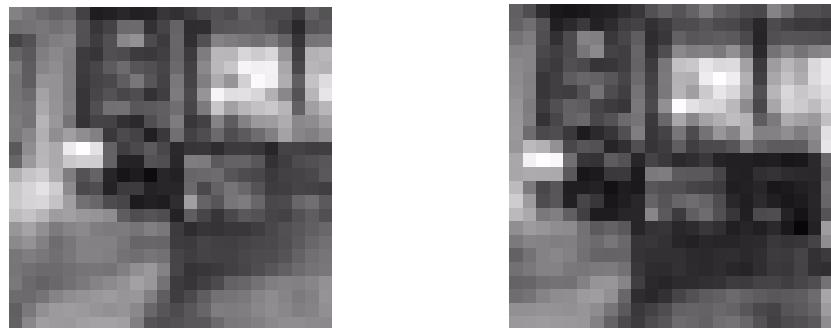


Fig. 4.8. Pictures from the mouse sensor at $t = 0$ ms and $t = 0,67$ ms

This is the basic processor for the optical mouse. It analyzes the surface imperfections and the other features and traces the motion through them. Two images are photographed in succession when the mouse moves to the right and upwards. Both displays show the same pictures. When the processor performs the algorithm, it finds the common things in the displays and determines the movement made. This information is translated into X and Y coordinates and sent to the computer.

Laser Mouse

Laser mice are optical mice. The only difference is that in laser mice a small infrared laser is used instead of a light emitting diode. It improves the resolution of the display and magnifies the mouse sensitivity to the surface by around 20 times. [40]

Optico-mechanical Mouse and a Trackball

The operation of these interactive manipulators is based on one and the same principle. The difference between them is in their construction.

The operating principle is explained by rolling a ball in a plane. Let the sphere is located at the origin of a chosen coordinate field and due to the effect of a certain force, it starts rolling (Fig. 4.9). The described trajectory of motion at each moment of time is determined by the current coordinates X_T and Y_T . If the motion is rectilinear, their ratio is expressed by the tangent of the angle which the trajectory has with the x-axis:

$$\frac{Y_T}{X_T} = \operatorname{tg}\beta . \quad (4.1)$$

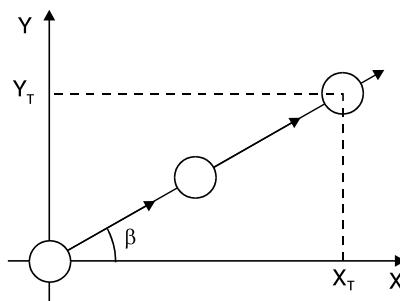


Fig. 4.9. Operating principle of trackball

The current values of X_T and Y_T coordinates are obtained in different ways, depending on how the rotary motion of the sphere is converted.

One of the ways of this conversion is illustrated in Fig. 4.10.

The diagram consists of a sphere (1) and two axis perpendicular to each other (2 and 3), and bearings lying in one plane. At the end of each axis a sensor is fixed (4 and 5), which converts the rotary motion into a digital code and has a gear ratio with the spherical surface. The operating principle is based on decomposition of the rotary motion of the ball along two axis perpendicular to each other (P_x and P_y).

Theoretically each of the axes is in contact with the spherical surface only at one point (A or B). When the ball rotates along the ordinate axis, rotary motion is transferred to the mechanical axis P_y , whose velocity depends on the gear ratio. The other axis P_x from the kinematic node slips at the point of contact A and remains static. The opposite case of rotation of P_x axis, when the sphere moves along the x-axis is similar. When the ball rotates in direction different from 0° or 90° towards the oriented coordinate system, the motion decomposes along two axes P_x and P_y . Each axis gets such rotation which is determined by the sphere motion in the coordinate field.

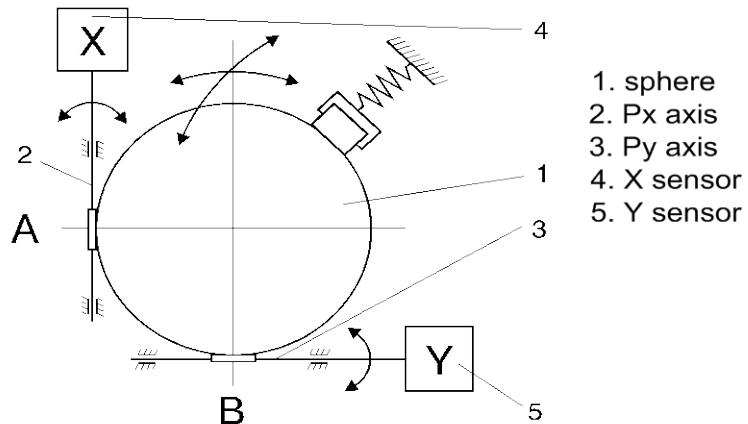


Fig 4.10. Conversion of the rotary motion of the sphere

The trackball (Fig. 4.11) is a kinematic system in which the sphere is rotated by the user's hand. In the optical-mechanical mouse the sphere is a small ball which moves when the manipulator is moved.

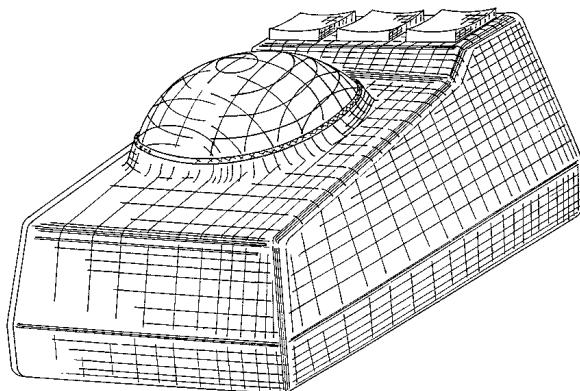


Fig. 4.11. Trackball

The sensors used are of angle-code type. In practice they are photo-raster transducers (PRT). Each of them generates two clock cycle sequences, electrically dephased one towards the other at 90° (Fig. 4.12). The φ_1 signal can lag behind or outrun φ_2 , which depends on the direction of sphere rotation.

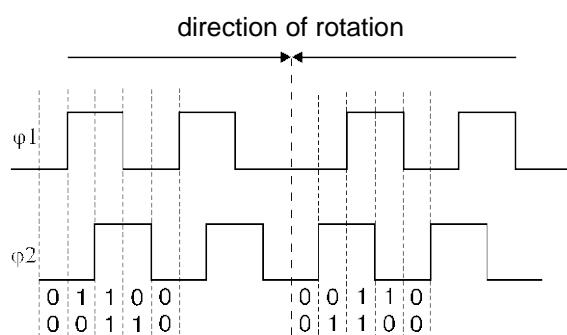


Fig. 4.12. Signals generated by the sensors

The photo-raster transducer is a device which comprises a raster grid and optoelectronic couple (Fig. 4.13).

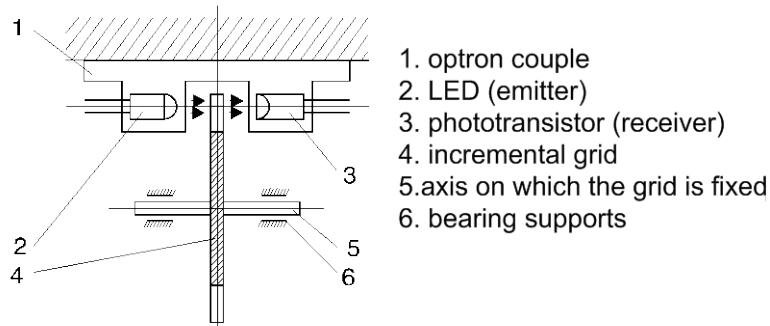


Fig. 4.13. Photo-raster transducer

The raster grid is a transparent disk on which dark spots with fill factor $K=0,5$ are applied along the radius. When the sphere rotates, the dark spots cross the light stream produced by the light emitting diode and as a result of this, the photodetector generates a clock cycle sequence of pulses.

To determine the direction of rotation when designing PRT, it is necessary to use two optron couples, positioned as shown in Fig. 4.14.

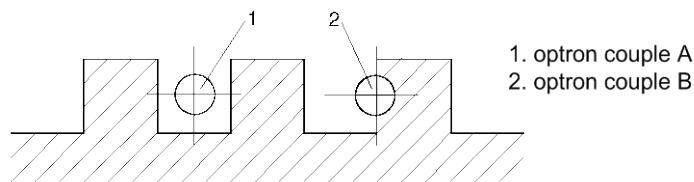


Fig. 4.14. Positioning of optron couples in PRT.

In the principle of converting rotary motion into a digital code, besides optron couples and raster grid, a circuit for output pulses formation is used (Fig. 4.15). It consists of an amplifier of the signal generated by the photodetector and a pulse former which converts the sinusoidal sequence into rectangular pulses needed for further digital processing. The number of output pulses is proportional to the magnitude of the angle of rotation, and the lead of one of the sequences towards the other and vice versa is determined by the direction of rotation.

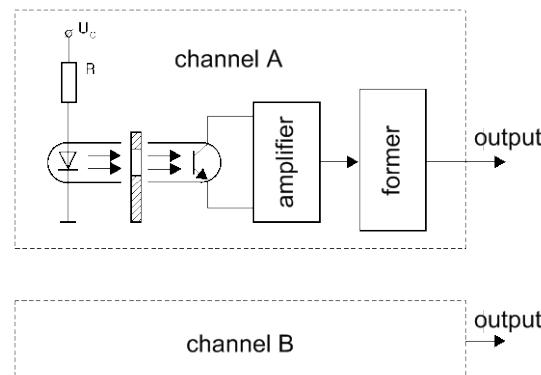


Fig. 4.15. Circuit for output pulse formation

When using PRT for the sensors (Fig. 4.10), during each sphere rotation the current coordinates (X , Y) are generated, expressed as a set number of pulses and the direction which correspond to the sphere displacement in the chosen coordinate field (Fig. 4.9).

The mouse and trackball can be presented in a block diagram as it is shown in Fig. 4.16. Each of them has two or three buttons which have their own functions during operation.

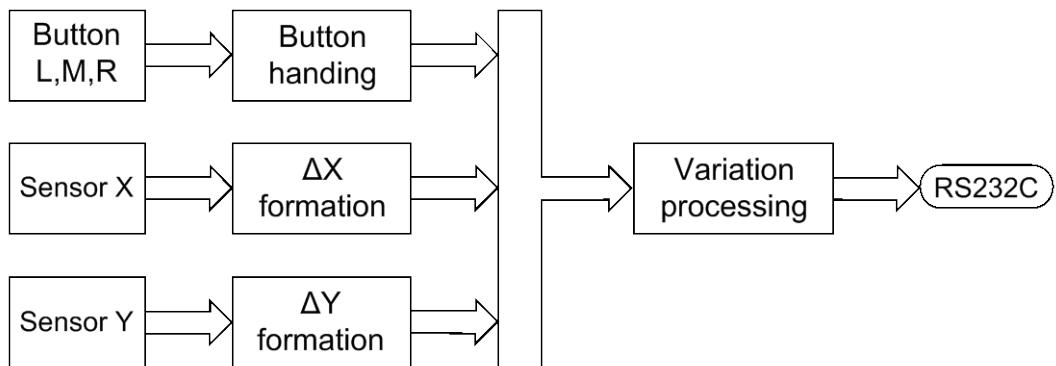


Fig. 4.16. Structural model of interactive manipulator of 'mouse' type

To form coordinates (X , Y) a graph diagram, shown in Fig. 4.17 is used. The graph nodes represent the discrete states of the two clock cycle sequences (Fig. 4.12). The transition from one state into another is expressed with an arc which defines the direction of rotation.

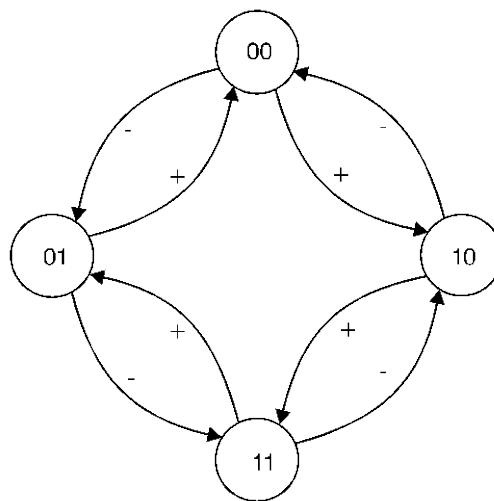


Fig. 4.17. Graph-diagram for determining the direction of mouse displacement

On the basis of the graph, Table 4.1 was made. An address is formed using the old and new state of each sensor. It gives the respective increase or decrease in the current coordinate by one as well as the lack of displacement.

BINARY ADDRESS		CONSTANT	COMMENT
old state	new state		
00	00	0	
00	01	-1	
00	10	+1	
00	11	0	forbidden state
01	00	+1	
01	01	0	
01	10	0	forbidden state
01	11	-1	
10	00	-1	
10	01	0	forbidden state
10	10	0	
10	11	+1	
11	00	0	forbidden state
11	01	+1	
11	10	-1	
11	11	0	

Table 4.1. Determining the direction of mouse displacement

Interactive manipulators of this type are controlled by a chip-computer. It serves the buttons, cyclically reads the state of signals of photo-raster transducers, forms relative coordinates and generates RS-232 output signal. The transfer protocol by RS-232 (Fig. 4.18) is three-byte for Microsoft Mouse driver and five-byte for PC Mouse. There is transfer when a button is pressed, released or moved.

Each protocol starts with the state of buttons and then relative coordinates follow. To serve the interruption by the interactive means, the personal computer employs a resident driver which receives the information and controls the display cursor.

Microsoft Mouse

Byte	b6	b5	b4	b3	b2	b1	b0
1	1	L	R	Y ₇	Y ₆	X ₇	X ₆
2	0	X ₅	X ₄	X ₃	X ₂	X ₁	X ₀
3	0	Y ₅	Y ₄	Y ₃	Y ₂	Y ₁	Y ₀

7information bits are used.

PC Mouse

Byte	b7	b6	b5	b4	b3	b2	b1	b0
1	1	0	0	0	0	L	M	R
2	X ₇	X ₆	X ₅	X ₄	X ₃	X ₂	X ₁	X ₀
3	Y ₇	Y ₆	Y ₅	Y ₄	Y ₃	Y ₂	Y ₁	Y ₀
4	X ₇	X ₆	X ₅	X ₄	X ₃	X ₂	X ₁	X ₀
5	Y ₇	Y ₆	Y ₅	Y ₄	Y ₃	Y ₂	Y ₁	Y ₀

Bytes 2 and 3 contain information about the previous displacement

Fig. 4.18. Transfer protocol.

Buttons L, M and R show the state of the left, middle and right button of the mouse, respectively. If the bit is 1 – the button is pressed, if it is 0 – it is released. The X and Y coordinates are in an additional code.

Modern mice use also a USB port for communication with the computer.

4.2. Digitizers

Various devices for graphical information input (DGII) are used which can be classified according to: methods of reading and registering a graphical image, the degree of automation of reading and operating principle.

According to the method of reading we can distinguish devices with continuous, discrete and continuous-discrete operation.

Depending on the degree of automation, two classes of digitizers are distinguished: automatic and semi-automatic. In the first case graphical information is read automatically while in the second case it is read with human participation.

Depending on the physical principle implemented for determining the coordinates of a given point, digitizers are subdivided into:

- electromechanical;
- optical-mechanical;
- potentiometrical;
- contact;
- acoustic;
- magnetostrictional ;
- capacitive;
- electromagnetic.

The electromechanical, optical-mechanical, potentiometrical, contact and acoustic physical principles for developing two-coordinate digitizers are not widely applied to the design of such devices and some of them are obsolete. They do not allow good technical performance to be achieved, and also their design solution is not effective. [14]

The most widespread digitalizing principles among leading world manufacturers of digitizers are: magnetostrictional, capacitive and electromagnetic.

The magnetostrictional principle is based on the long known physical phenomenon in which in certain materials a magnetostrictional wave can be excited. Its velocity of propagation is approximately equal to the velocity of propagation of sound in metals. The work plot of the digitizer operating on this principle, is made of magnetostrictional material consisting of a raster grid of thin strips or a sheet with the dimensions of the work format. A generator which produces a powerful current pulse excites a magnetostrictional wave which induces electro-motive force (e.m.f.) in an induction winding placed in the indicating means of the digitizer.

When the velocity of propagation of the magnetostrictional wave is known, it is easy to detect the coordinate point of the indicating means (sensor) location.

The advantages of this principle are simple control and lack of analogue apparatuses. At the same time, however, the magnetostrictional foil is not affordable and very expensive due to its amorphous metal structure. [1]

The capacitive principle uses the electrostatic effect. The work plot is formed as a grid of parallel conductors which contains two separate layers for

X and Y, and the indicating means is a capacitor plate. The simplest method for detecting the location of the indicating means is sequential activation of the conductors. The receiver registers pulsating voltage. The maximum amplitude is obtained when the strip is activated above which the indicating means centre is positioned. The strips are activated and then deactivated in one and the same sequence.. A cyclic sinusoidal signal is received in the receiver. After finding the phase difference between the beginning of the cyclic process in the strips and the one in the receiver, the address of the strip over which the indicating means is positioned can be determined very precisely. [1]

Capacitive digitizers have comparatively high resolution and precision. The electronic elements required for coordinate reading are commensurable with the ones of electromagnetic digitizers. Major shortcomings are sensitivity and a decrease in precision when the relative humidity of ambient air varies. The device has to operate in an air-conditioned room.

Electromagnetic principle. The physical process of electromagnetic induction is implemented in digitizers operating on this principle. According to it, an alternating magnetic field induces e.m.f. in a certain winding. The process takes place between the work field and the indicating means and each of them can be both an inductor and a receiver (Fig. 4.19). In both cases the plane of the winding is parallel to the digitizer work field and thus maximum mutual induction takes place. In order to determine the position of the indicating means, the absolute value of the induced voltage has to be found. The mutual induction varies by non-linear geometric law and depends on the relative position of the two windings. The inducing winding is a long conductor parallel to the receiving one, which is round with maximum small diameter, high quality factor and high internal inductance.

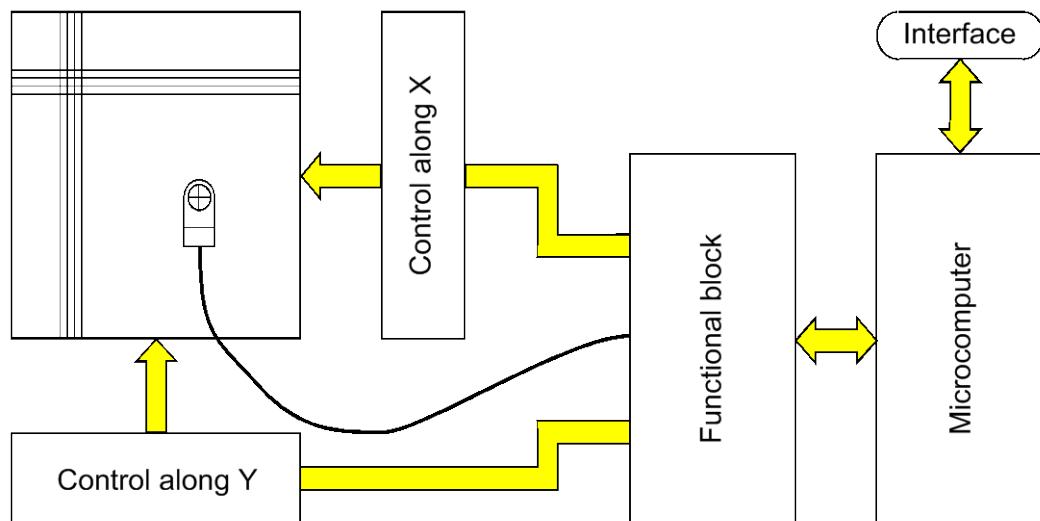


Fig. 4.19. Electromagnetic digitizer

The work plot is a rectangular grid of conductors arranged parallel along X direction and perpendicularly along Y.

Depending on the manner of conversion of the induced voltage, the conductors are positioned according to a certain law (linear, logarithmic, etc.). A common solution is by sectioning of the work field. Each section consists of

a lot of conductors arranged side by side. In this arrangement the reading is performed in two stages. First, the section in which the indicating means has fallen is read and then its position within the section..

4.2.1. Model of an Inductive Digitizer

The digitizing process in the inductive method can be considered from two aspects.

From the first aspect – the conductors arranged in the work plot are an inductor, while the coil (the position sensor) is a receiver. From the second aspect – the opposite – the coil emits, and the conductors are receivers.

The work plot of the inductive digitizer is a plane in which are positioned conductors arranged in parallel at constant pitch. Each of them can be considered as an infinitely long conductor through which sinusoidal current flows (Fig. 4.20):

$$i = I_m \sin \omega t . \quad (4.2)$$

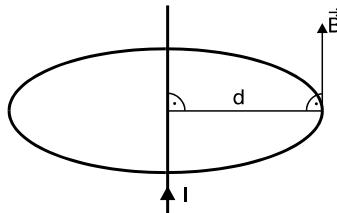


Fig. 4.20. Magnetic field around an infinitely long conductor

The current sets up a sinusoidal magnetic field around the conductor:

$$B = \frac{\mu_0 I_m \sin \omega t}{2\pi d} , \quad (4.3)$$

where: B is magnetic induction;

μ_0 is magnetic permeability of air;

d is the distance from the conductor to the vector of magnetic induction.

In the sensor, consisting of a flat coil with n windings, an electromotive force is induced, proportional to the change in magnetic flux:

$$U_d(t) = -n \frac{d\Phi}{dt} , \quad (4.4)$$

where: Φ is the flux through one winding of the coil;

U_d is the induced voltage in the sensor.

In digitizers the sensor moves in a plane parallel to the plane of the conductors (Fig. 4.21).

The magnetic flux is a function of inductance and it is presented as a double integral on a surface determined by the coil area:

$$\iint_s \vec{B} \cdot d\vec{s} = \iint_s B_n \cdot ds = \iint_s B \cos \theta \cdot ds , \quad (4.5)$$

where B_n is the normal inductance of the coil.

Since the distance from the conductor to the coil is constant - y_0 , this formula can be reduced to a single integral with limits determined along X

axis. Therefore all components with current values are presented only along this axis:

$$d = \sqrt{x^2 + y_0^2}, \quad (4.6)$$

$$B_x = \frac{\mu_0 I_m \sin \omega t}{2\pi \sqrt{x^2 + y_0^2}}, \quad (4.7)$$

$$\cos \theta = \frac{x}{d} = \frac{x}{\sqrt{x^2 + y_0^2}}, \quad (4.8)$$

$$ds = 2h \cdot dx = 2\sqrt{R^2 - (x - x_0)^2} \cdot dx, \quad (4.9)$$

$$\Phi = \int_{x_0-R}^{x_0+R} \frac{\mu_0 I_m \sin \omega t}{2\pi} \cdot \frac{2x}{x^2 + y_0^2} \cdot \sqrt{R^2 - (x - x_0)^2} \cdot dx, \quad (4.10)$$

$$U_d(t) = -\frac{d(n\mu_0 I_m \sin \omega t) \cdot 1/\pi}{dt} \int_{x_0-R}^{x_0+R} \frac{x}{x^2 + y_0^2} \sqrt{R^2 - (x - x_0)^2} \cdot dx, \quad (4.11)$$

$$U_d(t) = -\frac{n\mu_0 I_m \cos \omega t}{\pi} \int_{x_0-R}^{x_0+R} \frac{x}{x^2 + y_0^2} \sqrt{R^2 - (x - x_0)^2} \cdot dx. \quad (4.12)$$

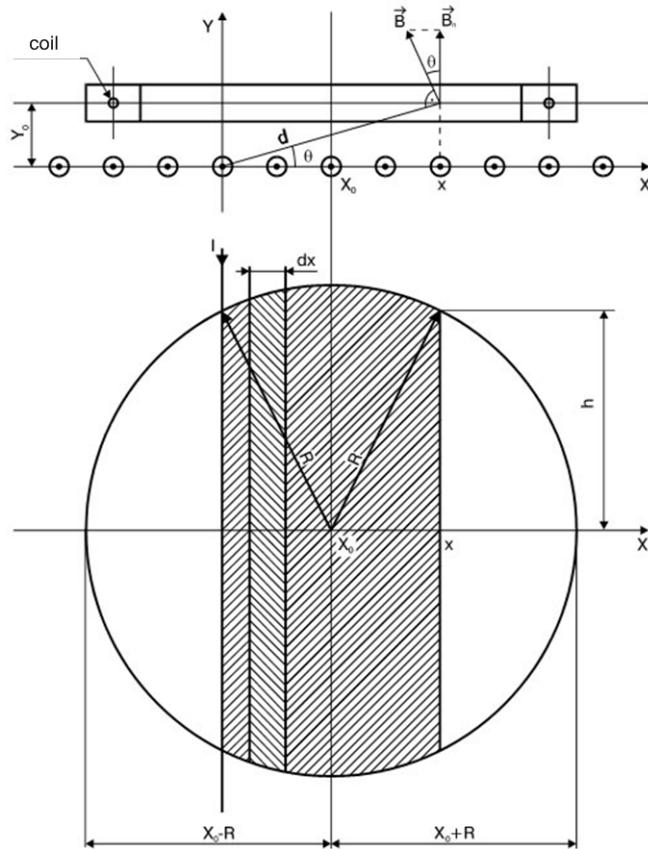


Fig. 4.21. Magnetic flux through the sensor

After solving the integral of the induced voltage in the sensor, it is obtained:

$$U_d(t) = U_0 \cos \omega t, \quad \text{when } x_0 \geq 0. \quad (4.13)$$

Since the characteristic is symmetrical, when written together it has the form:

$$U_d(t) = \pm U_0 \cos \omega t, \quad \text{for } -\infty < x_0 < +\infty. \quad (4.14)$$

The shape of the induced signal is shown graphically in Fig. 4.22.

When integration approach is implemented, it is possible to calculate the induced voltage as a function of its motion with parameters: the coil diameter and the distance between the winding and the conductors (Fig. 4.21).

From the second aspect, when the coil emits, an electromotive force is induced in the conductors having the shape shown in Fig. 4.22. This is proved by the law of mutual induction:

$$M_{12} = M_{21}. \quad (4.15)$$

The existence of a linear area is inherent to this characteristic. The area is expressed around the coil centre.

A deduction can be made that the straight line will be approximated best when the coil diameter is large and the distance to the conductors (Y_0) is small.

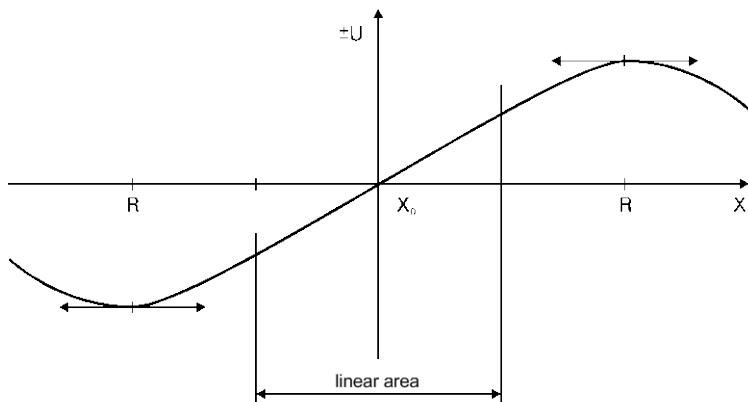


Fig. 4.22. Shape of the induced signal

Principle of Linear Interpolation

Let the position sensor is a round coil and sinusoidal current passes through its windings. Then an electric signal is induced in the conductors of the work plot as a result of the effect of the magnetic field of the coil (Fig. 4.23).

At moment t_n in conductor N voltage U_n is induced.

At moment t_{n+1} conductor $N+1$ is chosen. Then the signal induced in it is U_{n+1} . The interpolation value h is expressed in the following way:

$$\operatorname{tg} \theta = \frac{U_n}{h} = \frac{-U_{n+1}}{\delta - h}, \quad (4.16)$$

$$h = \frac{U_n}{U_n - U_{n+1}} \cdot \delta, \quad (4.17)$$

where $0 \leq h \leq \delta$.

It can be noticed that denominator ($U_n - U_{n+1}$) remains constant for all values of h , limited by pitch δ .

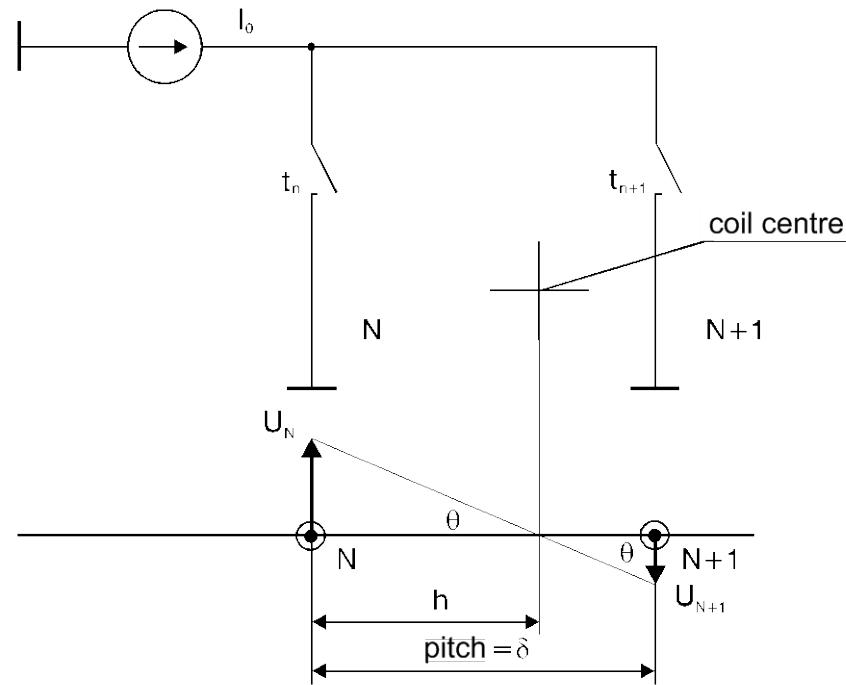


Fig. 4.23. Induced signal in the conductors of the work plot

Error in Linear Interpolation

The error is calculated for each position (x_0) of the position sensor by measuring the amplitudes of signals U_n and U_{n+1} . It is determined by the expression:

$$E = h - x_0 = \frac{U_n}{U_n - U_{n+1}} \delta - x_0, \quad (4.18)$$

where $a \leq x \leq \delta$ and a is the value of x , from which δ pitch begins.

In fact the error is relative because U_n is in both the numerator and denominator.

Practically, the error in the beginning and at the end of the pitch is equal to zero (Fig. 4.24). When $U_n = U_{n+1}$ the error is also equal to zero because the sensor is at equal distance from both conductors.

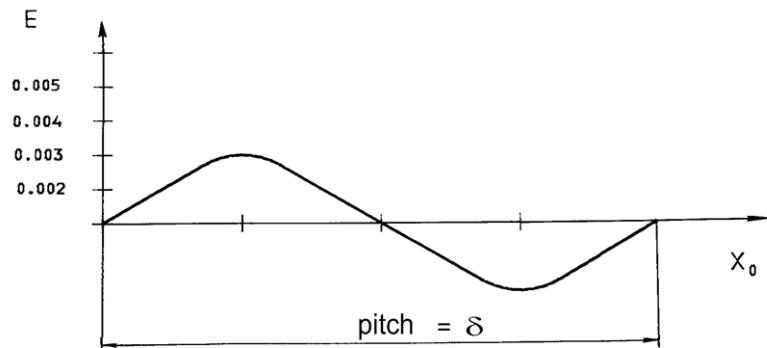


Fig. 4.24. Error in linear interpolation

Principle of Detecting N Conductors

Scanning of conductors placed on the digitizer work plot is shown in Fig. 4.25.

At moment t conductor 0 is chosen.

At moment t_1 conductor 1 is chosen.

At moment t_n conductor N , and at t_{n+1} conductor $N+1$ are chosen.

The N -th conductor is detected from the phase of the signal induced in the conductors. All conductors placed on the left of the coil centre correspond to negative motion, and those on the right – to positive motion. Hence, there is phase opposition (transfer from negative to positive phase) depending on the direction of switching the conductors (for instance, from N to $N+1$). By the detection of this opposition each conductor can be recognized. The address of the N -th conductor corresponds to N_p . Then the distance M_n from the beginning of the first conductor to the N -th is:

$$M_n = N_p \cdot \delta . \quad (4.19)$$

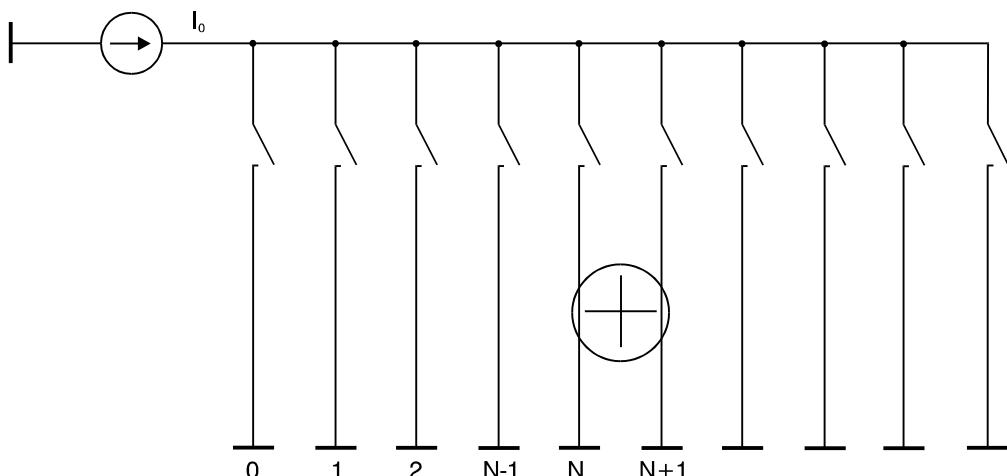


Fig. 4.25. Conductors on digitizer work plot.

The sensor location can be determined approximately as two fields of parallel conductors are situated so as to form a perpendicular grid. After determining the x coordinate of the sensor, the y coordinate can be found by a repeated scanning of the second field.

The coordinates (x, y) of the indicating means (sensor) are calculated in the following way::

$$\begin{bmatrix} x \\ y \end{bmatrix} = \delta \cdot \begin{bmatrix} N_{px} \\ N_{py} \end{bmatrix} + \begin{bmatrix} h_x \\ h_y \end{bmatrix} \quad (4.20)$$

Measurement accuracy is directly related to the accuracy of making the coordinate lattice. The best results are achieved by scanning a grid which is a printed-circuit board. However it has some deficiencies. It is impossible in practice to make an A0 format work plot and an identical error in the pitch of positioning the conductors to occur all over its surface

When forming a coordinate couple, both the error and each grid imperfection are taken into account. The principle is to divide the lattice into elementary surfaces and to connect each surface (N_{px} and N_{py}) to a compensating element..

This is described in this way:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \delta \cdot \begin{bmatrix} N_{px} \\ N_{py} \end{bmatrix} + \begin{bmatrix} h_x \\ h_y \end{bmatrix} + \begin{bmatrix} E_x \\ E_y \end{bmatrix}. \quad (4.21)$$

Components E_x and E_y for each elementary surface are stored in PROM. The advantage of this approach is that any type of error can be compensated for as far as the manufacture of the lattice is concerned. On the other hand, this requires a separate PROM to be recorded for each grid and measurements to be taken in advance with accuracy greater than the digitizer resolution.

Method of Switching N Conductors

When producing a digitizer with a larger format of the work field, the so-called block switching is employed (Fig. 4.26).

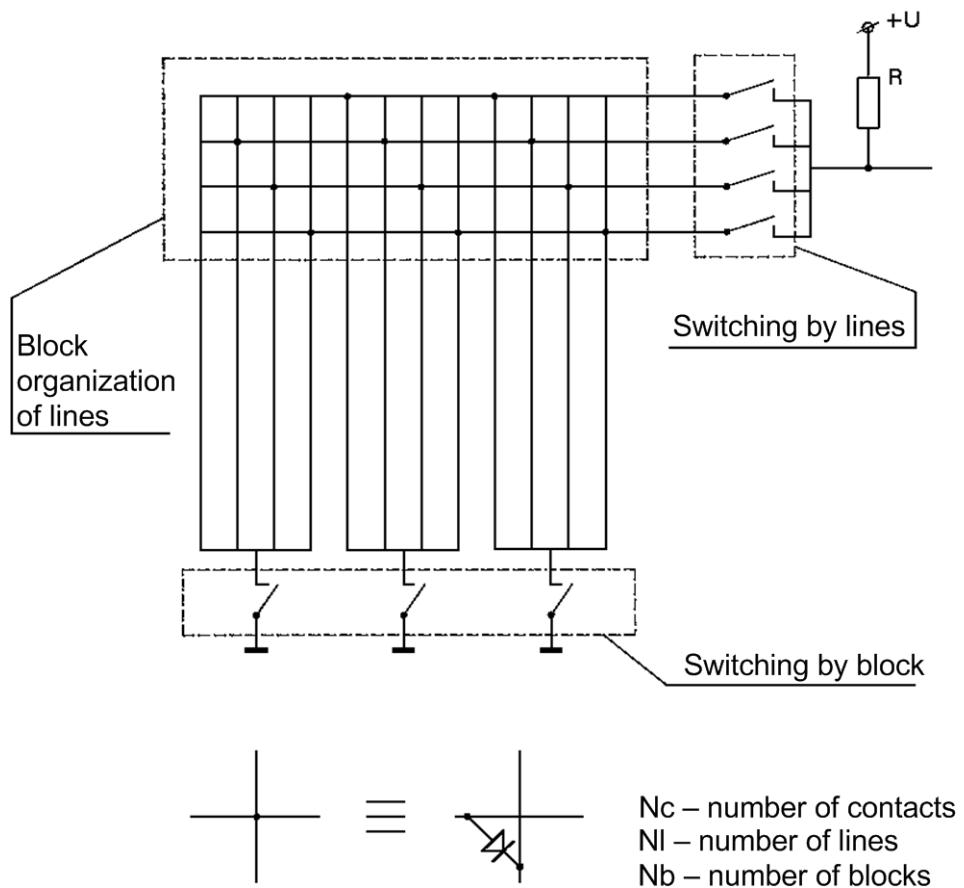


Fig. 4.26. Block switching of conductors

The grid is divided into blocks and an analogue switch is assigned to each block. Each conductor of a certain position (rank) in the block is connected to a strip of the same position (rank) by an opposite blocking diode. An analogue commutator is connected to each line. It is necessary to maintain the current positive at any time so that the lines connected in this way, can function. Thus the return between blocks is avoided.

For a given number of lines, there are a certain number of blocks which provide minimum analogue switches:

$$N_c = N_b + \frac{N_l}{N_b}, \quad (4.22)$$

where N_l is the total number of conductors according to the shape of the working area.

For the minimum number of switches it is obtained:

$$N_c = \min, \text{ when } \frac{dN_c}{dN_b} = 0.$$

Since $\frac{dN_c}{dN_b} = 1 - \frac{N_l}{N_b^2}$, then $\frac{dN_c}{dN_b} = 0$ for $N_b = \sqrt{N_l}$, (4.23)

and from this it follows that:

$$N_{c\min} = 2\sqrt{N_l}. \quad (4.24)$$

Resolution

Let the working field has Lx.Ly format, where:

$$\begin{aligned} L_x &= \delta N_{lx}; \\ L_y &= \delta N_{ly}; \end{aligned} \quad (4.25)$$

δ is the pitch between the conductors;

N_{lx} is the number of lines (conductors), forming the X coordinate;

N_{ly} is the number of lines (conductors), forming the Y coordinate.

Then resolution m is determined in the following way:

$$m = \frac{\delta}{F_m}, \quad (4.26)$$

where F_m is the number of pulses generated from moment t_i to moment t_{i+1} (Fig. 4.27).

Interpolation value h is divisible by pitch δ .

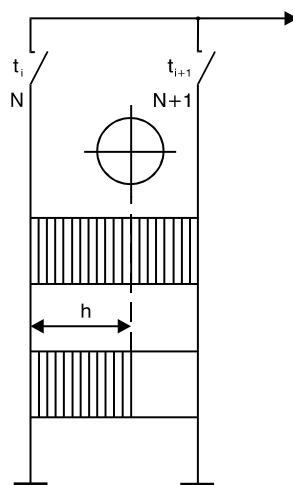


Fig. 4.27. Resolution of digitizer

Summarized Block Diagram of an Electromagnetic Digitizer

Digitizer block diagram is shown in Fig. 4.28. In this case, the position sensor emits a magnetic field. As a result of this, e.m.f. is induced in the conductors of the work plot.

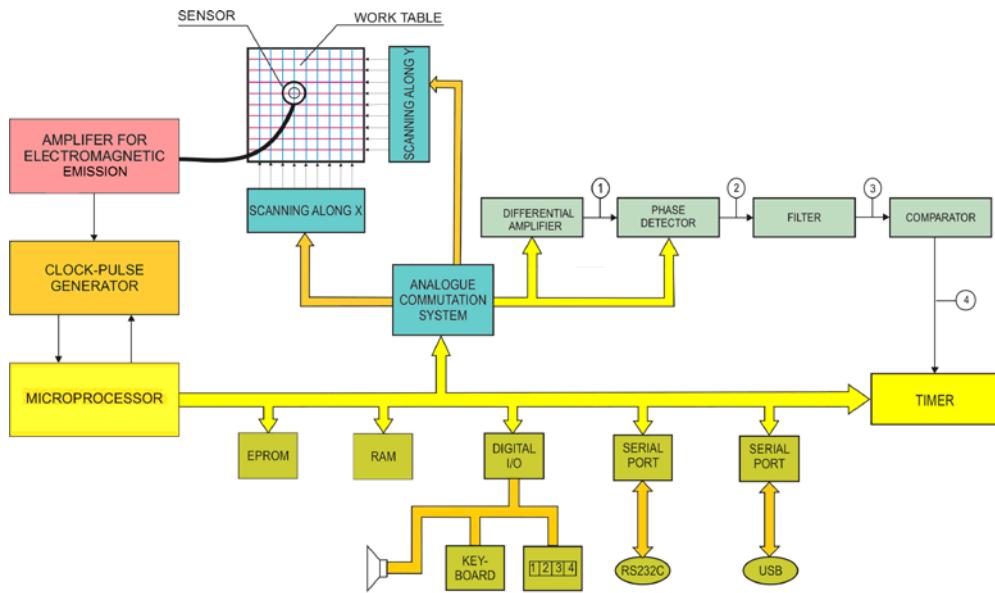


Fig. 4.28. Summarized block-diagram of a digitizer

After amplifying the e.m.f. by the differential amplifier a symmetric signal is obtained (Fig. 4.29) which is further detected by the phase detector. Passing through the filter, the modulating frequency part (envelope) of the signal is separated. The modulating frequency part (envelope) is a low-frequency signal which is also obtained by mathematical transformations as well (Fig. 4.22). The comparator is activated when the signal passes through the zero in order to stop the position counter. The work plot scanning is allowed by the rectangular sequence of pulses shown in the upper part of Fig. 4.29.

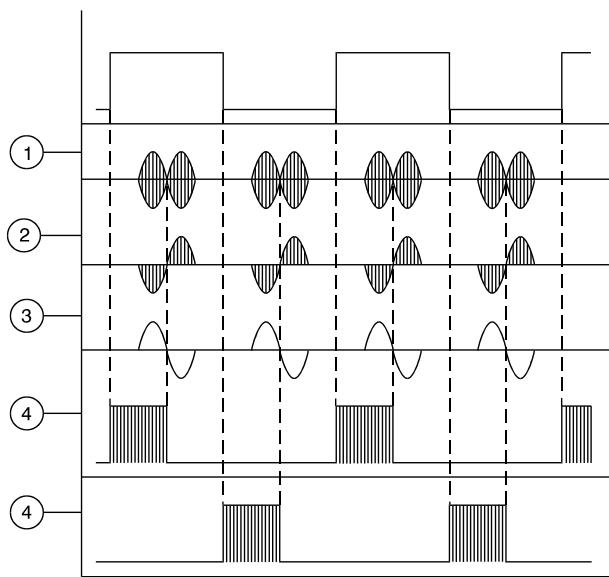


Fig. 4.29. Time diagram of coordinate couples digitalization

The pulse duration sets the scanning time of X coordinate, and the pause – of Y coordinate. At the beginning of each scanning a position counter (timer) is started which is stopped by the comparator. The state of the position counter is the value of the current coordinate. The signals from the time-diagram, marked with numbers (1,2,3,4) correspond to the outputs of the blocks in Fig. 4.28.

4.2.2. Formal Model of Digitizer Functions

Discretization of the Input Graphical Objects

In the general case graphical objects are described by means of continuous functions. Each continuous function $F(x, y)$ is presented as a set of geometric elements. They are called geometric primitives:

$$F(x, y) = \{f_1(x, y), f_2(x, y), \dots, f_n(x, y)\}, \quad (4.27)$$

where $f_1(x, y), f_2(x, y), \dots, f_n(x, y)$ are functions describing the primitives. When the graphic object is input, each of the component geometric elements is discretized separately. The set of discrete functions obtained models the discrete form of the whole graphical object. This manner of geometric element input as line segments, arcs and circumferences allows the elements to be described by several points only instead of being entirely discretized. On one hand, it makes it easier for the operator and on the other hand, it significantly reduces the amount of information transmitted and processed.

The problem of function discretization is related to its presentation in a rectangular discretizing grid. This grid is formed by parallel lines positioned at a constant pitch of dx for the X axis and dy for the Y axis. The pitch is equal to the resolution m of the device (Fig. 4.30):

$$dx = dy = m. \quad (4.28)$$

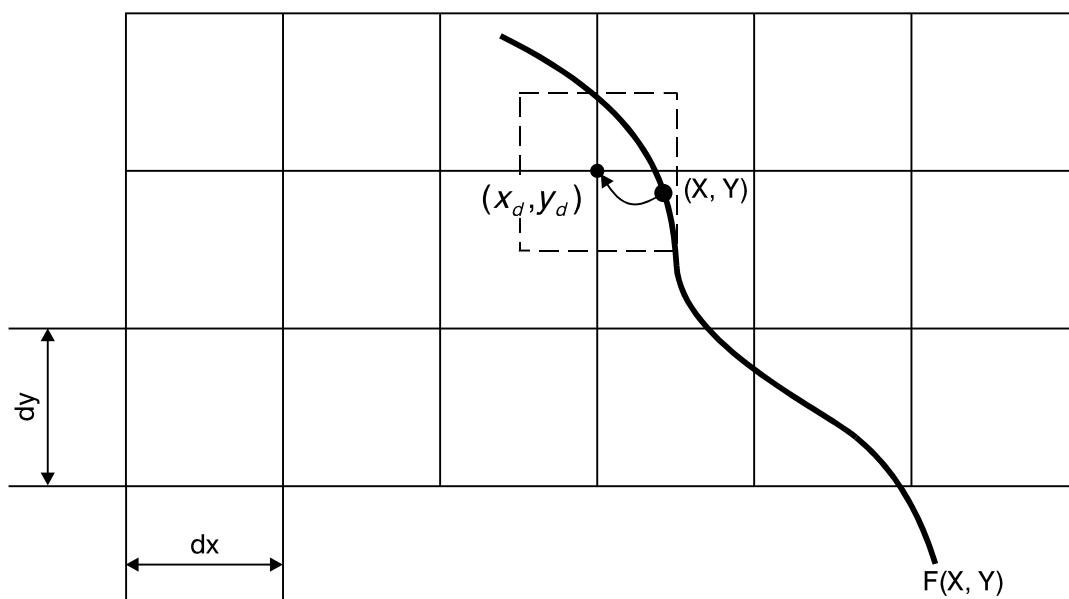


Fig. 4.30. Presentation of a function in a rectangular discretizing grid

Discretization is similar to ROUND function which converts real numbers into integers. Thus the graphic of a real function is replaced by a

broken line or more exactly by the points of this line which have integer coordinates / whole coordinates.

The conversion is precisely described mathematically as follows:

$$\begin{aligned}x_d &= \left\lceil \frac{x}{dx} + 0,5 \right\rceil \\y_d &= \left\lceil \frac{y}{dy} + 0,5 \right\rceil\end{aligned}, \quad (4.29)$$

where the function [a] returns the largest integer which does not exceed a.

The point with x_d and y_d coordinates is the nearest to (x, y) point and is defined by the formulae for x_d and y_d .

For the input of a certain graphic primitive $f_i(x, y)$, additional information about the type of primitive, about the thickness and type of the contour line, about colour, etc. is required. These parameters are called the attributes of the primitive. The set of coordinate information and attributes of a given geometric primitive is called input primitive. If it is denoted with IP, then:

$$IP = \{IA, IL\}, \quad (4.30)$$

where $IL=\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ is the set of points determining the primitive;

$IA=\{a_1, a_2, \dots, a_n\}$ is the set of its attributes.

The digitalized graphic object is presented as an ordered set F_i of input primitives:

$$F_i = \{IP_1, IP_2, \dots, IP_n\} \quad (4.31)$$

$$\text{or} \quad F_i = \{IA_1, IL_1, IA_2, IL_2, \dots, IA_n, IL_n\}. \quad (4.32)$$

It can be seen that the digitalizing process has two components: coordinates and attributes.

Coordinate Transformations

One of the major tasks of digitalization is to perform coordinate transformations. A number of its functions are based on them. Some of these functions are related to the necessity the coordinates of the particular points to be calculated in a coordinate system different from the one of the digitizer. The reason is positioning the graphic object in its own coordinate system, called user's system.

Translation. Each point $P(x, y)$ is translated into a new point $P'(x', y')$ (Fig. 4.31) with coordinates:

$$\begin{aligned}x' &= x + D_x \\y' &= y + D_y\end{aligned}, \quad (4.33)$$

where D_x is the displacement along X axis, and D_y is the displacement along Y axis.

The translation of point P into P' can be written in vector form. Let:

$$P = \begin{bmatrix} x & y \end{bmatrix} P' = \begin{bmatrix} x' & y' \end{bmatrix} T_T = \begin{bmatrix} D_x & D_y \end{bmatrix} \quad (4.34)$$

Then equalities (4.33) are presented in the form of:

$$\begin{bmatrix} x' & y' \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} + \begin{bmatrix} D_x & D_y \end{bmatrix} \quad (4.35)$$

or in a shorter form:

$$P' = P + T_T . \quad (4.36)$$

Each object is translated (moved) by using the equalities (4.33) for each of its points.

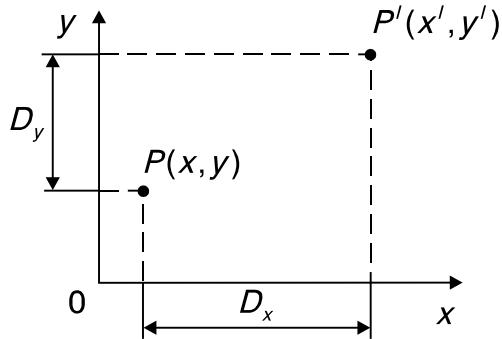


Fig 4.31. Translation

Scaling. Each point $P(x, y)$ is scaled (extended) into point $P'(x', y')$ by means of a scaling coefficient:

$$\begin{aligned} x' &= x \cdot S_x \\ y' &= y \cdot S_y , \end{aligned} \quad (4.37)$$

where S_x and S_y are scaling coefficients.

By means of the matrix:

$$S = \begin{bmatrix} S_x & 0 \\ 0 & S_y \end{bmatrix}, \quad (4.38)$$

equalities (4.37) are written in matrix form:

$$\begin{bmatrix} x' & y' \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} S_x & 0 \\ 0 & S_y \end{bmatrix} \quad (4.39)$$

or in a shorter form:

$$P' = P \cdot S . \quad (4.40)$$

Rotation. Point $P'(x', y')$ is obtained from point $P(x, y)$ through rotation to θ angle, as it is shown in Fig. 4.32.

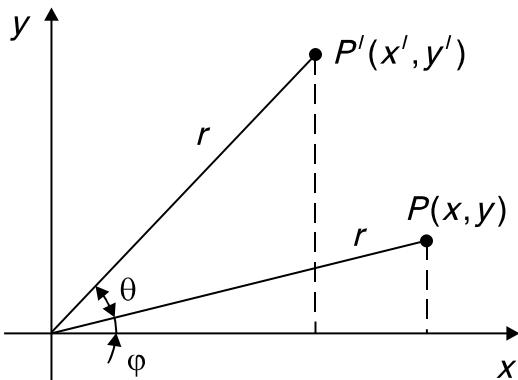


Fig. 4.32. Rotation

The coordinates of points P and P' are presented in the following way:

$$\begin{aligned}x &= r \cdot \cos \varphi \\y &= r \cdot \sin \varphi\end{aligned}\quad (4.41)$$

$$\begin{aligned}x' &= r \cdot \cos(\theta + \varphi) = r \cdot \cos \varphi \cdot \cos \theta - r \cdot \sin \varphi \cdot \sin \theta \\y' &= r \cdot \sin(\theta + \varphi) = r \cdot \cos \varphi \cdot \sin \theta + r \cdot \sin \varphi \cdot \cos \theta\end{aligned}\quad (4.42)$$

or also

$$\begin{aligned}x' &= x \cdot \cos \theta - y \cdot \sin \theta \\y' &= x \cdot \sin \theta + y \cdot \cos \theta\end{aligned}\quad (4.43)$$

The coordinate dependency of the two points can be represented in matrix form as well:

$$\begin{bmatrix} x' & y' \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \quad (4.44)$$

or shorter

$$P' = P \cdot R, \quad (4.45)$$

where R denotes the matrix of rotation.

Menu Function

This function is used for input of various commands, modes, input primitive attributes, and also for text information (Fig. 4.33). The menu configuration is a rectangular table consisting of a certain number of fields. It is positioned at an arbitrary place in the digitizer operating format so that it guarantees convenient service.

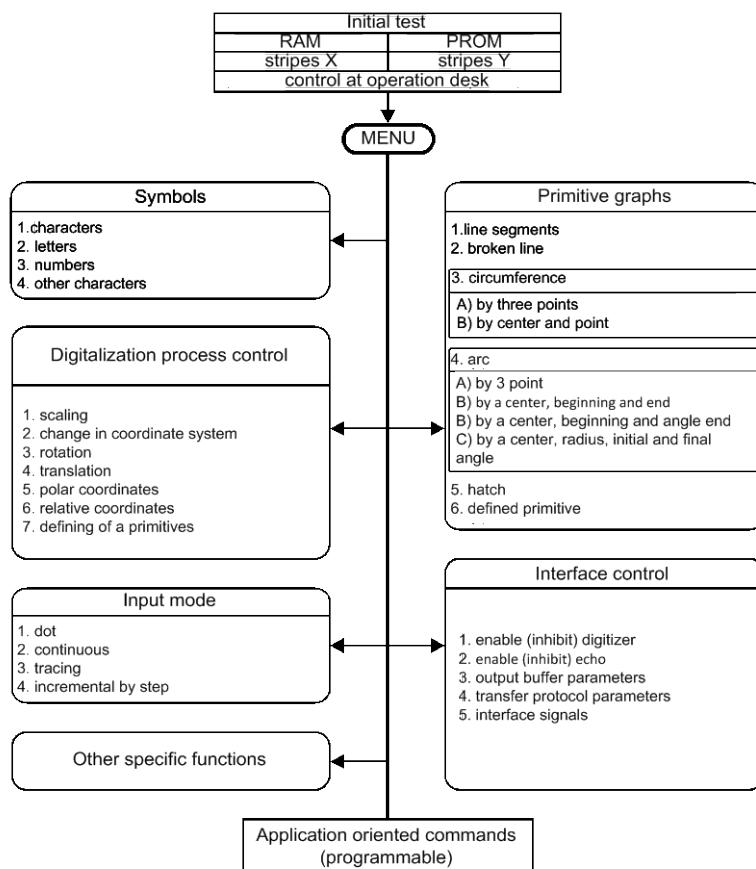


Fig. 4.33. Complete functional description of a digitizer.

An identifier corresponds to each field in the table. It is sent to the graphic programs instead of the coordinates of the digitalized point. Each field is a set of points. A program check is necessary to determine to which of these sets the digitalized point belongs. This procedure includes comparison of input point coordinates to the menu coordinates given in tables. Thus with one comparison a group of sets belonging to a line or column, is rejected. After choosing the respective menu field, the respective identifier is also retrieved which means a certain function, attribute or character chosen from the menu.

Graphic Information Input Modes

Various ways of graphic information input are available. Generally they fall into four modes:

1. Point (key) mode. The operator moves the sensor to the desired point in the work field and then inputs current coordinates by pressing a button whereat they are recorded in the output queue.
2. Continuous mode. The digitizer continuously generates the current coordinates of the sensor location to the output queue. Coordinate couples are output through a certain period of time, necessary for performing the process of digitalization.
3. Continuous-key mode (tracing). By pressing the button the operator inputs a complicated line. Data are sent to the output only when there is a change in the sensor position.
4. Incremental mode. The digitizer sends coordinates to the output queue only when there is a precisely identified sensor displacement on the work field (current coordinate growth). This mode is equivalent to the tracing mode with an increased pitch of the discretizing grid.

4.3. Scanners and Digital Cameras

Video Signal Matrix Photo-convertisers

Multi-element video signal converters (photodetectors) are the most important part of scanners and digital video cameras. In these converters, similarly to cathode ray tubes, a video signal is produced at the device output which can be converted into digital form and processed by a computer. [4]

A general view of the structure of a semiconductor matrix photo-converting device is shown in Fig. 4.34. It incorporates a number of basic blocks:

- a matrix of photosensitive elements;
- unit for outputting the charge generated by photo-sensitive elements and its possible storage;
- unit for converting charge packets into electric video signals;
- electronic unit for generating clock cycle and synchronizing pulses;
- unit for preventing blurring of the image as a result of charge overload.

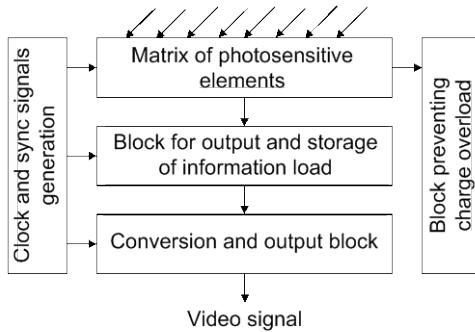


Fig. 4.34. Block-diagram of a matrix photo-converter

Semiconductor matrix video-signal converters are classified by different symptoms:

1. Type of photosensitive elements (MOS capacitors or PN junctions).
2. Spectral range (photodetectors for the visible or for the infrared region).
3. Character of photosensitive matrix (linear or two-dimensional).

The photosensitive elements are Charge Coupled Devices (CCD). Under the effect of external voltage, depleted areas are formed in them in which photon-generated current carriers collect. The charge accumulated for a period of time depends on the luminescence of the element. The unit for output and storage of information load retrieves the load sequentially accumulated in the matrix elements and sends it to the converting and output unit where the video signal is formed.

Scanners

Scanners are automatic DGII. They convert graphic information into digital without being necessary for the user to input separately each graphic object (dot, line, circumference, etc.). Thus digital raster images are produced from photographs, slides, texts, illustrations, etc. The operating principle of the scanner is shown in Fig. 4.35. The original image is illuminated by a light source built in the device. The reflected beams are caught by a photo-converter with linear structure which forms an analogue electric signal. This signal is converted into a digital code by an analogue-to digital converter and it is sent to be processed by a computer by means of suitable interface. The linear structure of photo-converter requires the image to be scanned line-by-line.

Depending on the scanning method, there are several types of scanners.

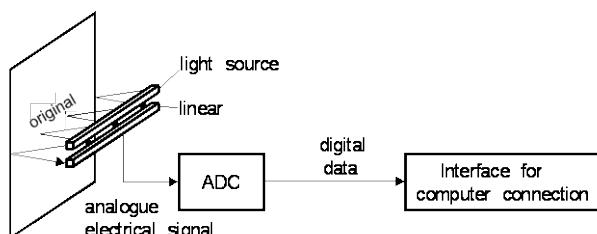


Figure 4.35. Operating principle of a scanner

Hand-held Scanners

The user has to move the scanner along the original in order to input the desired image (Fig. 4.36). Hand-held scanners usually input images with dimensions of up to 4 inches (10 cm) and their resolution is from 200 to 400 dpi (dots per inch). Therefore they are good for small-size images. Since quality depends on human hand stability, these scanners are used for not so exacting processing. Their major problem is their narrow scanning range which is solved by suitable software. Most software products provide the 'gluing' function – the object to be input is divided into strips which are input separately and then they are glued together automatically (~*~). It is impossible to move the scanner at constant speed in manual scanning. Therefore, most hand-held scanners have a speed-measuring sensor. By means of this sensor, the image being scanned is synchronized to the real scanning speed. In spite of this, if the movement is accelerated, there will be some extended areas. Some models have a light-emitting diode indicator which gives signals when the admissible scanning speed is exceeded. Also, when there are deviations from the ideal straight line of scanning, folds appear in the image. Therefore, most models have guides (built-in additional rollers) which ensure straight-line movement. Scanning results are stored in files (TIFF, PCX, IMG, MSP, CUT,etc.) or they are printed on a printer.

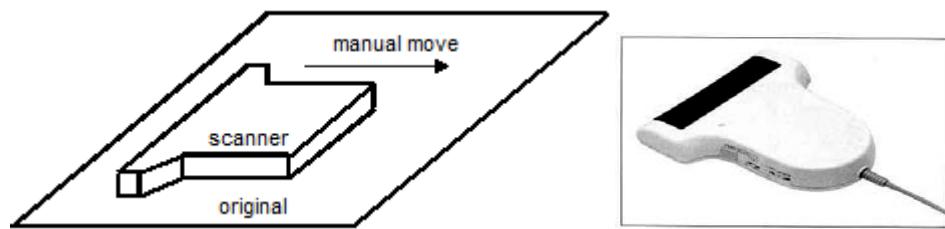


Figure 4.36. Hand-held scanner

Flatbed Scanners

In flatbed scanners the sensor matrix is moved by an automatic mechanism (Fig. 4.37). They look similar to photocopier machines. The image on paper is placed on a flat glass surface. The motion pitch of photo-converter is exactly regulated line by line and thus the quality of the produced image is improved. The resolution is from 300 to 1600 dpi (with interpolation), and the work field size is from A4 to A3. The price of these devices is higher owing to the precise scanning mechanism used.

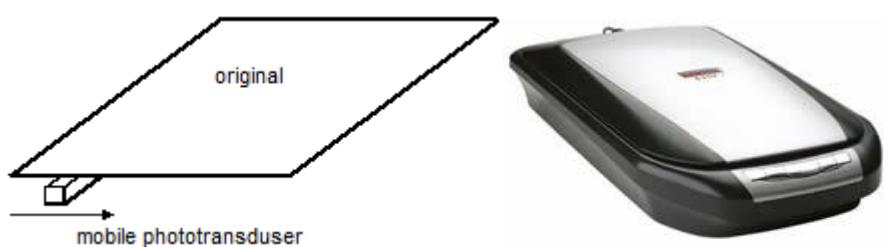


Figure 4.37. Flatbed scanner

Drum Scanners

In drum scanners the original is fixed on a transparent drum (Fig. 4.38). An optical sensor moves linearly simultaneously with the rotating drum and thus information about colour and optical density of each point of the original is obtained. The resolution of this type of scanners is up to 4200 dpi. They are distinguished for superior quality but high price as well. They are used in professional publishing. [33]

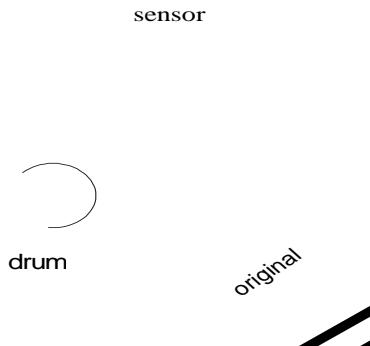


Figure 4.38. Drum scanner

Broad-format Scanners

Broad-format scanners are designed for large drawings in mechanical engineering, architecture design, cartography, etc. (Fig. 4.39). They reach up to A0 format. The sensor used is linear and stationary, operating on the short side of the format. The document is driven and passes through the scanning band and its length can be arbitrary. [33]



Figure 4.39. Broad-format scanner

Slide Scanners

These scanners achieve maximum resolution - 4200 dpi. Their work field is usually with the size of the slides 35x35 mm or strips 35mm wide. The operating principle is entirely different from most types of scanners usually processing reflected beams. These devices process light passing through the original. Drum and flatbed scanners have been offered lately with attachments for transparencies.

Specialized Scanners

Bar-code scanners. They are used mainly in retail applications for scanning bar codes (black-and-white bar image). The operating principle of such a device is shown in Fig. 4.40.

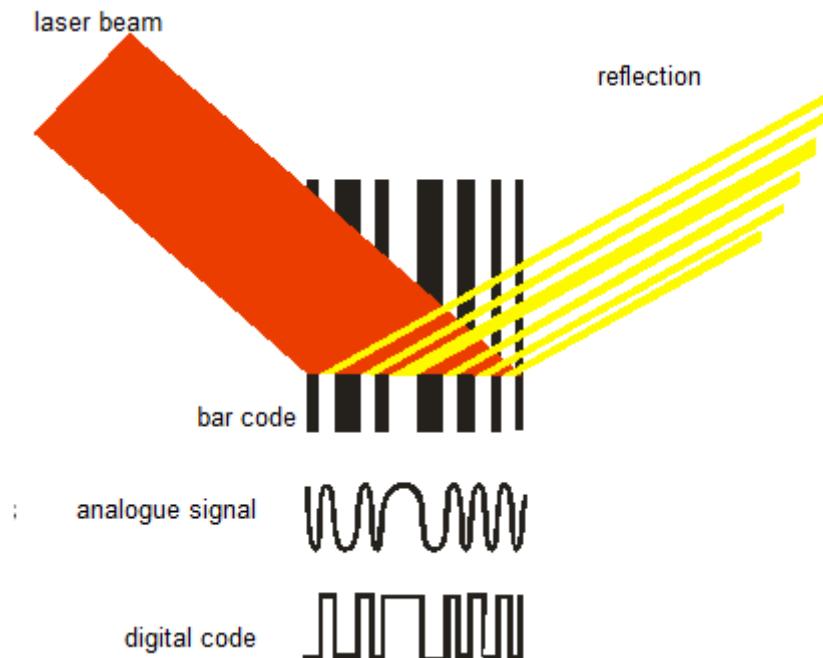


Figure 4.40. Operating principle of a bar-code scanner

Other types of specialized scanners exist designed for:

- taking fingerprints in access systems;
- passport check-ups;
- cheque handling;
- input of documents of different shape and purpose.

Depending on whether scanners can recognize colour images, they are divided into black-and-white and colour. Black-and-white scanners are characterized by the number of degrees of gray that they can perceive. Depending on type of document to be scanned, different number of gray tones is used. For example, in drawings only two degrees are sufficient: black and white, while in photographs a greater number of shades are necessary. Therefore most scanners allow this parameter to be set for each scanning. The choice is between two shades of gray up to the maximum number for each individual scanner. The basic parameter for colour scanners is the number of colours used. It can also be set depending on the type of image. In modern scanners the maximum number of colours reaches several million, and for each dot 24 bits are used. Colour scanners allow operation in black-and-white mode when texts or drawings are scanned, to reduce the information volume and to enhance the scanning speed.

The scanner resolution is determined by the smallest pitch for moving the scanning mechanism. It can be improved by means of software. Interpolating algorithms are implemented to fill in some more pixels between the scanned ones. This approach does not increase the image details but makes the image smoother.

Scanning results are stored in files (TIFF, PCX, IMG, MSP, CUT, JPEG, etc.) or they are printed on a printer. Scanners are connected to computer environment by means of accessible interfaces (most frequently a USB port for new models) or by means of a PCI controller.

Digital Cameras

Video-signal converters with two-dimensional photosensitive matrix (CCD matrix) are employed in digital cameras. The matrix is a complex semiconductor element consisting of a number of photosensitive elements composed of photo-diodes. These elements are grouped and located in a certain way with possible addressable access to each active element in it. The number of sensitive elements composing the CCD matrix determines the resolution of the image.

Since the pixel is the smallest element in an image, this term is suitable for describing the capacities of the matrix and for some calculations as well. If the camera takes a picture with a resolution of 1280x960 pixels, this means that the total number of pixels from the matrix is equal to 1 228 800. In this case, the camera has a megapixel matrix.

Most digital cameras of low and middle class contain one CCD matrix (image sensor). Modern digital cameras of higher class use three separate sensors for RGB-components of the input signal. Each of these sensors receives its colour – red, green or blue and the extra colour components are separated by filters transmitting only signals with specified wavelength.

Different manufacturers implement different algorithms for producing an image. Matrices operate in different modes. For example, two sensors can be determined for green colour, and the third to be combined for red and blue colours. After intercepting the image, the information is interpolated.

Tiny lenses are used in the matrix located over each pixel, serving to focus and direct photons exactly to the right pixel element (Fig. 4.41). Then photons are transformed into electrons by means of a silicon photodiode which is located immediately underneath the micro-lens in the light-sensitive region, and the region itself preserves its electric charge on the principle of capacitor. The matrix must be initialized (cleared) periodically at a frequency of several tens of times per second, in order to be able to take the next photo.

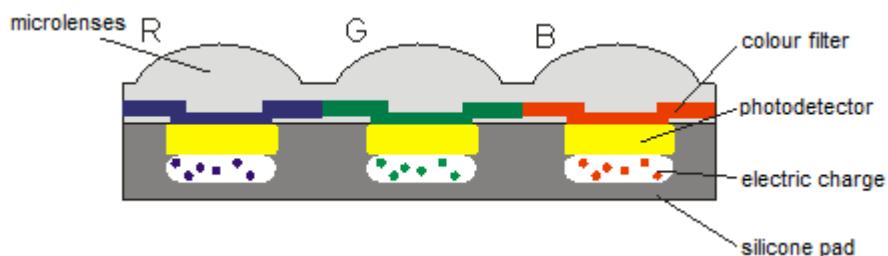


Fig. 4.41. Light-sensitive matrix

Matrix manufacturers use also various constructions of filters which transmit light with definite wavelength (i.e., definite colour). The most frequently implemented technology is Bayer Pattern, in which green filters are twice as much as red and blue ones. Besides that, the red, green and blue filters are positioned in chess-board order which makes colour defects less likely to appear and the mutual effect of adjacent pixels of the same colour to

be avoided. Manufacturers of digital cameras build micro-programs in their products which execute various algorithms for improving and additional processing of images.

Cameras that have been developed over recent years incorporate a colour matrix. Complex colour-coding filters of mosaic type are used in them without exception (Fig. 4.42). The only photo-converting matrix has to have high resolution and uniform spectral characteristics.

According to the type of sensor, matrix semiconductor photo-transducers fall into three groups:

- CCD matrices;
- Super CCD matrices;
- Super CCD SR/HR matrices.

W	Cy	W	Cy
W	Cy	W	Cy
Ye	W	Ye	W
Ye	W	Ye	W

W	Cy	G	Ye
G	Ye	W	Cy
W	Cy	G	Ye
G	Ye	W	Cy

W	Ye	W	Ye
W	Ye	W	Ye
Cy	W	Cy	W
Cy	W	Cy	W

Ye	Cy	Ye	Cy
Mg	G	Mg	G
Ye	Cy	Ye	Cy
G	Mg	G	Mg

W - white

, G – green, Ye – yellow,

Mg – purple, Cy - grey

Figure 4.42. Model structures of mosaic colour-coding filters

CCD Matrices

In CCD matrices MOS capacitors or photodiodes are used as photosensitive elements. MOS capacitors are less applicable than photodiodes. Light-sensitive elements have square or rectangular shape. They are positioned in the way shown in Fig. 4.43. The video signal received at the output is a sequence of pulses whose amplitudes correspond to the light intensity at the respective point of the image for a certain moment of time. The scanning process is automatic and is performed by the image receiver itself and therefore these devices are called self-scanning devices. [41]

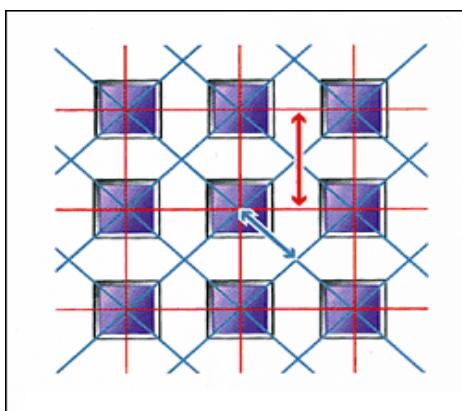


Fig. 4.43. CCD matrix

Super CCD Matrices

In the new Super CCD matrices, the elements are located at the angle of 45° and have the shape of an octagon (Fig. 4.44). This allows the area taken by the current-conducting strips to be reduced. The new principle of location of photosensitive elements looks like honeycombs. Thus the effective resolution is enhanced (by around 60%) compared to conventional matrices. This new arrangement of elements increases the size of each pixel and as a result the sensitivity is enhanced and the noise is reduced. [41]

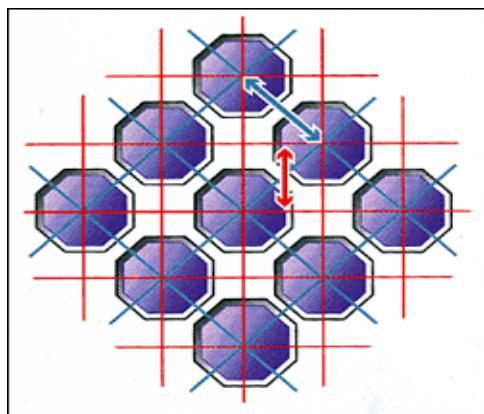


Fig. 4.44. Super CCD matrix

Super CCD SR / HR Matrices

These matrices have a very wide dynamic range – fourth generation. The structure of Super CCD SR matrix consists in combining larger photodiodes with high sensitivity (S-pixels - Fig. 4.45), with smaller photodiodes with low sensitivity (R-pixels - Fig. 4.46).

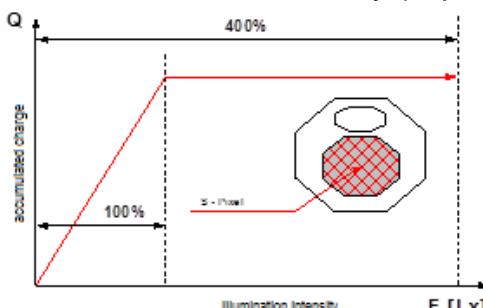


Figure 4.45. S – pixels

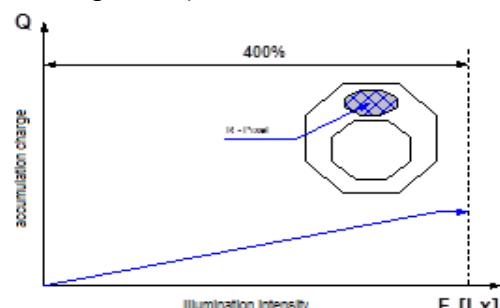


Figure 4.46. R – pixels

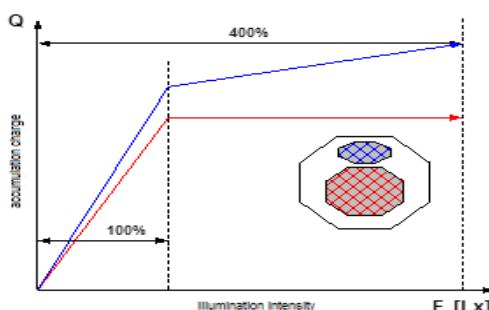


Figure 4.47. Resultant photo-converting characteristics of a matrix

Photographs taken by SR-matrix are of very high quality without any loss of details in the darkest and lightest areas of the image. The resultant photo-converting characteristic of the matrix appears to be a sum of the photo-converting characteristics of the two matrices (Fig. 4.47). The structure of these sensors also looks like honeycombs.

Video-signal converters with two-dimensional photosensitive matrices are used in digital cameras. A system comprising three matrices is considered to be classical. [4] The three primary colours are sent to these matrices by means of an optical colour-separating device. A very good colour image is produced with the help of the three-matrix devices. However, they have higher price, larger size and larger mass.

A model block-diagram of a digital video camera is shown in Fig. 4.48. The signals for the three primary colours are converted into digital by analogue-to-digital converters. The system has microprocessor control which reads the data from the ADC and sends them along the interface connection to the computer.

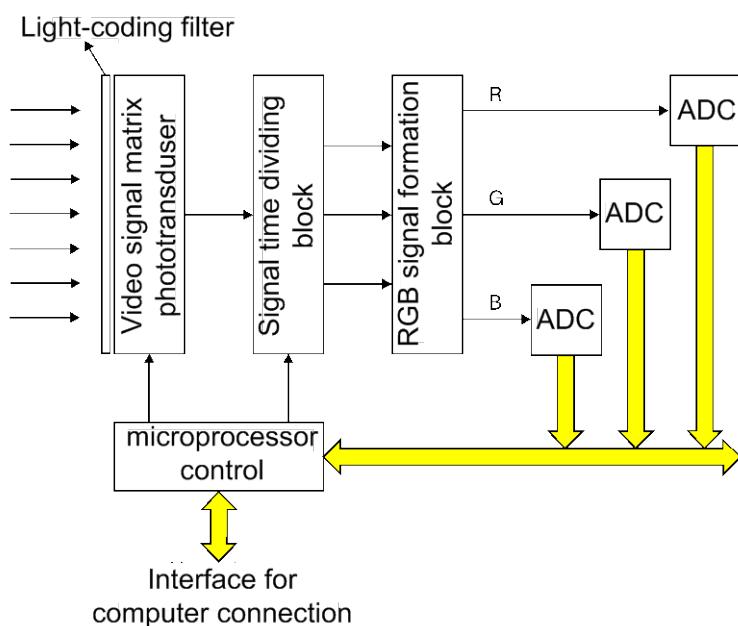


Fig. 4.48. Model block-diagram of a digital video camera

Nowadays specialized controllers are more widely used in video processing. They allow a television signal from PAL, SECAM or NTSC systems to be input and they can be directly connected to most video cameras and video recorders. The construction of such a controller is shown in Fig. 4.49. The analogue television signal is converted into digital form by ADC and then it is decoded with the help of a suitable algorithm. To reduce the data volume, information is compressed, before it is transferred to the system RAM or to the hard disk. When information is retrieved, it is decompressed first and then digitally colour-coded. Data processed in this way are converted into an analogue television signal by DAC.

Audio-input and audio-output are incorporated in the controller for input and output of the sound accompanying the picture.

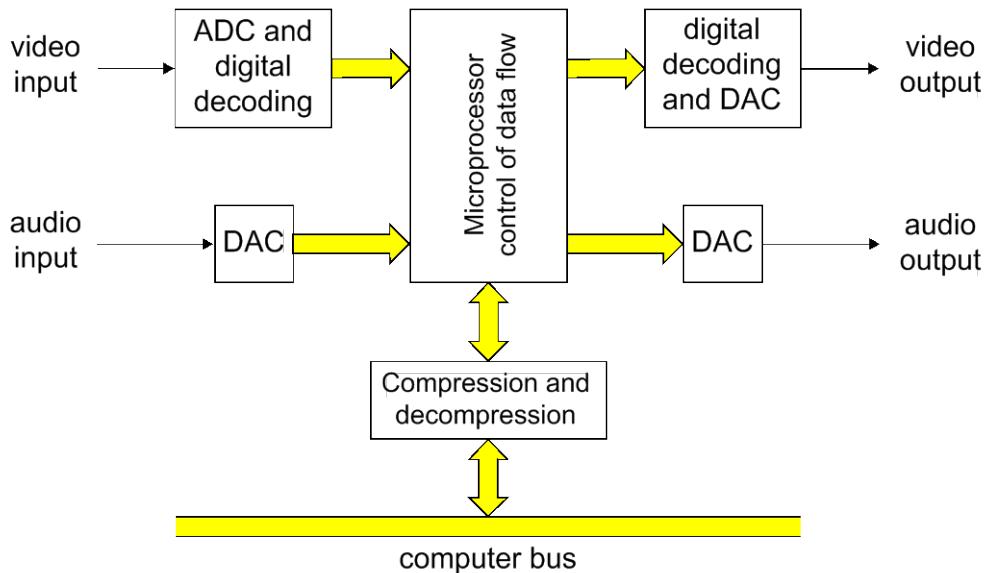


Fig. 4.49. Specialized controller for input and output of video information

Digital Cameras

Conventional photography had its origin in 1830 and the inventors in this field were Louis Daguerre and Henry Fox Talbot. The ordinary camera uses a polymer film with light-sensitive coating (emulsion) applied on it, to form the obtained image (negative or positive). The frame is received as reflected light by the lens and passing in the form of photons through the lens system, affects the emulsion applied on the film and causes a chemical reaction. To fix the results of this reaction, a chemical process, known as development, is run. Digital cameras are much more complicated than conventional and comprise similar elements and systems. They also have a lens which focuses the stream of reflected light which falls on a special light-sensitive element - *CCD matrix* (image sensor). The signals generated by the matrix are converted by ADC. A specialized microprocessor controller analyzes the incoming information and presents these signals in digital format. The software processes the data further up to final results which are recorded on the respective carrier and are shown on the display (LCD panel). The quality of the picture, the format in which it will be recorded (*.jpg, *.tiff or another one); the degree of compression, various options as: frame for each photograph, time and date indication and a lot more functions and capabilities, depend on the preset algorithms.

4.4. Bar Codes

Introduction

A typical bar-code is shown in Fig. 4. 50. Data are presented by a series of bars of varying widths and spacings between them. The coded numbers are often printed below for convenience, as it is shown in the figure.



Figure 4.50. Bar-code

Codes are read easily by means of a light pen or scanner. The light pen has a photosensitive sensor and when it is passed over the bar-code, the dark and light bars are converted into a sequence of logical levels. These levels are decoded to obtain the data. When a scanner is used the bar-code is moved transversely to the window of the device. This window is scanned quickly from the bottom upwards by a narrow light beam and the light reflected by the bar-code is caught by a photosensitive sensor, which converts it into a sequence of pulses. In both variants – by a light pen or scanner, it is indicated when an invalid code has been read. The operator can try to read the code again. [18]

Specifics of Coding

Each digit in Fig. 4.50 is presented by a sequence of alternating two black and white bars. The total width of these black and white bars is the same for all symbols. Therefore, the speed of moving the bar-code during scanning is not critical - the scanning time for each bar is expressed as part of the time for scanning the whole digit. These parts are relatively independent of the speed of displacement, because it varies insignificantly during scanning.

Each digit is divided into seven elements. They are black and white as it is shown in Fig. 4.51.

The digits are coded in different ways depending on whether they are in the lefthand or righthand portion of the bar-code. All digits in the lefthand portion have an odd number of black elements and always start with a white bar, while those in the righthand portion have an even number of black elements and always begin with a black bar. This enables the scanner to determine the direction of scanning automatically.

The two portions of the bar-code are divided by two black dividing bars. Such pairs of bars also mark the beginning and end of the bar-code.

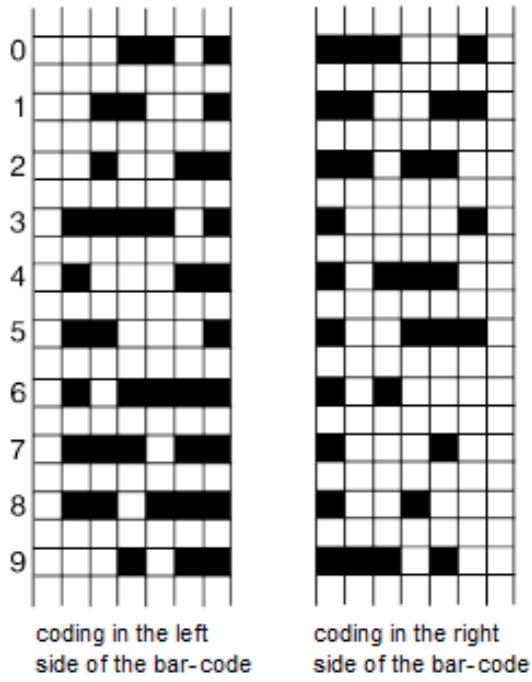


Fig. 4.51. Bar-coding

Formats

One of the systems implemented for giving numbers to products is Universal Product Code (UPC). It is shown in Fig. 4.52. The numbers of manufacturer and product are placed in the lefthand and righthand portions of the bar-code, respectively. The system character identifies the code version and the type of article. The check digit is used for error detection and correction.

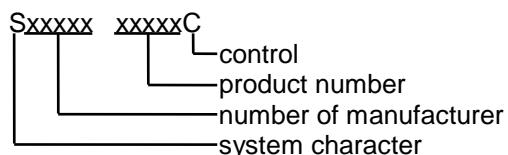


Fig. 4.52. UPC format.

The European Article Number (EAN) system includes a two-digit code which identifies the country. This system is compatible with UPC.

Various bar-code formats exist, even for internal use in some organizations.

Errors

Errors in reading bar-codes are rare in practice. Invalid scanning happens, for instance, if the light pen moves at acute angle and for this reason does not scan the whole code. This is easy to detect because the edge markers for start or end are missing and probably part of the data.

Spots on the code may cause white elements to be interpreted as black and vice versa. In most cases this results in an unrecognizable code. If an error is found only in one digit, it can be corrected by means of the check digit. Error scanning indication is activated if more than one invalid digits are

detected. It is possible the spot to cause forming a recognizable code of another digit. This error is detected by the check digit. This digit is a check sum of all digits in the code. Thus errors are detected in more than one digit. Multiple errors in which a correct check sum value is obtained are extremely rare.

4.5. Optical Carriers (CD)

The compact disk (CD) is an optical carrier of digital information in the form of disk.

The optical technology standards are divided into two major types:

- CD (CD-ROM, CD-R, CD-RW);
- DVD (DVD-ROM, DVD-RAM, DVD-RW, DVD-R, DVD+RW, DVD+R).

According to the type of information stored in optical carriers, they are divided into several types:

CD-DA (Compact Disk Digital Audio). These are the first optical carriers designed for storing Audio information.

CD-ROM (Read Only Memory). These are disks only for reading. The information on these carriers is written by the manufacturer and the user cannot make a recording.

CD-R (Read). These disks allow a single user's recording.

CD-RW (ReWritable). These are rewritable disks and allow multiple user's recordings – around 1000 times.

DVD (Digital Video Disk). This standard allows storing of video information only. Later the Digital Versatile Disk standard was developed. It is faster and besides video information it can store data as well.

DVD-R (Read). They allow a single user's recording.

DVD-RW (Read/Write). They allow multiple user's recordings – around 1000 times.

Operating Principle of an Optical Disk Device

An optical disk device (Fig. 4.53) reads data from a disk. Information is stored along a single spiral path which usually has a length of 40000 revolutions. The tracks are divided into sections and each section has a unique address. The reflective capability of the disk active surface varies in accordance with the recorded data. A laser beam of low power is focused by means of a lens on to a tiny dot (usually it has a diameter of $1\mu\text{m}$). It is reflected when there is an active level or it is absorbed by the surface altered as a result of the recording. The reflected beam is caught by a detector and it is converted into an electric signal. A specified section is accessed in two steps. In the first step the optical system moves closely to the track containing the desired section. This is not accurate because the distance between the tracks is very small. After reading the section addresses, it is determined how many tracks aside from the desired one the initial positioning is. In the second step precise positioning on the track containing the specified section is done by means of an adjustable reflecting surface placed on the laser beam path and thus the needed data can be selected fast. The same surface is included

in a feedback system which directs the beam accurately on to the track. The causes for the beam shift are spiral location of data, disk eccentricity, etc. [18]

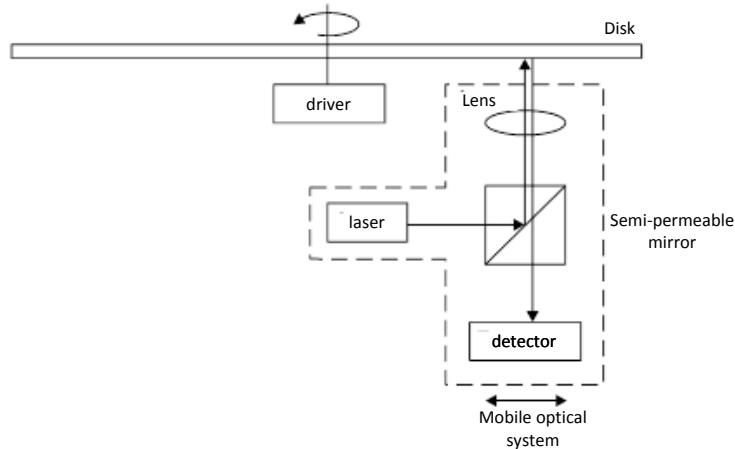


Figure 4.53. Optical disk device

Owing to the two-step positioning the access time to an adjacent track is too long – 1 ms, and the access to a remote track takes less than 150 ms.

Disks are manufactured with different sizes - 8cm (3,15 inches), 12 cm (4,72 inches), 13,3 cm (5,25 inches), 20 cm (8 inches) and 30 cm (12 inches). The information capacity of optical disks varies from 500 MB up to 4 GB. "Jukebox" systems have larger capacity. They automatically load disks from a stack containing around 100 disks in one or two disk devices.

It is characteristic of optical disks that they are not so demanding of the operating conditions compared to magnetic disks. The optical system is located far from the disk surface and they are not in danger of being hit. Besides, optical disks are coated with a transparent protective film and thus they are not affected by slight scratching and dust particles.

The compact disk is made of 1,2 mm thick polycarbonate coated with a very thin layer of aluminum (gold was used at the beginning) with a protective coating of lacquer . The label is usually printed on it (Fig. 4.54).

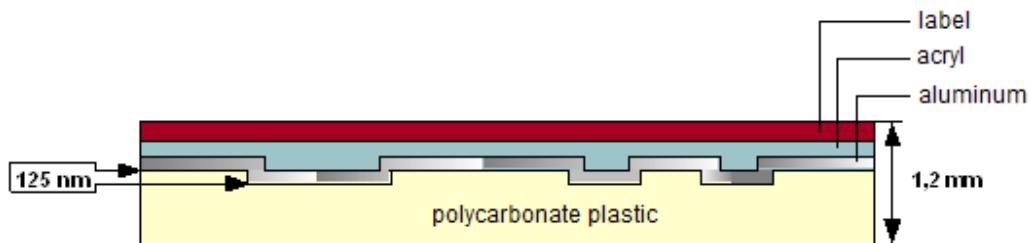


Figure 4.54. Construction of CD

A compact disk usually has a diameter of 12 cm and during the early years of manufacture it could store up to 650 MB of information (74 minutes of sound). Later, disks with capacity of 700 MB (80 minutes of sound) became more common. Carriers with capacity of 800 MB (90 minutes) are also available. However, they cannot be read by some devices for optical carriers. There are also mini CD with a diameter of 8cm which can hold about 120MB of data or 21 minutes of sound as well as CD whose shape resembles a credit

card (the so called disk-card). The format for data storing on the disk is known as „Red-Book“. It was developed by Philips. In accordance with this format sound with two channels with 16-bit pulse-code modulation (PCM) and discretization frequency of 44,1 kHz. is recorded on the CD. Owing to error correction by means of Reed-Solomon code slight scratches do not affect reading the disk.

Recording on Compact Disks

Data are recorded as non-homogeneous sections on an optical surface. These sections absorb the laser beam. They are 0.5 microns wide, at least 0.83 microns long and 125 nm high (Fig. 4.55).

Recording on compact disks in factories is performed by stamping using a nickel die. The tracks and pits (non-homogeneous sections) are engraved on the polycarbonate disk substrate. The nickel matrix is manufactured by galvanic copying of glassmaster. The glassmaster is made by exposure to a special laser recording device on a light-sensitive (photo-resistive) film applied on a glass substrate. Recording on CD-R and CD-RW at home is done employing optical recording devices. The recording process is called ‘burning’ the disk. [42]

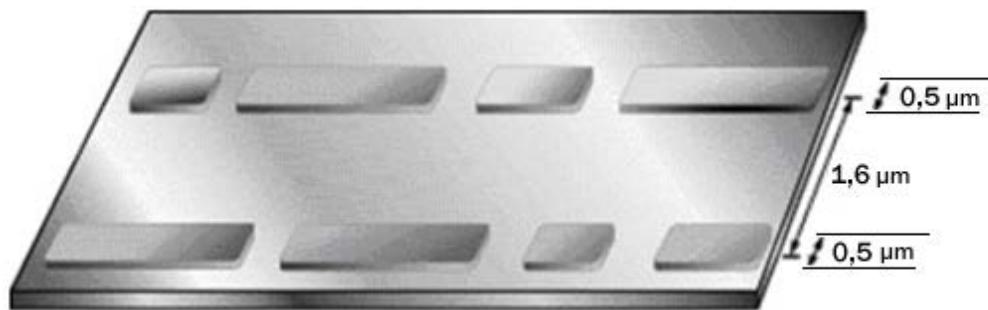


Figure 4.55. Data recording on a compact disk

By means of the standard DVD (Digital Versatile Disk) data can be written on a disk from 4,7 GB (DVD5) to 9,4GB (DVD9), respectively, for single or dual layer recordings. 18,8 GB (DVD18) disks are not common. The difference between a DVD and CD is in the recording depth, track size as well as in the wave length of the recording beam. Besides, a DVD is made up of two substrates 0,6 mm thick. They are glued together and that makes 1,2 mm, which is the thickness of a CD. Each of the DVD substrates carries one or two information layers and hence comes the versatility in the physical structure and capacity of disks (DVD5, DVD9, DVD14, DVD18). DVD standard is used mainly for distribution of films and multimedia. Unlike ordinary CD, DVD carriers are protected by additional technologies against copying and illegal distribution. DVD Region Code, CSS (Content Scrambling System), etc. are used to this purpose. [43]

DVD format employs four standards depending on the volume of the recorded information (Table. 4.1).

standard	DVD 5	DVD 9	DVD 10	DVD 18
sides	one	one	two	two
layers	one	two	one	two
capacity	4.7 GB	8.5 GB	9.4 GB	17 GB
4 hours	5 hours	6 hours	7 hours	8 hours

Table 4.1. DVD standards.

Until recently standards different from DVD5 were used only for DVD video and DVD - ROM. The boom in technologies was caused by the appearance of dual layer carriers. (DVD9). The capacity was increased from 4,5 GB to 8,5 GB. In this format, as far as hardware of recording devices is concerned, the change is very small – the optical heads are the same. The difference is only in firmware (firm software of devices) and in recording device controllers. They can operate with two layers and have an additional option for error correction. Emblems of recording devices operating with these technologies are shown in Fig. 4.56.



Figure 4.56. Emblems of recording DVD devices.

Construction of a dual layer DVD is shown in Fig. 4.57. The successive layers in them differ little in terms of transparency. [44]

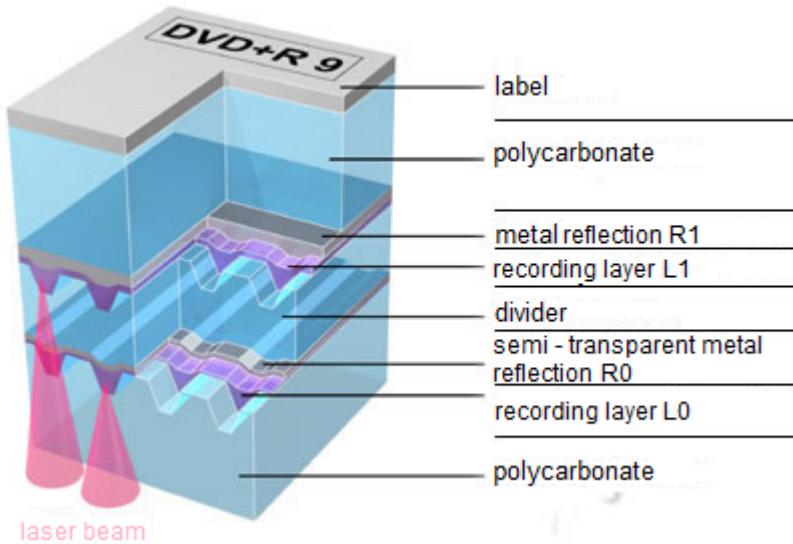


Figure 4.57. Successive layers in DVD

The major problem is to find a medium R0 reflecting layer which is transparent enough to allow the laser beam to reach the second layer L1, but at the same time to be reflective enough to allow information contained in the first layer L0 to be read. To be compatible with standard dual layer DVD-ROM

(or DVD Video) devices, the respective layers R0 and R1 must have reflective capacity of at least 18%. At the same time, to be compatible with DVD standard, R0 layer must transmit at least 50% of the laser beam (for reading and recording). Therefore the R0 layer is made of silver alloy and the format of recorded holes has been optimized. The dividing layer is only 55 µm thick and serves for physical division of the two layers. When the disk is being read or recorded, the laser beam is focused on one of the two layers by adjusting the lens focus. The recording starts from the DVD centre on the first layer and when the outside end of the disk is reached the lens sight alters the focus so that the laser can have an access to the second layer. The recording is resumed but this time from the outside to the centre. By implementing this technology theoretically it is possible to store up to 8,5 GB in a single disk. In practice, however, exactly 7,96 GB are stored. In video mode with DVD quality four hours of video information can be recorded and with VSH quality – up to 16 hours of video information. On the other hand, when dual layer mode of recording is applied, the speed falls to 24x. or 16x. Nowadays manufacturers prefer the dual layer technology and make unceasing efforts to achieve higher and higher speeds of reading and recording. At present, +R disks have maximum speed of 16x. At this speed the disk rotates at about 10 000 revolutions per minute and the recording device reaches its physical limits.

PLOTTERS

Plotters are graphic output devices in computer graphic systems. They differ from displays mainly in creating a durable visible image on portable physical carrier – paper, plastic foil, photoplate, etc. Plotters are automatic devices – they convert digitally coded information into an image without direct participation of an operator. [1]

5.1. Classification of Plotters

Depending on the manner of building the image, plotters are divided into vector and raster plotters (Fig. 5.1). Classical plotters implement vector principle of building the image as it is set in description of objects and their geometrical processing in generating computer graphic systems. In fact plotters perform movements following preset parameters for each vector and the writing tool leaves a track equivalent to this movement on the respective carrier (paper).

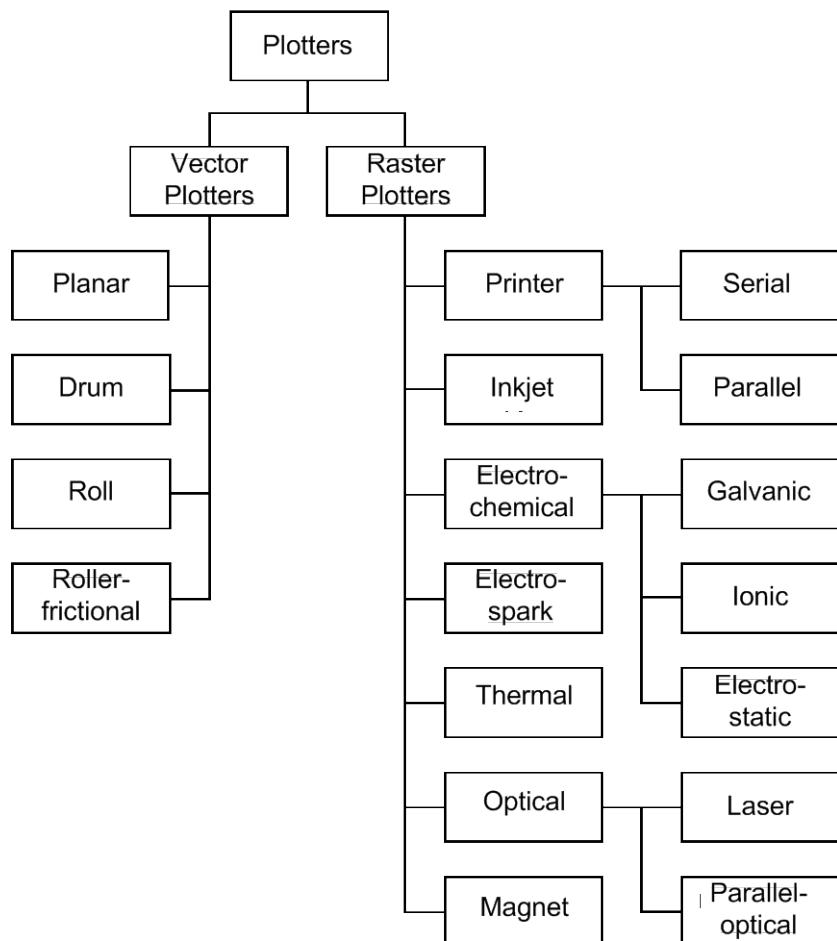


Figure 5.1. Classification of plotters

Raster plotters develop on the analogy of raster displays and building the image on television screen. They borrow the principle of building the image by scanning the screen area along a strictly repeating motion trajectory. The image is drawn on graphic information carrier by points. These point elements are located at strictly specified sites and the axial lines connecting them along the two coordinates form the so called raster.

In raster plotters various principles of plotting the image are implemented: mechanical, electro-spark, electro-chemical, thermal, optical, magnet and ink-jet (Fig. 5.1). Each principle has found a certain field of application but the most common are electro-chemical and optical plotters. The use of laser in raster plotters is a premise for very fast speed.

5.2. Vector Plotters

5.2.1. Mechanical Constructions of Vector Plotters

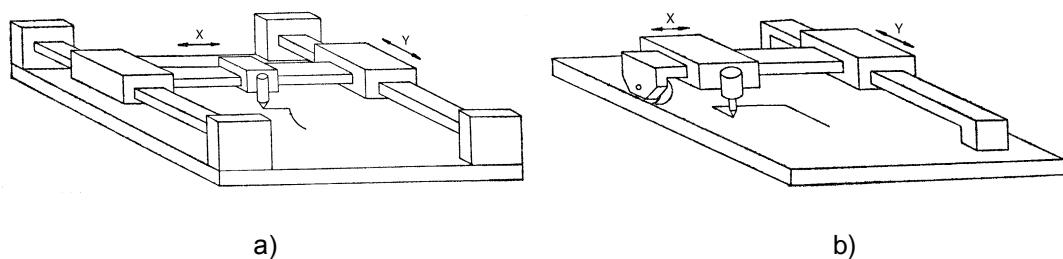
Each vector plotter comprises a mechanical construction and control block. The mechanical construction consists of a static supporting construction and mechanical executive units. The mechanical executive units, electrical drive and its control form a positioning system. Mechanical executive units are often called positioners.

Depending on the type of the static supporting construction and the positioning system construction, plotters are divided into several types: planar, drum, roll and roller-frictional.

In planar plotters (Fig. 5.2 a) and b)) the image carrier (most frequently paper) is static and lies on a plot. The positioner moves the writing tool over the plot which draws the image on the surface.

In drum plotters (Fig. 5.2 c)) the carrier is wound and fixed in a suitable way on the cylindrical surface of the drum, whose rotary motion is controlled so that one of the two coordinates to be traced tangentially. The writing tool moves along the other coordinate parallel to the drum axis. This scheme has an advantage over the planar one – the construction performing the motion along one of the coordinates is not carried by the part performing the motion along the other coordinate and thus the requirements towards the drive are reduced.

In roll plotters the idea of two independent constructive solutions to motion along either coordinate has been implemented, but the carrier is wound on rolls in order to be always tight. They have important advantages – it is not necessary the paper to be replaced after drawing each image and long images can be drawn.



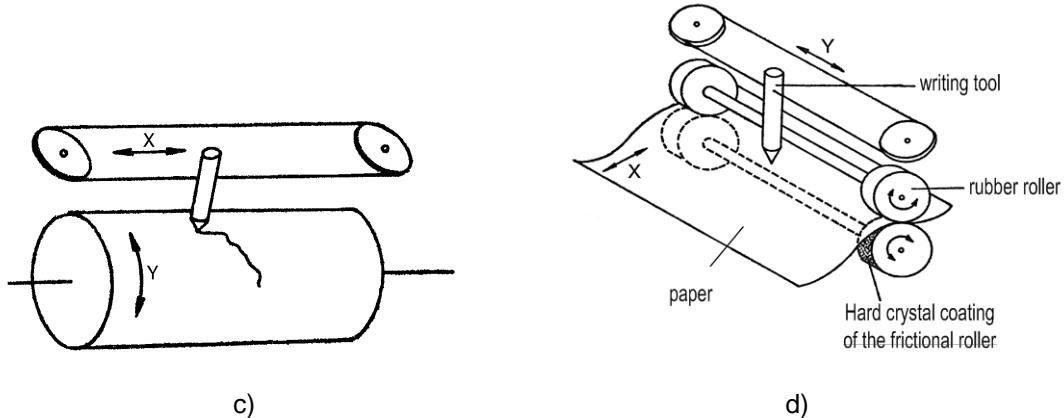


Figure 5.2. Principal constructions of drawing vector plotters:

- a) and b) planar;
- c) drum;
- d) roller-frictional.

The best constructive solution so far is the roller-frictional drive of paper (Fig. 5.2 d)). Here also the motions along both coordinates are constructionally independent as in drum constructions but the paper is not fixed on a drum. It is placed horizontally on guiding rollers and then it is pressed from above by parasite rubber rollers. By employing modern technologies, the frictional coefficient between the frictional surface and paper is almost equal to 1. To this end, on the especially treated surface of the guide roller, a layer of abrasive particles with precisely determined size and high mechanical strength is applied.

Besides the simplified mechanical construction, the roller-frictional drive of the paper has another advantage – low inertia moment of the driving rollers. This is a premise for enhancing the maximum acceleration and speed of motion of the mechanism. The positioning system dynamics is improved. The roller-frictional drive of the paper is the most prospective constructional solution to drawing vector plotters so far and it is preferred by most manufacturers of such devices in the world.

In practice, roller-frictional and planar plotters are most widely used. They allow high speed of drawing which is achieved due to low inertia moments in the constructive design. Unlike them, drum plotters are more awkward because of the large mass of the rotating drum on which the paper is fixed. Roll plotters are not successful either.

5.2.2. Control of Positioning Systems

The system for executive mechanism positioning is the basic part of plotter control. Requirements set to that sort of systems are: high speed of motion, high acceleration and high precision of positioning. Achieving good parameters to a great extent depends on the dynamic characteristics of the positioning system along X and Y axes. The motion control is facilitated when the static and dynamic characteristics are the same. The inertia moments along both axes determine the maximum acceleration which can be reached applying the given system and implementing a respective control method. Most frequently mechanical characteristics along both axes of motion differ

significantly which requires the use of complicated algorithms of motion along the set trajectory.

The executive motor is not a less important unit in the positioning system. In spite of the great variety of executive motors in plotters, mainly two types are applied: stepping and low-inertia direct current motors. The advantages of these two types of motors compared to the rest, are: direct current power supply and easy control by digital circuits.

Depending on the type of executive motor, the system can be digital or digital-analogue. The use of stepping motors results in simplification because an open digital system is obtained without position and velocity feedback. In this case it is counted upon processing each step set by the control device. A major problem is combining high speed of motion with high precision of positioning. Speed is limited by the maximum operating frequency of the motor and the precision – by the number of steps per revolution. Resonance phenomena which can occur in the system when the mechanic resonance frequency coincides with the command pulse frequency of the stepping motor are also a problem. These deficiencies can be avoided when a direct current motor is employed. For this purpose it is necessary to use a digital-analogue structure of the control system closed by position-of-the-operating-mechanism feedback. Information is provided by a special sensor which makes the system more complicated.

The choice of a system is determined by the purpose of the device. The use of stepping motor is purposeful in small inexpensive plotters which do not have demands for high speed and high precision of drawing. Higher speed, acceleration and precision are achieved by more complicated and expensive control systems realized by direct current executive motors.

Method of Estimate Function

Fig.5.3 shows the fundamental blocks of a system with digital-analogue structure which incorporates a direct current low-inertia motor and a photo-raster transducer connected to it.

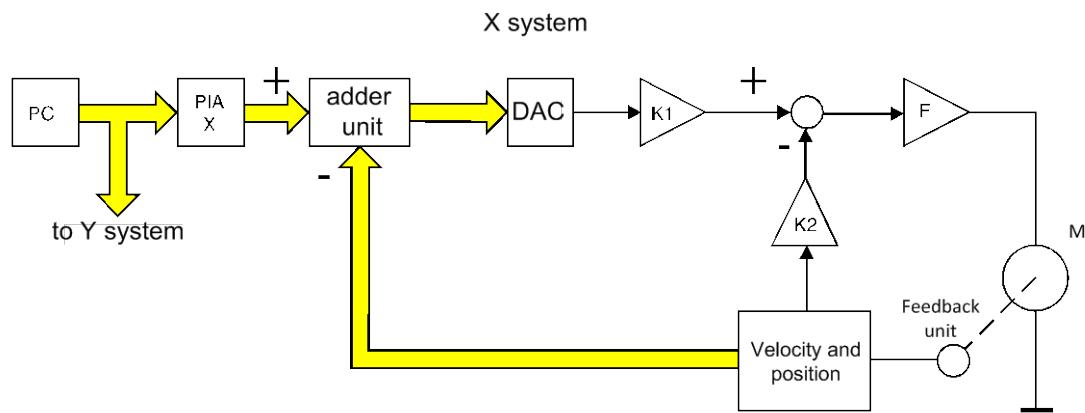


Fig. 5.3. Control system with speed and position feedback

Over a certain period of time the microcomputer loads a new value of the set position into the peripheral adapter. The value is determined by interpolation by the method of estimate function which is illustrated in Fig. 5.4.

Let the executive mechanism is located at point $M(X_i, Y_j)$, after performing the serial elementary step, and let $f_{i,j}$ be the difference between the angular coefficients of segments OK and OM. Then:

$$f_{i,j} = \frac{X_i}{Y_j} - \frac{X_k}{Y_k} = \frac{X_i Y_k - X_k Y_j}{Y_j Y_k}. \quad (5.1)$$

If denominator $Y_j \cdot Y_k$ is ignored, then

$$F_{i,j} = X_i Y_k - X_k Y_j. \quad (5.2)$$

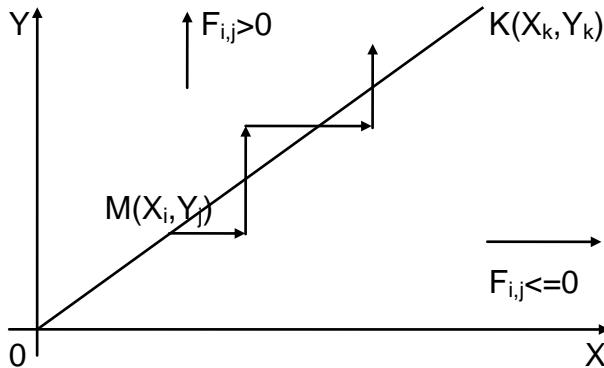


Fig. 5.4. Interpolation by the method of estimate function

Quantity $F_{i,j}$ is called estimate function and it is positive or negative depending on which side of the straight line OK the current point M lies. After a step along X axis, the new value of the current coordinate X_{i+1} will be $X_i + 1$, then:

$$F_{i+1,j} = Y_k(X_i + 1) - Y_j X_k = Y_k X_i - Y_j X_k + Y_k = F_{i,j} + Y_k \quad (5.3)$$

or $F_{i,j} > 0$, whereat there will be an increment along Y.

If the step is along Y axis, then the estimate function has the form:

$$F_{i,j+1} = X_i Y_k - X_k(Y_j + 1) = X_i Y_k - X_k Y_j - X_k = F_{i,j} - X_k, \quad (5.4)$$

or $F_{i,j} < 0$, whereat there will be an increment along X.

The algorithm by which incrementation is carried out is the following:

a) after each elementary step along X or Y the new value of the estimate function is calculated by the set formulae;

b) if $F_{i,j} > 0$, the coordinate is incremented along Y;

c) if $F_{i,j} \leq 0$, the coordinate is incremented along X.

The set position value calculated in this way does not differ by more than a step. The position feedback has been calculated so that for each elementary step, a certain number of pulses enter from the feedback unit. In this unit pulses are decoded by the photo-raster transducer, the direction of motion is determined, a signal is generated proportional to velocity of motion and information about the current position of the operating pen is given. In the adder unit the set and current positions are added. Thus at the adder outputs the difference between these two positions is obtained, it is converted by the

digital-to-analogue converter and it is used for setting the required speed of motion. The feedback unit produces also an analogue signal proportional to the speed of motion. In the adding point enter the position error signal with coefficient K1 and speed setting signal with transmission coefficient K2 which has a negative sign. The remainder from the two signals is passed to the terminal amplifier. When the difference between the set and processed positions is the least the amplification is the strongest. Thus the elementary step is processed with the required precision. When the differences are greater, corresponding to motion at higher speed, amplification decreases more and more and in this way stability of the system is achieved and the voltage of the motor is limited.

Method of Control by Speed

This method allows more flexible motion control based on the same executive mechanism. In practice, the system allows a speed of 60 cm/s, maximum acceleration of 40 cm/s² and minimum displacement of 25 μm to be reached. Fig. 5.5 shows a block diagram of the positioning system.

Unlike the first system, here there is no position feedback activated by a circuit, which allows the system to be controlled by speed. At a regular period of time the microcomputer sets the speed of motion in digital form which is converted into analogue form by a digital-to-analogue converter.

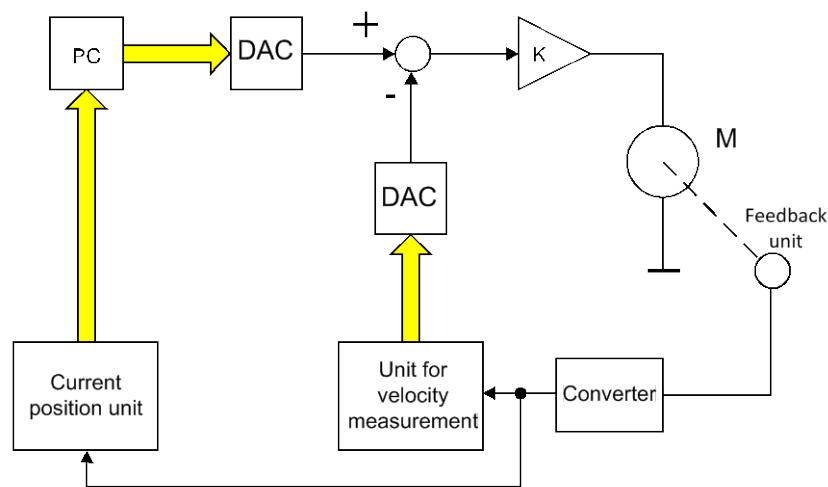


Figure 5.5. Control system with speed feedback

The pulses from the photo-raster sensor are passed to transducer input which multiplies and decodes them in order to determine the direction of motion. The transduced pulses go to the units for measuring speed and position. The unit for measuring the current position consists of bidirectional counters where the pulses accumulate. The difference between the set and measured speed is amplified by the terminal amplifier which has a linear characteristic. Periodically the microprocessor receives information about the current position by reading the contents of the counters in the unit for position measuring. The system maintains a constant vector speed which does not depend on the direction of motion.

When there is motion at angle α towards X axis

$$\begin{aligned} V_x &= V \cos \alpha \\ V_y &= V \sin \alpha , \end{aligned} \quad (5.5)$$

where V is vector speed, V_x is the speed along X, and V_y is the speed along Y. Speed variation along X and Y axes has the form shown in Fig.. 5.6.

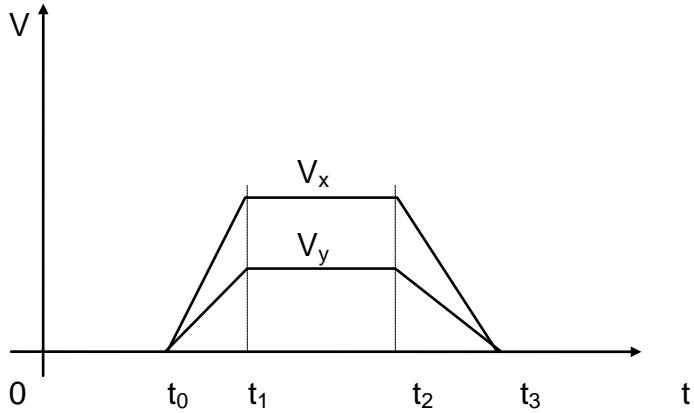


Fig. 5.6. Speed variation along X and Y axes.

To keep the motion along a set trajectory unvaried, which in this case is a straight line, it is necessary the relation between speeds to be constant. In the first section ($t_0 \div t_1$) acceleration is positive. Then, for each cycle speeds increase by ΔV_x and ΔV_y , whereat:

$$\Delta V_x = \frac{V_x}{n}, \quad \Delta V_y = \frac{V_y}{n}, \quad (5.6)$$

$$V_x = n\Delta V_x, \quad V_y = n\Delta V_y. \quad (5.7)$$

The number n is preset and it determines the required speed of motion. In this section acceleration along both axes must be kept constant. In the section between t_1 and t_2 the motion has constant speeds and acceleration equal to 0. The control aims at keeping V_x and V_y constant. In the third section ($t_2 \div t_3$) the speed decreases by ΔV_x and ΔV_y , by creating constant acceleration with a negative sign. [25]

5.2.3. Graphic Language for Plotter Control HP-GL

The graphic language HP-GL, used as an input language for plotters, consists of instructions with two-letter mnemonics. The HP-GL syntax is illustrated in Fig. 5.7. [2]

Parameter fields are filled in format specified by the syntax of the respective instruction. For example, the following formats are used in plotter HP7475A:

Integer – the parameter is an integer between -32768 and +32767. Fractional parts of parameters which must be integers are cut out. If there is no sign, the parameter is assumed to be positive.

Decimal – a number between -128,0000 and +127,9999 with non-compulsory decimal point and fractional part which contains up to four significant digits. If there is no sign, the parameter is considered positive.

Scaled decimal format – a number between -32768,0000 and +32767,9999 with non-compulsory decimal point and fractional part which contains up to four significant digits. If there is no sign, the parameter is considered positive.

Label boxes are a combination of texts, numerical expressions and variable strings.

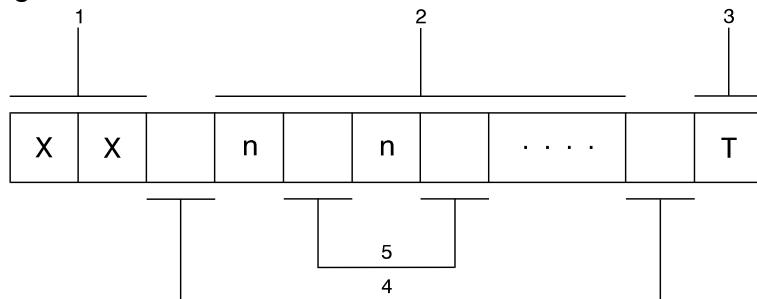


Figure 5.7. Instruction syntax in HP-GL.

1. Two-letter mnemonic code.
2. Parameter field (if necessary).
3. Instruction terminator.
4. Optional divider.
5. Compulsory divider.

Instructions for Initial Positioning

Positioning instruction by default - DF.

Syntax: DF t

Positioning instruction by default sets some functions to the plotter in a state specified in advance. This instruction is used to return the plotter to a certain state. As a result, the undesired states of graphic parameters as: character size, slope or scale are not transferred to the next program. Parameters are not used. If a parameter is set, the instruction will be executed and an error will be found.

Initialization instruction - IN.

Syntax: IN t

Initialization instruction IN gives initial state to the plotter which is set when the power is switched on. This instruction does not affect the exchange protocol and the state of the plotter (programmed turned-on or programmed turned-off).

The instruction is used to return the plotter to a certain state at the beginning of the graphic program, so that some undesired graphic parameters as size and slope of the letters and scaling from the previous program to be eliminated.

Parameters are not necessary. If parameters are set, the instruction will be executed and an error will be found.

Input mask instruction - IM.

The input mask instruction IM controls the conditions under which an error state message is received; the states which cause interrupt request message and the states which cause a positive answer to a parallel query.

This instruction is used to change the conditions under which error state message is given.

Syntax: IM E-mask (, S-mask (, P-mask)) t
 or
 IM t

In configuration with RS232, S- and P-masks are not used and if they are created, they are ignored.

The E-mask value determines which HP-GL errors cause the bit error to be found and the ERROR LED to be lit up on the plotter front panel.

Instructions for Determining Limits and Units

Instruction for inputting scaling points P1 and P2 - IP.

Syntax: IP P1x, P1y, P2x, P2y t

Scaling points P1 and P2 determine the size of the drawing area (the size of the drawing) and the direction of inscriptions. The P1 and P2 coordinates must be in absolute units. It is not obligatory to preset P2. If it is missing, P2 is determined according to P1, so that the distance between P1 and P2 remains unchanged.

Instruction IP t inputs P1 and P2 by default for the given format A3 or A4.

Instruction for outputting scaling points P1 and P2 - OP.

Syntax: OP t

After accepting the instruction, the plotter outputs P1 and P2 in plotter units as four integers in the following form:

P1x, P1y, P2x, P2y t,

where t is the output terminator.

Instructions for scaling - SC.

Syntax: SC Xmin, Xmax, Ymin, Ymax t

It sets the coordinate system in user's units and acquires the values of the scaling points P1 and P2. The drawing area is divided into plotter units (1 plotter unit is equal to 0,025 mm). By SC instruction the users define their own users' units.

SC t instruction cancels scaling.

Instruction for window input - IW.

Syntax: IW X1, Y1, X2, Y2 t,

where:

X1,Y1 are the coordinates of the lower left corner;

X2,Y2 are the coordinates of the upper right corner.

This instruction allows the pen motion to be limited up to a rectangular part of the drawing area called 'window'. Parameters are set in absolute plotter units.

IW t instruction returns the system in its initial state.

Instruction of window output - OW.

Syntax: OW t

It outputs the coordinates of diagonal points of the defined window in plotter units in the following order:

X1, Y1, X2, Y2 t,

where:

X1,Y1 are the coordinates of the lower left corner;

X2,Y2 are the coordinates of the upper right corner.

Instruction for outputting physical limits - OH.

Syntax: OH t

It outputs the coordinates of the upper left corner and the lower right corner of the current physical limits, i.e. of the maximum available area as integers in ASCII form:

X1, Y1, X2, Y2 t,

where:

X1,Y1 are the coordinates of the lower left corner;

X2,Y2 are the coordinates of the upper right corner.

Instruction for rotating the coordinate system - RO.

Syntax: RO (angle in degrees) t

It rotates the coordinate system in plotter or user's units to 90°, i.e. it orients the drawing area horizontally or vertically. Allowed values for the angle are 0° and 90°.

RO t is equivalent to RO 0 t. This instruction returns the coordinate system to its initial state.

Instructions for Pen Control and for Drawing

Instructions for pen control - PU and PD.

These instructions are used to lift and drop the pen during drawing. They can be used with parameters in order to perform drawing or motion up to the points given by the parameters.

Syntax: PU t or PD t

PU X,Y(...) t or PD X,Y(...) t

When parameters are not set, the instruction for lifting the pen PU lifts the pen without moving it to a new point. The instruction for dropping the pen PD lowers the pen without moving it to a new point if the pen is inside the defined window.

If parameters are set, the pen will move in turn to the given coordinates X,Y. Coordinates are interpreted in plotter units if scaling is off and in user's units if scaling is on. If parameters are set, both coordinates X,Y must also be set. An odd number of parameters will find an error state but the pen will draw until all X,Y pairs preceding the single parameter are exhausted.

Instruction for pen selection - SP.

This instruction is used to put a pen in the holder. It selects the colour and thickness of the pen during executing a drawing program. If the

instruction has a zero parameter or it is without parameters, it returns the current pen from the holder to its nest before the end of the program.

Syntax:

SP pen number t
or
SP t

The number of pens can be different. It depends on the plotter model. The fraction part is removed. Zero parameter or lack of parameter returns the pen to where it was taken from. If the nest is occupied, the plotter will send an error message. When the pen parameter is out of range, it is ignored and the pen number is not changed.

Instruction for drawing in absolute coordinates - PA.

The instruction for drawing in absolute coordinates PA moves the pen to the point shown by the values of X and Ycoordinates.

The PA instruction is used together with PD instruction, to draw a line or with PU, to move the pen to a given point in the drawing. The instruction can be executed without any parameters in order to specify drawing in absolute coordinates for instructions PU and PD with parameters. In this case, the parameters of PU and PD are interpreted as absolute coordinates of X and Y, until the instruction for relative coordinates PR is adopted.

Syntax:

PA coordinate X1, coordinate Y1,(, coordinate X2, coordinate Y2,.., coordinate Xn, coordinate Yn) t
or

PA t

Instruction for drawing in relative coordinates - PR.

The instruction for drawing in relative coordinates PR moves the pen with regard to the current position by the number of units specified in the increment parameters along X and Y. The instruction for drawing in relative units can be used for drawing lines and for moving the pen to a set point. With the PR instruction the pen motion is relative towards its current position. The instruction can also be executed without parameters. In this way relative drawing is set by PU or PD instructions. The PR instruction is often used for multiple drawing of a figure on the drawing, for example, to draw several rectangles with one and the same size.

Syntax:

PR increment along X1, increment along Y1 (, increment along X2, increment along Y2,..., increment along Xn, increment along Yn) t
or

PR t

The PR instruction requires the increments along X and Y to be in pairs. An odd number of parameters will determine an error state but the pen will draw until all X,Y pairs preceding the single parameter are finished.

The parameter sign determines the direction of motion: positive values move the pen in positive direction, and negative values – in negative direction.

Instruction for drawing a circumference - CI.

The instruction is used for drawing circumferences. All calculations are carried out by the plotter which relieves the computer load.

Syntax: CI radius (,chord angle) t

The radius parameter is a positive or negative number in integer or scaled decimal format. The sign determines the initial point of drawing the circumference: a circumference with a positive radius starts being drawn from point with 0° , and with a negative radius – from the point with 180° . The circumference centre is in the current position of the pen (Fig. 5.8). If scaling is off, the radius is in plotter units. If scaling is on, the radius is in user's units. If user's units are not identical along X and Y axes, an ellipse will be drawn. The chord angle parameter is set in integer format and accounts for the smoothness of the circumference. It is interpreted in degrees and sets the maximum angle corresponding to the chord which is drawn instead of the arc from the circumference (Fig. 5.8). The real angle used can be altered by the plotter so that all chords have one and the same length. The parameter sign is ignored.

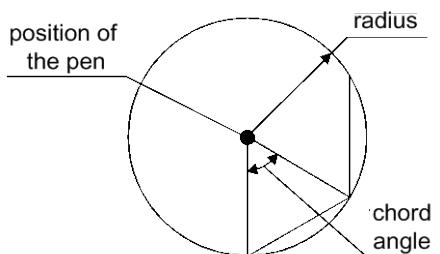


Fig. 5.8. Drawing a circumference by CI instruction

The most common values of the chord angle are within the range from 0° to 180° . Angle of 0° produces the smoothest circumference, and increasing the values decreases the number of points used.

Instruction for drawing an arc in absolute coordinates - AA.

It allows an arc to be drawn whose centre is at a point with set absolute coordinates. The arc is drawn clockwise and counterclockwise.

Syntax:

AA coordinate X, coordinate Y, arc angle (, chord angle) t

The AA instruction requires both X and Y coordinates to be set in integer or scaled decimal format. They are interpreted in plotter units if scaling is off and in user's units if scaling is on. X and Y coordinates specify the arc centre (Fig. 5.9). It can be located out of the drawing area.

The initial point of drawing an arc is the current position of the pen. The arc angle is in degrees in integer format. When the value of the angle is positive, the arc is drawn counterclockwise regarding the current position of the pen, and when it is negative – clockwise. The chord angle parameter is in integer format and accounts for the smoothness of the arc in the same way as in the instruction for circumference CI.

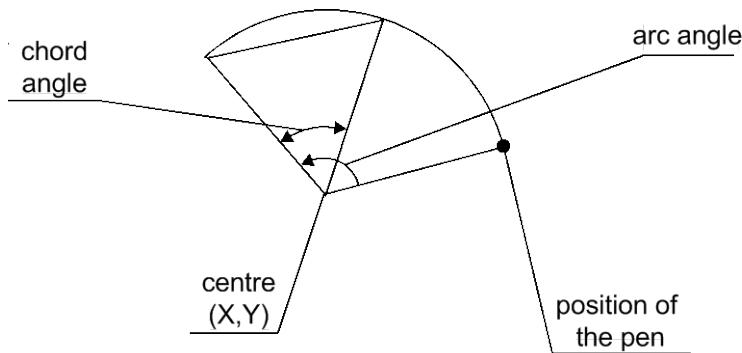


Fig. 5.9. Drawing an arc

Instruction for drawing an arc in relative coordinates - AR.

This instruction sets the drawing of an arc in relative coordinates clockwise and counterclockwise. The arc angle and chord angle are set. The instruction is used for drawing arcs with arbitrary radii, length and smoothness. The arc is drawn starting from the current position of the pen, and the position of its centre is set in relative coordinates.

Syntax:

AR increment along X, increment along Y, arc angle (,chord angle) t

The instruction requires both increments along X and Y to be set as well as the arc angle. The increments along X and Y determine the arc centre with regard to the current position of the pen. The signs of these values define the relative position of the arc centre.

The initial point of drawing the arc is in the current position of the pen. When the angle is positive, the arc is drawn counterclockwise, and when it is negative – clockwise. The chord angle parameter is in integer format and accounts for the smoothness of the curve.

Instruction for drawing a rectangle in absolute coordinates - EA.

This instruction draws the contours of a rectangle.

Syntax: EA coordinate X, coordinate Y t

This instruction requires both coordinates X and Y to be set. Drawing starts from the current position of the pen. The parameters set the absolute coordinates of the opposite vertex of the rectangle.

Instruction for drawing a rectangle in relative coordinates - ER.

Syntax: ER coordinate X, coordinate Y t

The instruction draws the contours of a rectangle, similarly to EA instruction. The current position of the pen is at the initial point. The instruction parameters specify the displacement of the coordinates of the opposite vertex of the rectangle. The values for X and Y are set and interpreted as in EA instruction.

Instructions for Adding Specific Features to the Drawing

Instruction for selecting speed (velocity) - VS.

The instruction is used for setting speed different from the speed by default and acceleration different from the acceleration by default. This instruction allows speed to be reduced to minimum. When speed is reduced, the line produced on each carrier (paper) becomes a bit thicker.

Syntax: VS speed of the pen t

or

VS t

The VS instruction without parameters sets the speed of the pen by default. The parameter depends on the type of plotter.

Instruction for the type of filling-in - FT.

The FT instruction chooses the type of filling-in in pie charts and bar charts and in other kinds of diagrams.

Syntax: FT type of filling-in (, distance (, angle)) t

or

FT t

There are 5 types of filling-in:

1. Dense (the lines are spaced at a distance specified by PT instruction, filling-in is bidirectional).
2. Dense – unidirectional filling-in.
3. Parallel lines.
4. Crossed lines.
5. No filling-in.

The filling type parameter must be always an integer between 1 and 5. If the filling type is not preset, it is assumed to be one.

The distance parameter sets the distance between the parallel lines in the area to be filled-in.

If the distance is not preset, it is assumed to be by default equal to 1% of the diagonal distance between point P1 and P2. 0 value of the distance is ignored and it is assumed to be equal to the currently determined pen thickness in PT. The distance parameter is ignored for dense filling-in of type 1 and 2 and the distance is specified by PT instruction.

The angle parameter determines the slope of the line and it is set for every 45°, starting from 0°. When an angle of 0° is set, horizontal lines are produced, when the set angle is 45° - inclined lines, and when the set angle is 90° - vertical lines. If no angle is set, it will be set by default at 0°.

Parameters set out of the range are assumed to be an error and the instruction is ignored.

Instruction for pen thickness – PT.

This instruction is used with instructions FT, RR, RA and WG for producing dense filling-in of pie charts, bar charts and graphs.

Syntax: PT pen thickness t

or

PT t

The pen thickness is a decimal number, setting the physical thickness of the pen in millimeters within the range of 0,1mm to 5,0mm (the optimum range is from 0,3 to 0,7mm). If the pen thickness is not preset, it is assumed to be 0,3mm by default.

Values set out of the range are assumed to be an error and the instruction is ignored. If more parameters have been set, the plotter executes the instruction only with the first parameter, finds an error and ignores the remaining parameters. The PT instruction concerns only the currently chosen pen. It is valid until:

a) a new pen is chosen with SP instruction or manually from the front panel of the plotter;

b) a new PT instruction is received.

Instruction for filling-in a rectangle in absolute coordinates - RA.

Syntax: RA coordinate X, coordinate Y t

The RA instruction requires both coordinates X and Y to be set (coordinate pair). The initial point of the rectangle is the current position of the pen, and X and Y coordinates determine the opposite vertex of the rectangle.

The RA instruction without parameters is ignored but an error is not found. Parameters set out of the range are considered to be an error and the instruction is ignored. If only one parameter is set, the instruction is ignored and an error is found. If more parameters are set, the instruction is carried out with the first two parameters and an error is found. The remaining parameters are ignored.

Instruction for filling-in a rectangle in relative coordinates - RR.

Syntax: RR increment along X, increment along Y t

The instruction is used for defining and filling-in a rectangle in relative coordinates. The initial point of the rectangle is the current position of the pen, and the increments along X and Y determine the opposite vertex. The rectangle is filled-in by using the current position of the pen and the current type of the line. The values for X and Y are set and interpreted as in RA instruction.

Instruction for filling-in a circular sector (wedge) - WG.

The instruction is used for drawing pie charts. It is also possible to draw triangles, rhombuses, pentagons, hexagons, octagons, etc.

Syntax:

WG radius, initial angle, arc angle (,chord angle) t

The instruction determines and fills-in a circular sector (wedge) (Fig. 5.10), by employing the current pen and line type. The current position of the pen determines the centre of circumference. The radius determines the size of circumference and it is set in integer or scaled decimal format. The radius sign specifies where the point with 0° is located, towards which the initial angle and arc angle are measured.

The initial angle determines the point from which the drawing starts. It has values from -360° to 360° . A positive initial angle positions the radius counterclockwise from the point with 0° . The arc angle is in integer format and determines the number of degrees corresponding to the circular sector being drawn. A positive angle draws counterclockwise.

The chord angle is in integer format from 0° to 120° and accounts for the smoothness of the arc. The chord angle is assumed to be 5° by default.

After the pen finishes drawing it returns to its initial position.

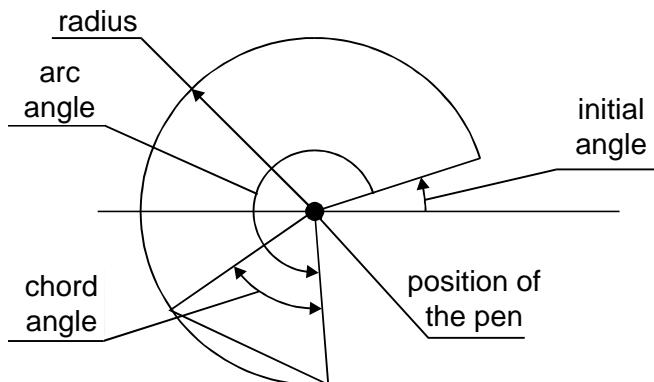


Fig. 5.10. Parameters which determine a circular sector (wedge)

Instruction for drawing a circular sector (wedge) - EW.

Syntax:

EW radius, initial angle, arc angle (, chord angle) t

The instruction is used for drawing separate arcs which after combining can form a pie chart. The instruction draws a circular sector (wedge) by employing the current pen and line type. The values and range of the instruction parameters are similar to the ones in WG instruction.

Instruction for ticks - XT, YT.

The XT instruction draws vertical ticks along X axis, and YT instruction – horizontal ticks along Y axis. The drawing of ticks starts from the current position of the pen.

Syntax: XT t
 YT t

The instructions do not require parameters. The length of the tick is specified by TL instruction. By default it is assumed to be 0.5% of ($P2x - P1x$) for YT and 0.5% of ($P2y - P1y$) for XT.

Instruction for tick length - TL.

Syntax: TL positive part (, negative part) t
 or
 TL t

The instruction sets the length of the ticks which are drawn by XT and YT instructions. The tick length is specified in percentage of the horizontal and vertical distance between scaling points P1 and P2.

The instruction sets positive and negative part of ticks. If only one parameter is set, the negative part of the tick will be suppressed. If first parameter 0 is set, the positive part of the tick will be suppressed.

The negative value of the parameter changes the direction of drawing. If negative value is set to the positive part of the tick, this part will be drawn as

negative and vice versa – if the negative part has a negative value, it will be drawn as positive part of the tick.

If a tick value of 100 is set, a coordinate grid is drawn. The value of each parameter by default is 0.5.

Instruction for symbol mode - SM.

The instruction is used together with instructions PA and PR and allows writing a symbol centered at the end of each vector. Drawing in symbol mode is used for writing a specific symbol at a given point. In this way dispersion time diagrams, geometrical drawings, or charts with many lines are produced.

Syntax: SM symbol t

or

SM t

Any printed symbol can be set as instruction parameter (decimal value of ASCII code from 33 to 126) except the symbol ; (ASCII код 59), which serves as a terminator and switches off the symbol mode. The first symbol after the mnemonics is interpreted as parameter.

After carrying out instruction SM, instructions PA and PR act as usual, but at the end of each vector the symbol set in the parameter margin is written. This goes on until the symbol mode is cancelled.

The SM instruction without parameters switches off the symbol mode.

Instruction for the line type - LT.

Syntax: LT type number (, type length) t

or

LT t

The instruction sets the type of line which is drawn with instructions PA and PR. The first parameter (type number) is given values from 0 to 6 and specifies the type of line. When the value of this parameter is larger than 6, it is ignored, the type of line remains the same and no error is found. The types of lines used are shown in Fig. 5.11.

The second parameter (type length) determines one cycle of the type and it is set in percentage of the diagonal distance between the scaling points P1 and P2. By default the length is determined as 4% of $|P_2 - P_1|$.

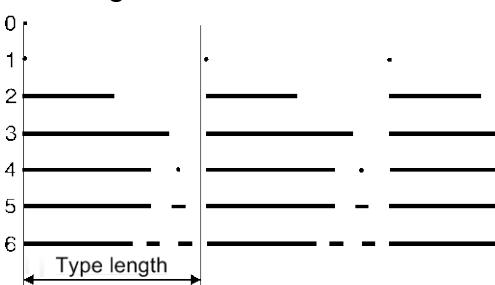


Fig. 5.11. Types of lines:

1. Dot at the beginning;
2. Line of dots;
3. Broken line;
4. Line broken by a dot;
5. Line broken by a dash;
6. Line broken by two dashes.

The LT instruction without parameters defines drawing a continuous line.

Instruction for Inscriptions

Plotters can inscribe with one of the 19 internally built-in sets of symbols. The symbols and punctuation are different in the different sets and correspond to several different languages. It is possible to operate with two sets of symbols: one is standard and the other is alternative. When the plotter is initialized, 0 set is chosen automatically as standard and alternative. To write texts in Bulgarian, it is necessary to choose and appoint a set containing the Cyrillic alphabet.

Instruction for appointing a standard set of symbols - CS.

Syntax: CS symbol set number t
 or
 CS t

The instruction replaces the appointed standard set of symbols with another one. Depending on the language chosen for inscriptions the respective set is chosen. It is used in all inscription operations when the standard set of symbols is chosen. The CS instruction without parameters chooses 0 set.

Instruction for appointing an alternative set of symbols - CA.

Syntax: CA symbol set number t
 or
 CA t

The instruction replaces the appointed alternative set of symbols with another one. This set is chosen according to user's requirements and it is used in all inscription operations. The CA instruction without parameters chooses 0 set.

Instruction for choosing a standard set of symbols - SS.

Syntax: SS t

The instruction chooses the standard set of symbols for inscription. It is used for replacing the chosen alternative set of symbols with the standard one. The choice of the standard set is made by CS instruction or by the controlling symbol SI (ASCII code 15). The instruction does not require any parameters.

Instruction for choosing an alternative set of symbols - SA.

Syntax: SA t

The instruction chooses the alternative set of symbols for inscription. It replaces the chosen standard set of symbols with an alternative one. The choice of the alternative set of symbols is made by CA instruction or by the controlling symbol SO (ASCII code 14). The instruction does not require any parameters.

Instruction for defining an inscription terminator - DT.

Syntax: DT inscription terminator t
 or
 DT t

The inscription terminator is the final symbol in the inscription string. When such a symbol is reached, inscription is stopped and the next symbols are interpreted as HP-GL instructions. By default as an inscription terminator ETX (ASCII code 3) is used but it can be changed by DT instruction.

If one of the controlling symbols (ASCII codes 1÷32 и 127) is used as an inscription terminator, at the end of the inscription the symbol is not written, though its function is fulfilled. If one of the symbols with ASCII codes 32÷126 is used, the symbol is written at the end of the inscription. The symbols: NULL (ASCII code 0), ESC (ASCII code 27) and ENQ (ASCII code 5) cannot be used as terminators.

Instruction for inscription - LB.

Syntax: LB c...c t

The instruction writes a string of symbols with the currently chosen set of symbols. The inscription starts from the current position of the pen. After writing a symbol the pen moves to the lower left corner of the next symbol space. The values set by default are used for the direction of writing, the size and slope of symbols, unless they are preset by instructions DI, DR, SI, SR or SL. The inscription mode is stopped when a terminator for ending the inscription is sent, chosen by DT instruction.

Instruction for absolute direction - DI.

Syntax: DI abscissa, ordinate t
 or
 DI t

The instruction is used for changing the direction of writing into a new, absolute direction i.e. direction independent of the position of P1 and P2. The slope angle of the inscription (α) is determined as: $\alpha = \arctg \frac{y}{x}$, where: $x = \cos\alpha$ - abscissa, $y = \sin\alpha$ - ordinate (Fig. 5.12).

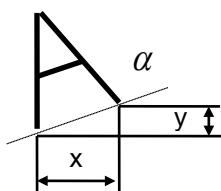


Fig. 5.12. Direction of inscription

The abscissa and ordinate are set in decimal format and at least one of the parameters must be different from 0. An instruction without any parameters is equal to instruction DI 1,0; This instruction determines a horizontal inscription directed from left to right.

Instruction for relative direction - DR.

Syntax: DR abscissa, ordinate t
 or
 DR t

The instruction is used for changing the direction of inscribing into a direction which is set relatively with regard to the positions of points P1 and P2. The change in the scaling points affects the direction of inscribing. The angle of inscription slope (α) is determined as: $\alpha = \arctg \frac{y}{x}$, where: $x = \cos\alpha$ - abscissa, $y = \sin\alpha$ - ordinate (Fig. 5.12).

Unlike instruction DI, the abscissa and ordinate here are set as percentage of the distance between points P1 и P2.

abscissa = the desired % of $|P2x - P1x|$;
 ordinate = the desired % of $|P2y - P1y|$.

Instruction without any parameters is equivalent to instruction DR 1,0; This instruction determines a horizontal inscription directed from left to right.

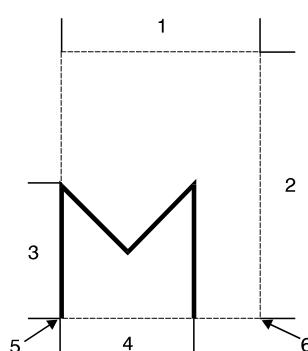
Specifications for the size of symbols.

Each symbol is placed in symbol space which has a height twice the height of the symbol and a width which is 1.5 of the width of the symbol. The space above and by the sides of the symbol serves as spacing between lines and spacing between symbols. The parameters shown in Fig. 5.13 are defined to describe symbols.

Instruction for absolute size of symbols - SI.

Syntax: SI width, height t
 or
 SI t

The instruction specifies the real size of symbols in centimeters. By default are set: W=0,187 and H=0,269 for A4 format and W=0,285, H=0,375 for A3 format. Negative values of parameters result in producing a mirror image of the inscription.



1. Width of symbol space = W.
2. Height of symbol space = H.
3. Height of symbol = 0,5H.
4. Width of symbol = 0,67W.
5. Initial point of the symbol.
6. Initial point of the next symbol.

Figure 5.13. Parameters of symbol space

Instruction for relative size of symbols - SR.

Syntax: SR width, height t
 or
 SR t

The instruction specifies the size of symbols as percentage of the distance between the scaling points respectively along X and Y axes:

$$W = \% \text{ of } |P2x - P1x|, \\ H = \% \text{ of } |P2y - P1y|.$$

The values of the width and height by default are:

$$W=0,75\% |P2x - P1x|, H=1,5\% |P2y - P1y|.$$

Instruction for inserting symbol spaces - CP.

Syntax:

CP number of symbol space widths, number of symbol space heights t

or

CP t

The instruction moves the pen to an arbitrary number of symbol fields or lines from a certain point. The first parameter moves the pen to the specified number of widths of symbol fields to the right (if the parameter has a positive value) or to the left (negative value). The second parameter moves the pen to the specified number of heights of symbol fields upwards (if the parameter has a positive value) or downwards (negative value). The directions: upwards, downwards, left and right depend on the direction of the inscription.

If no parameters are set, the CP instruction is used for moving to a new line. The position of the pen (up or down) is not changed by instruction CP.

Instruction for slope of symbol - SL.

Syntax: SL tg(α) t
 or
 SL t

The instruction specifies the slope angle of symbols. Instruction without parameters determines writing symbols without any slope (Fig. 5.14).

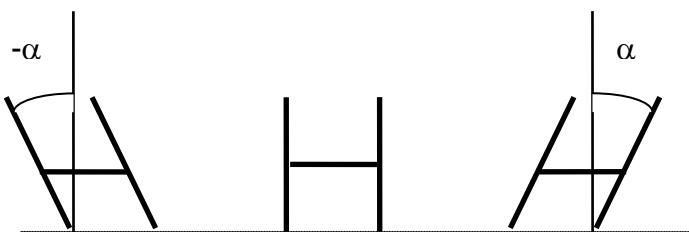


Figure 5.14. Specifications for slope of symbols.

Instruction for symbols defined by the user - UC.

The instruction is used for writing symbols which are not included in the sets of symbols of the plotter and for creating proper fonts.

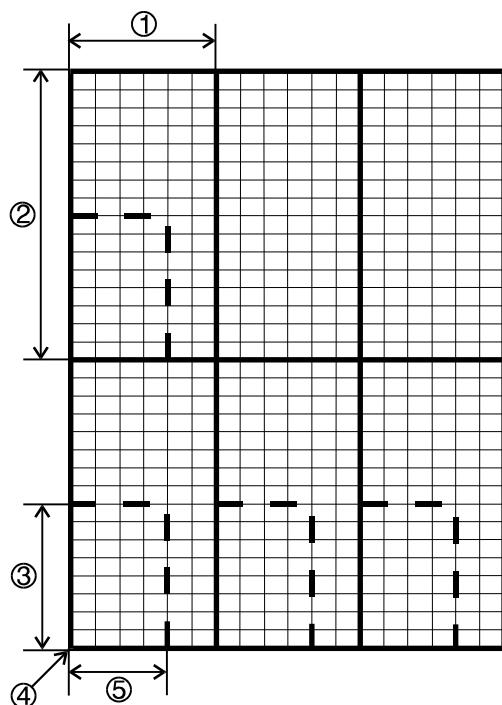
Syntax:

UC (pen control,) increment along X, increment along Y,
(pen control,) increment along X, increment along Y.....t
or
UC t

Each element of the symbol is drawn on a symbol grid. This grid is positioned on a symbol field which is divided into 6 horizontal and 16 vertical units (Fig. 5.15). The size of the symbol field and respectively of the grid units is determined by the current instruction for size. The size of the symbol space and respectively of the grid is always twice the height and 1.5 times the width of the symbol. To produce a user's symbol with the same size as the symbols written following the instruction for inscriptions, it must be placed in the lower left corner of the grid and must have a width of four units and height of eight units.

A symbol defined by the user is written in the following way:

1. Each X, Y increment is drawn by employing the latest specified state of the pen (up or down) with the latest parameter for pen control. When the instruction starts, the plotter places the pen at point (0,0) of the coordinate grid in lifted state.



1. 6 grid units.
2. 16 grid units.
3. Height of symbol specified by the current instruction for size.
4. Initial point of symbol.
5. Width of symbol specified by the current instruction for size.

Figure 5.15. Symbol grid

2. The pen moves sequentially to each point with a couple of X and Y increments. The X and Y increments must be set in couples and to be within the interval from -99 to +99. The increment along X defines in decimal format (from -98,9999 to +98,9999) the number of grid units which the pen must

pass in horizontal direction from its current position. Positive increments move the pen to the right, while negative – to the left.

The increment along Y specifies in decimal format (from -98,9999 to +98,9999) the number of grid units which the pen must pass in vertical direction from its current position. Positive increments move the pen upwards, while negative – downwards. The directions: left, right, up and down refer to the current direction of writing. Mirror images of symbols defined by the user are produced in the same manner as of the rest of the symbols. If only one increment of the X,Y couple is specified, an error is found.

3. The parameter for pen control is specific for UC instruction. The parameters for pen control are the following:

- integers larger than +99 command the pen to be down;
- integers smaller than -99 command the pen to be lifted;
- integers larger than +127,9999 or smaller than -128 find an error – an out-of-range parameter.

The position of the pen after fulfilling UC instruction becomes an initial point of the symbol. The initial increment along X,Y is towards the initial point of the symbol, and each subsequent motion is towards the latest position reached by the pen. After finishing the symbol defined by the user, the pen automatically moves by one symbol space to the right of the initial point of the symbol. This point becomes current position of the pen and therefore the initial point of the next symbol (if there is one).

The symbols defined by the user are written by using current sizes, slope and direction of symbols. It is also possible to vary the size of the symbol defined by the user, by multiplying all increment parameters along X and Y by one and the same constant.

Instructions for Digitalization

Instruction for digitalization of a point - DP.

The instruction is used for inputting data for a graphic program or for obtaining the coordinates of points from a certain curve.

Syntax: DP t

Parameters are not used. When DP instruction is received, the automatic lifting of the pen is suppressed. The current indicator for paper size on the front panel starts blinking and the plotter is ready to input a digitalized point by pressing the INPUT button on the front panel.

When INPUT is pressed, X and Y coordinates of this point and the position of the pen (up or down) are stored in order to be retrieved by OD instruction. A b2 bit of the sense byte is set which is an indication that a digitalized point is ready to be output. The automatic lifting of the pen is activated again and the paper size indicator stops blinking.

Instruction for clearing digitalization - DC.

This instruction is used as a procedure for interruption in the digitalizing program in order to make a transition to another drawing function.

Syntax: DC t

When DC instruction is received the digitalization mode is interrupted and the paper size indicator stops blinking. The automatic lifting of the pen is activated again.

Instruction for input of digitalized point and pen position - OD.

This instruction is used after instruction DP and pressing the INPUT button, in order to give the coordinates of the digitalized point to the computer.

Syntax: OD t

Parameters are not used. The position and state of the pen are output to the computer as integers in ASCII code in the following form:

X, Y, P, t , where:

X is the abscissa of the digitalized point in plotter units;

Y is the ordinate of the digitalized point in plotter units.;

P is the state of the pen when the point was input (0 - up, 1 - down);

t is a terminator for output.

The boundaries of the range of X and Y coordinates are equal to the physical boundaries of the plotter.

Instructions for Receiving Information from the Plotter

Instruction for output of the actual position and state of the pen - OA.

This instruction is used for determining the current position of the pen in plotter units. The location of an inscription or figure can be pointed or the parameters of the desired window can be defined by this instruction.

Syntax: OA t

The output is always in plotter units. Parameters are not used. The position and state of the pen are output to the computer similarly to instruction OD.

Instruction for output of the position and state of the pen from the latest instruction for the pen - OC.

This instruction is used for determining the position of the pen from the latest command in plotter or user units depending on whether the scaling is on or off. This information is useful for locating an inscription or figure, or for determining the parameters of the instruction which has moved the pen to the boundaries of a window.

Syntax: OC t

The output is in decimal format, in user's units when scaling is on and in plotter units when scaling is off. Parameters are not used. The position and state of the pen are output to the computer similarly to instruction OD.

Instruction for error output - OE.

This instruction is used for defining the type of the first error. It is useful for program debugging or for checking if all data or instructions have been received correctly by the plotter.

Syntax: OE t

Parameters are not used. When OE instruction is received, the plotter converts the first error from HP-GL into a positive integer in ASCII code which is output in the form:

error number t

The error number is defined in Table 5.1.

error number	Meaning
0	no error
1	unrecognized instruction
2	wrong number of parameters
3	out-of-range parameters
4	not used
5	unknown number of symbols
6	positional overflow
7	not used
8	vector received when paper has not been fixed

Table 5.1. Error numbers

Instruction for output of coefficients - OF.

This instruction is used for the output of the number of plotter units per millimeter along each axis.

Syntax: OF t

Parameters are not used. For HP7475A plotter the following is output:
40, 40 t

These coefficients mean that along X and Y axes there are approximately 40.2 plotter units per millimeter (resolution 0.025 mm).

Instruction for identification output - OI.

It is used for the output of the plotter identifier. It is especially useful for remote control for identifying the model of the operating plotter.

Syntax: OI t

Parameters are not used. The plotter always outputs the following string of symbols:

7475A t

Instruction for output of options - OO.

It is used for the output of eight optional parameters. It is especially useful for remote operation to know what options are available in the operating plotter.

Syntax: OO t

Parameters are not used. The plotter outputs a combination of eight integers in ASCII code, divided by commas. The options incorporated in the plotter are marked with ones as it is shown in Fig. 5.16.

0, 1, 0, 0, 1, 0, 0, 0 t

It means that instructions are included for arcs and circumferences.

It means that an option for pen selection is included.

Fig. 5.16. Format of option output

5.3. Raster Plotters

According to the known operating principles, raster plotters are classified into:

- electrochemical (electrolytic);
- electro-spark (erosion);
- thermal;
- electrostatic;
- ink-jet.

Electrochemical Plotters

These were the earliest plotters. Their operating principle is based on an electrochemical reaction which results in a dot (colour) image. It is used in telegraph facsimile receivers. Several variants of the electrochemical principle are known: anode dissolution with introducing ions into paper carrier; formation of colouring compounds; oxidizing reactions, etc. All these require special kind of paper, impregnated with a certain chemical solution, most frequently, electrolyte (hence the second name of these devices – electrolytic). The main disadvantage is the requirement: electrodes should be made of a special metal or they should often be replaced because they take part in the reaction.

Electro-spark Plotters

Electro-spark plotters use a metal-coated carrier (usually aluminium foil) with paper stuck to it, which is dot-burnt. Between this carrier and the paper a dark colouring agent is applied. On the spots where the metal foil is eroded, dark dots are produced. A method of dot burning of non-transparent synthetic foil is also known. The obtained image is used for copying. The electro-spark method also requires frequent replacing of electrodes, due to their wear. This makes the construction of the positioning head complicated and does not allow high resolution to be achieved. Another deficiency is that harmful to health products are released during burning.

Thermal Plotters

They are devices of small format working with thermo-sensitive paper. The operating principle is based on changing the colour of some substances, when heated. The paper is impregnated with the substance or the latter is applied on foil. The image is built by means of a matrix head with dot-heated electrodes or by light warming. A lot of measuring devices are equipped with such output devices. A lot of fax machines work with thermo-sensitive paper.

Electrostatic Plotters

Electrostatic plotters operate on the principle of creating an electrostatic charge on a surface which can retain a colouring substance (toner). Depending on the fact whether this surface is only a intermediate carrier or a final one, we distinguish electrostatic plotters and electrostatic printing devices.

Electrostatic printing devices were developed by the firm Xerox and at the beginning were offered as photocopiers. In these devices certain materials (e.g. selenium) can be charged electrostatically, to fix a specified colouring agent (toner) selectively and to print the built intermediate image on normal paper. Laser printers operate on the same principle. Their construction is treated in Chapter 6 – Printers – of this book.

Raster (non-printing) electrostatic plotters use special paper for plotting the image. They do not use an intermediate printing carrier and therefore they have quite simple mechanical construction. Their velocity is very high and their resolution is satisfactory.

The operating principle of electrostatic plotters is shown in Fig. 5.17. Paper coated with a film having dielectric properties is run over an electrifying head which creates dot charges on the dielectric film. In the next phase of its path, a liquid colouring agent with electrostatic properties is poured over it. Where the dot charge on the carrier is large enough, a respective amount of colouring agent (usually black) sticks to it. In the final phase of the process the carrier surface is blown and then the volatile component of the colouring agent evaporates and at the same time, the image is fixed to the surface.

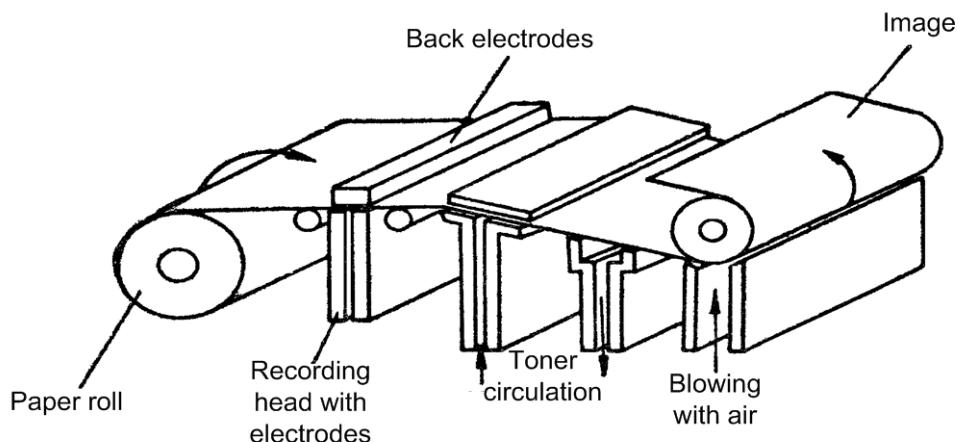


Fig. 5.17. Diagram of an electrostatic plotter

Inkjet Plotters.

By the inkjet method for producing graphic images, a non-contact process for applying ink in liquid or solid state on the surface of a paper carrier. Usually inks are magnetic or current-conductive. A number of conceptual constructions of inkjet heads are known. In one of them direct applying of information by raster-scanning method is employed (Fig. 5.18) without deflecting the jet from one or several nozzles simultaneously. Each head with independent control can have different colour. Another construction comprises a local deflection of the jet and it is used primarily for alphanumeric printing devices, the so called inkjet printers which are treated in Chapter 6 - Printers of this book. A modification of a jet head is also known. It uses solid ink which is sprayed by means of plasma charge between electrodes on which controllable high voltage is applied.

Inkjet plotters can move continuously along one of the coordinates (e.g. along the lines) as in serial printing devices or drum constructions with continuous motion. The inkjet method of producing images on paper has

been known for a long time but became widespread after a piezo-ceramic jet head was developed with suitable construction and high-frequency parameters.

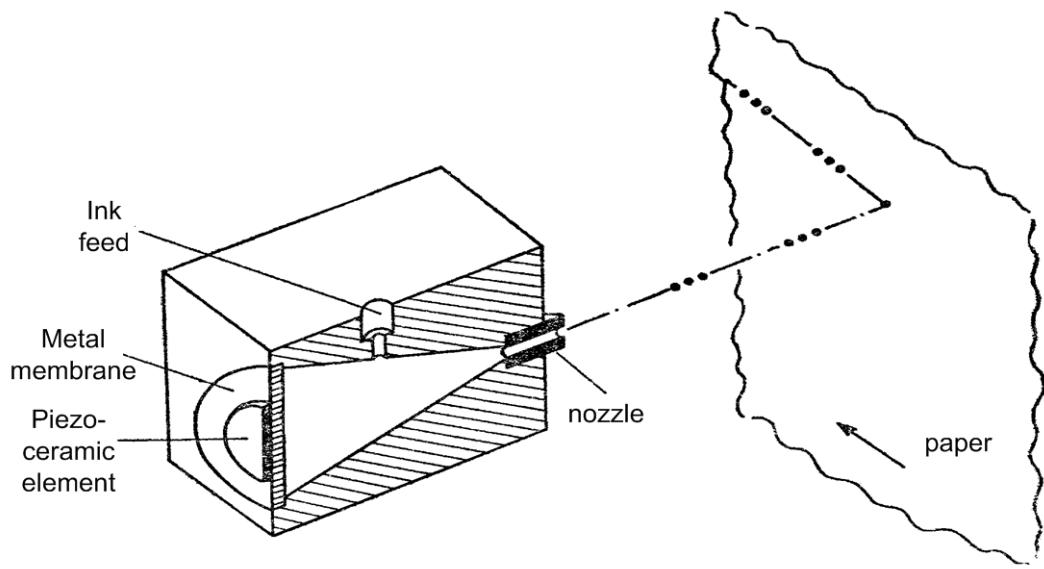


Fig. 5.18. Conceptual construction of an inkjet head

The inkjet method advantages are: non-contact and reliable operating principle at relatively low price, the usage of an ordinary paper carrier and the possibility of applying colour images. A shortcoming is the low velocity due to serial plotting and low-frequency operation of the jet head.

Inkjet plotters are used mainly for colour images. They have low prices, commensurable to the prices of serial printing devices with graphic operating mode (printer-plotters) but they have greater reliability and noiseless operation.

6. PRINTERS

Printing devices are discussed in this chapter. They are divided into two principal types: impact and non-impact. In impact mechanisms printing is performed by means of a print head which strikes an inked ribbon, placed between the head and paper. Various constructions of print heads are available. The most commonly used are described below. Non-impact printing mechanisms use different printing techniques which do not require mechanical hammers or print heads to strike the paper. Characters are formed in a non-mechanical way. Laser and jet technologies are treated in this chapter.

6.1. Daisy-wheel and Matrix Printers

Daisy-wheel Printers

Fig. 6.1 shows a 'daisy' type print head. It consists of a number of spokes fixed to an axis. The spokes resemble the petals of a daisy and hence comes the name of the print head. At the end of each spoke the relief image of a character is fixed. Usually 96 characters are used and therefore the head has 96 spokes. The ribbon and paper are positioned in front of the print head, and behind it – a small electromechanical hammer. The hammer consists of a coil with a movable core which strikes the spoke in front of it.

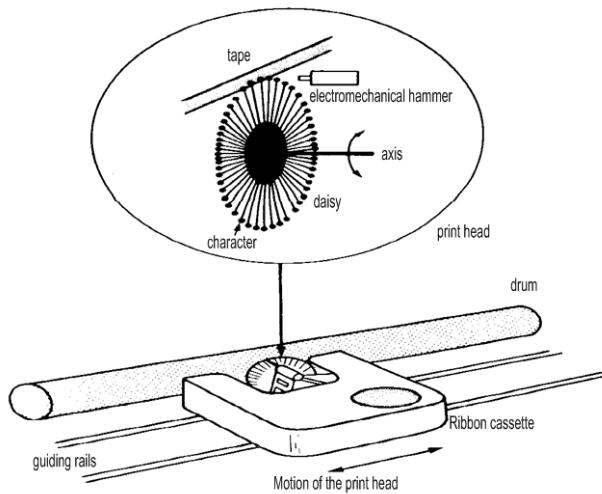


Fig. 6.1. Daisy-wheel printer

Affected by the hammer, the spoke bends and hits the ribbon which causes the printing of the character on the paper. Various characters are produced by rotating the print head before the hammer is activated. Rotation happens in two directions up to the angle of 180°. The direction of rotation is chosen so that the desired character passes the shortest distance before being printed. The most common characters are mounted on adjacent spokes. That minimizes the rotation of the print head. Daisy-wheels can be

replaced which allows working with various character sets. They are made of metal or plastic and metal daisies are of higher quality.

Printers have a built-in microprocessor which controls the printing process. The rotation direction of the print head is specified by the microprocessor. It also controls the rotation speed of the motor so that the needed character is always chosen for one and the same time. By controlling the current through the coil, the strength of the stroke of the electromechanical hammer can be adjusted. Thus different printing effects are produced.

Matrix Printers

The most important part of the dot-matrix printer is the needle print head. The characters are formed by a matrix of dots, usually 7×7 or 9×7 . One column of the matrix is printed by a corresponding number of needles, striking the ribbon and paper under the effect of the electromagnetic coils. For a column of 7 dots, 7 coils are necessary, for 9 dots - 9 coils. After the column is printed, the head moves to the position of the next column of the character space and prints it as well. So for a matrix of 7 columns, the head moves 7 times. Thus the whole matrix is printed and the head moves to the position of the next character. The process is repeated until the whole line is printed.

The manufactured solenoids are large compared with the needles and cannot be mounted without bending the needles. In the early print heads the solenoids were arranged linearly (Fig. 6.2 a), but in modern heads, they are arranged in a circle (Fig. 6.2 b) or in two lines (Fig. 6.2 c). Thus the bending of the needles is reduced and

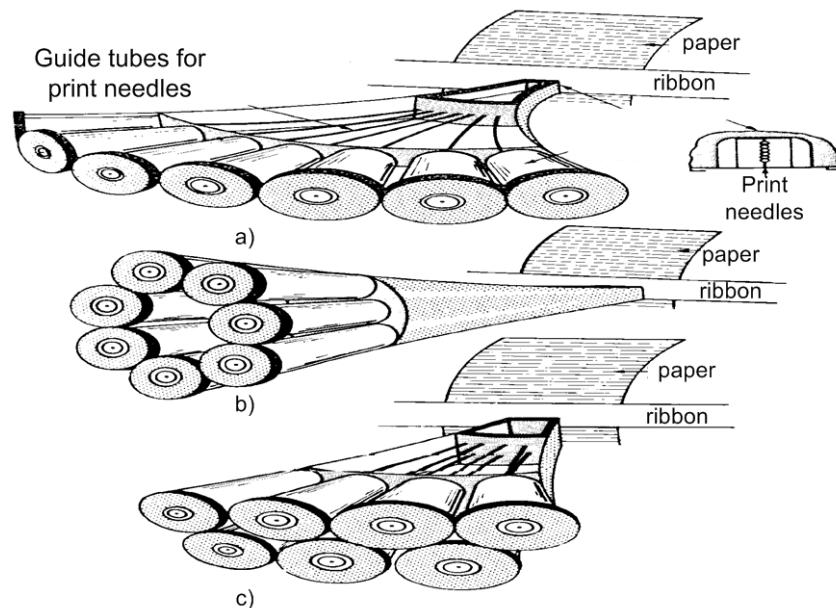


Figure 6.2. Needle print heads with:
a) linearly arranged solenoids;
b) solenoids arranged in a circle;
c) coils arranged in two lines.

the lifetime of the print heads is lengthened. Needles are made of tungsten and they can endure 10^8 strokes. Their tips wear out and with time they become shorter which results in poor print quality.

Dot-matrix printers are industrial standard for low and medium speed printing, providing reasonable quality and high flexibility, owing to the matrix format. The character set is easy to replace by means of the character generator recorder in the ROM of the printer. By several runs of the print head on the paper, double (or arbitrary) height of characters can be obtained. Thus when the 7-needle head runs twice, a 14 dot high matrix is produced. In a similar manner the width of the characters can be changed – double width is the most frequently used. It is also possible to print graphic information.

The considered print heads require flexible needles which is a serious deficiency. It is avoided by ballistic print heads. The needles in them are not connected to the cores of the solenoids, and are driven by hammers (Fig. 6.3). These hammers can be mounted closely enough to one another, to avoid bending the needles. Thus the lifetime of the print head gets longer. Moreover, in the new construction the power of the electromagnets is lower and therefore they release less heat. This allows the print head to work longer and faster (180 characters per second).

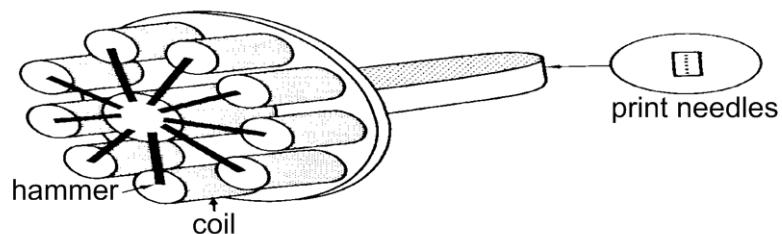


Figure 6.3. Ballistic print head

When several copies are printed simultaneously, the needle passes a shorter distance before reaching the paper and its kinetic energy is greater. Consequently the action of the hammer provides a stronger stroke which is needed when printing several copies. The construction of the ballistic head does not require adjustment when from one to six copies are printed. This means that different number of copies of documents can be printed.

6.2. Line printers

In the printers described so far, characters are formed consecutively and the printing speed reached is 10÷500 characters per second. Printers with a special construction are used to increase this speed. They allow all characters in a line to be printed simultaneously. When only one character is printed, their speed does not differ from the speed of printers with consecutive printing but when we work with a great amount of information – their advantage is obvious – a speed of 200÷2000 characters per second is reached.

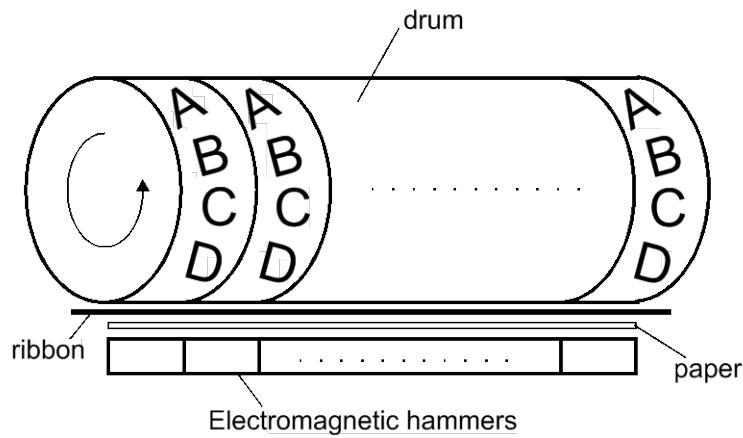


Figure 6.4. Drum printer

A drum line printer is shown in Fig. 6.4. Its operating principle is similar to the one of the daisy-wheel printer. The difference is in the use of as many electromagnetic hammers as the number of characters in a line. All hammers work simultaneously and thus for one revolution of the drum, a whole line is printed. The printing speed of these printers is n times higher than the speed of daisy-wheel printers where n is the number of characters in a line.

Chain line printers have similar construction (Fig. 6.5) but the drum is replaced by a chain on which the relief images of the characters are fixed. When the chain moves, they move along the printed line. At the moment when the needed character is opposite a position in the line, the electromagnetic hammer is activated and it prints the character on the paper.

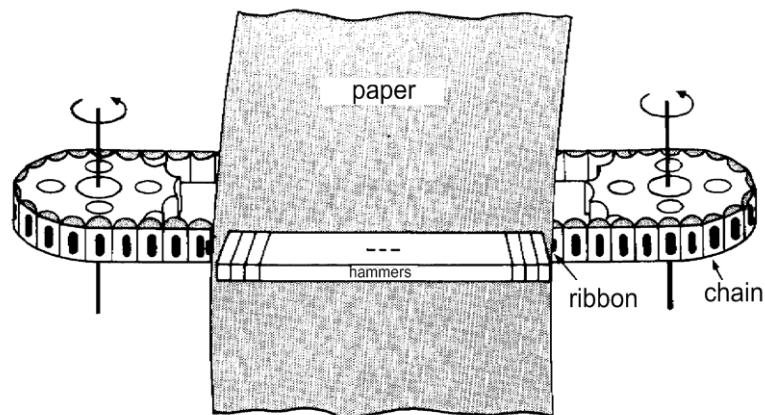


Figure 6.5. Chain printer.

After the whole line is printed in this way, the paper moves and printing the next line starts.

Besides the line printers described so far, which employ relief images of the printed characters, there are constructions in which the characters are represented as a matrix of dots. These are comb printers and needle line printers.

The printing mechanism of comb printers is shown in Fig. 6.6. The comb is made of flexible metal and has one tooth for each character in the line. At

one end of the tooth a ball of hard steel is fixed, which is used as a hammer when

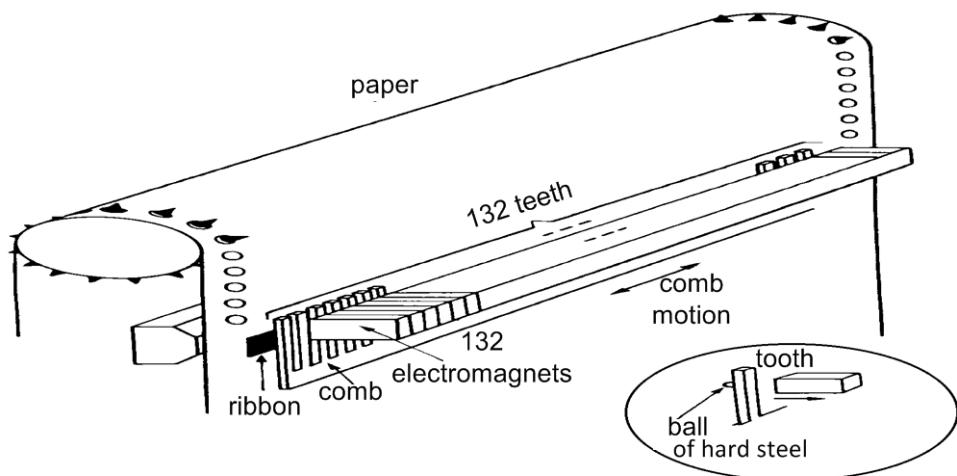


Fig. 6.6. Comb printer

forming the dots. The comb is mounted transversely to the page and the ribbon is placed between the comb and paper. A small electromagnet is positioned opposite each tooth of the comb. When the electromagnets are activated, they attract the respective teeth of the comb, and when they are deactivated the released teeth strike the ribbon and paper. Thus the first dot in the first line is formed simultaneously for all character matrices. If a certain dot should not be printed, the corresponding electromagnet is not activated. Then the comb moves to the next column of the matrix and the process is repeated until the first lines of all matrices are printed. When this is done, the paper moves vertically to print the next line of character matrices. To minimize the comb motions, even lines are printed from right to left instead of from left to right. After finishing the whole line of text, the paper moves so that the comb stands opposite the position of the next line.

The printing speed of this mechanism is within 125 lines to 300 lines per minute. The comb is driven by a step motor.

Fig.6.7 shows the construction of a needle line printer. Along the printing line electromagnetic coils are mounted, and each of them controls one printing needle. The bending of the needles is avoided in this mechanism and therefore they are considerably shorter than the needles of the printers for consecutive printing. Usually 22 coils are used which allows 22 dots to be printed simultaneously. The printing mechanism moves transversely to the paper in the same way as the comb in the comb printer, to print a line of dots. If the text line contains 132 characters and they are represented by a matrix of 5×7 dots, then for the printing of the first lines of all character matrices 30 steps are necessary ($5 \text{ dots per line} \times 132 \text{ characters} / 22 \text{ needles}$), and for printing the whole text line – 210 steps ($7 \text{ lines} \times 30 \text{ steps per line}$).

Since bending the print needles has been avoided in this printer, they can be made of a harder material which prolongs their lifetime – more than 500 million characters. A typical printing speed is 125 lines per minute (275 characters per second).

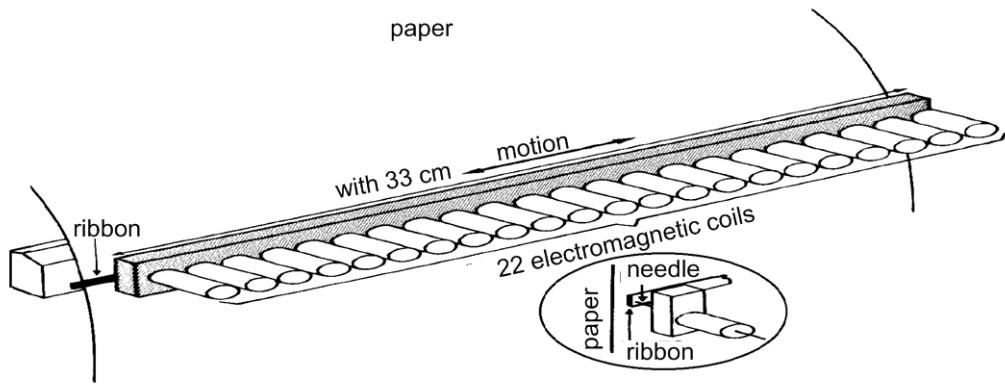


Figure 6.7. Needle line printer

6.3. Laser Printers

Laser printers offer the best quality of images, respectively, the highest resolution. They, as well as LED printers, are based on one and the same technology, used for the first time in copier machines. This process, known as **electrophotography** was invented in 1938 and was developed by Xerox and Canon at the end of the 1980s (Fig. 6.8).

The electrophotographic process comprises several major steps in laser printers.

The critical component of this process is a drum. The drum is an aluminium cylinder coated with photosensitive material – selenium. The photoconductor is uniformly charged with static electricity by means of a coronary discharge. An image is exposed on the so charged photoconductor by means of light which discharges it selectively and creates a hidden or invisible image.

The development in the printer takes place by transferring toner to the photoconductor. It sticks only to the charged areas and thus develops the hidden image into a visible one. Then the developed image is transferred to the paper electrostatically. By pressure and heating, the toner is baked and fixed to the paper. The last step is cleaning the photoconductor from the residual toner and electrostatic charge and thus the drum is prepared for the next cycle. [45]

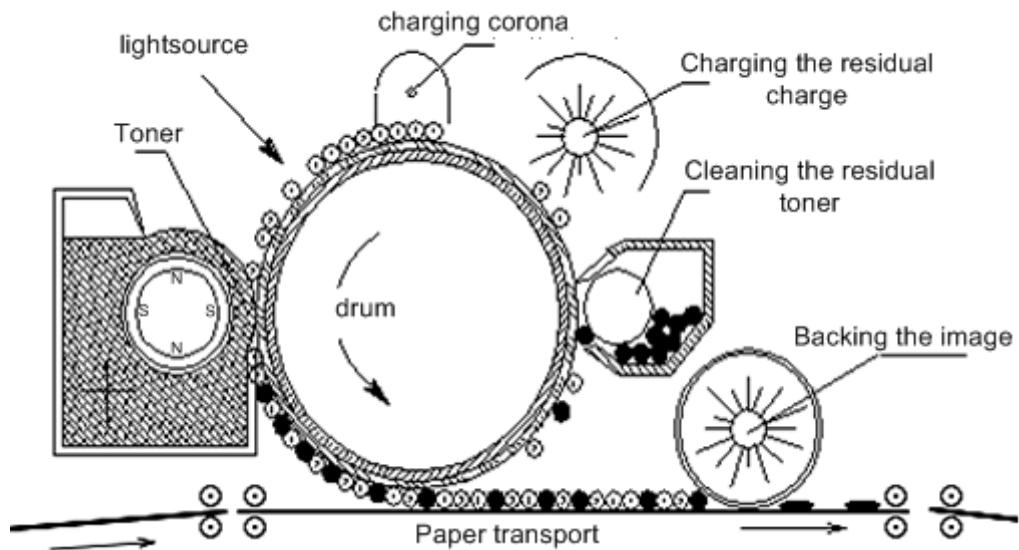


Figure 6.8. Electrophotographic process

In laser printers the hidden image is created dot by dot by the scanning module (Fig. 6.9). The light beam scans the drum surface parallel to its axis by means of a polygonal mirror prism. The drum and prism rotate continuously and synchronously. Thus the scanning of the complete surface is ensured. The image is formed similarly to the image in the cathode ray tube by using a light beam instead of an electron beam (Fig. 6.10.).

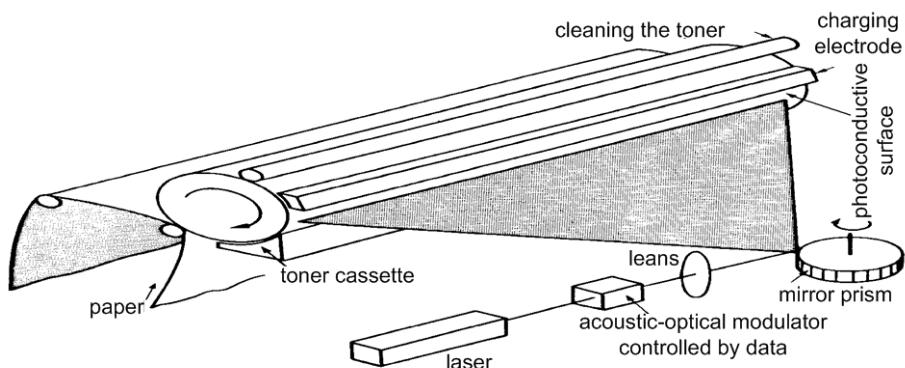


Figure 6.9. Scanning module of a laser printer



Figure 6.10. Image formation

The beam is turned on and off by means of an acoustic-optical modulator placed between the laser and mirror prism. Usually a low-digit matrix (18×24 dots) is used in which the dots partially overlap so that a high quality image is achieved.

If the laser beam is divided into several beams, several lines of the image will be formed simultaneously which will increase the printing speed.

Block diagram of a laser printer is shown in Figure 6.11. [11] The data to be printed enter by the interface for connecting to the computer. The central processor retrieves from the font memory the dot matrices of the characters and saves them in the dot memory from where they are passed to the acoustic-optical modulator. To avoid distortions in the produced image, devices for maintaining constant speed of rotation of the mirror prism and drum are incorporated. The heater serves to fix the toner to the paper before the printed page is released from the printer.

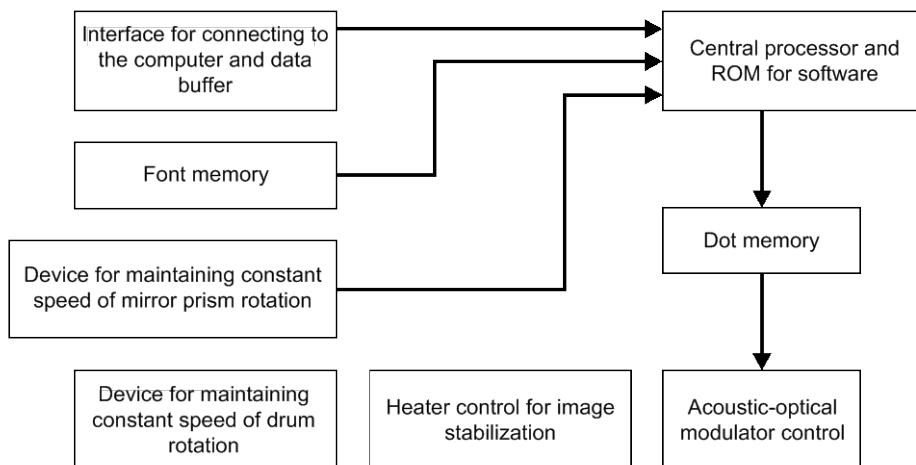


Figure 6.11. Simplified block diagram of a laser printer

Initially, laser printers were designed to operate at very high speed (of the order of 100 pages per minute) and to print a large amount of information. The price of such printer is high and its usage is justified only in very large systems. Later less expensive laser printers were developed, designed for meeting more moderate requirements. They operate at a speed of 10÷20 pages per minute. Their mechanisms have limited lifetime and they are not suitable for prolonged printing. A cheap laser prints 3000 pages per month, i.e. it works 10 minutes a day. The drum surface can be coated by a cheap non-toxic organic photoconductive material instead of selenium. Such a drum prints 3000÷5000 pages while if selenium is used, 50÷100 thousand pages are printed. Usually, but it is not obligatory, drums with organic coating are replaced every time when the toner-cassette is fed.

Laser and LED printers are very suitable for applications which require high-quality and fast printing of images and texts (e.g. in publisher's systems)

6.4. Inkjet Printers

In the 1970s **IBM** developed a **continuous** inkjet technology. It is based on the principle of controlling a continuous stream of ink drops to the image carrier by applying an electrostatic field to previously charged ink droplets (Fig. 6.12).

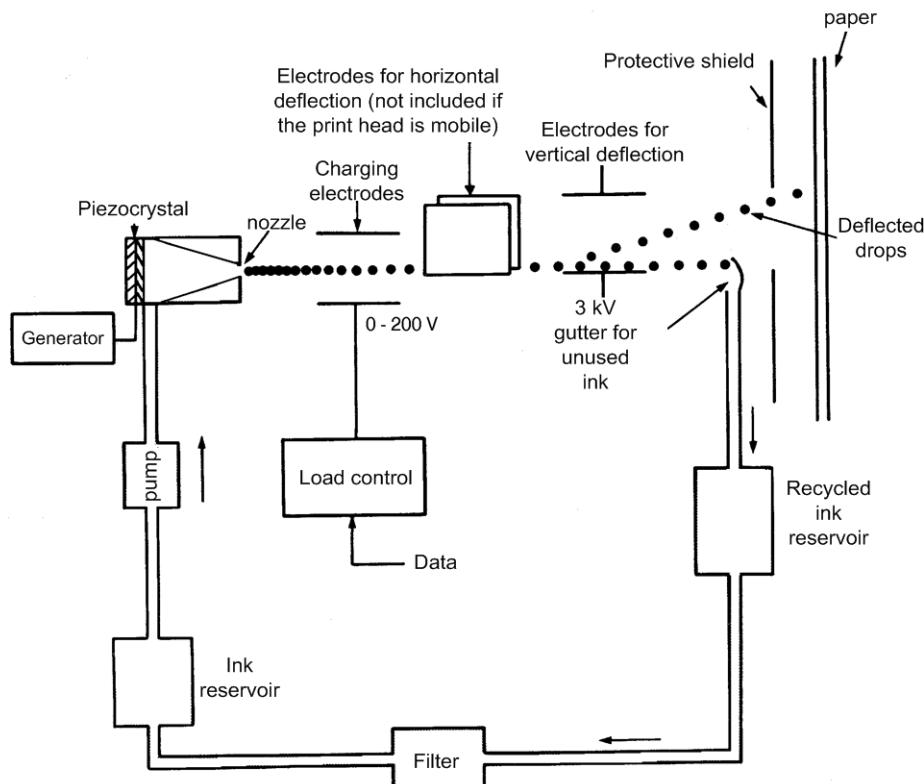


Figure 6.12. Inkjet printer

Electroconductive ink is forced through a very small nozzle to produce a high speed stream or jet of drops of ink. By means of a piezo crystal vibrating at ultrasonic frequency, the size and spacing of these drop is made constant. The vibrating frequency is around 100kHz, the drop diameter is typically 0,06 mm, and the spacing 0,15 mm. Each drop of ink, after leaving the nozzle, is given a specific charge as it passes through charging electrodes located next to the nozzle. The drops are deflected by electrodes for vertical deflection and strike the paper. The horizontal position is normally controlled either by moving the system mechanically or by electrodes for horizontal deflection.

The angle of deflection of the drop is determined by the charge which it has received when passing through the charging electrodes. With no charge there is no deflection and these drops are collected in a gutter placed close to the paper. Increasing charge increases deflection so the voltage applied to the charging electrodes must be around 200V. The voltage of the deflecting electrodes remains constant – around 3kV.

Characters are formed in a dot matrix. The information about the characters is stored in digital form (in compressed form) and controls the voltage of the charging electrodes during printing. For high quality printing usually 10^3 drops are required per character. With 10^5 drops per second released from the nozzle, 100 characters can be printed per second. The first ink jet printer was manufactured in 1967. Some improvements were made later in its construction. For example, a system was developed which allows two successive drops to fall at two remote dots on the paper. Thus inaccuracies due to electrostatic repulsion between the drops are avoided.

The inkjet printers described so far are called continuous jet printers, since the ink flows continuously through the nozzle. The unused ink is collected in the gutter. In practice constructions are used in which drops are produced only when required and hence there is no need for a collecting gutter and deflecting system. In this case a piezoelectric element can also be employed but then the process becomes comparatively slow. To enhance speed other techniques of pushing out the ink are implemented.

In the 1980s ‘**drops on request**’ technology appeared in the manufacture of printers. Printers based on this technology shoot drops to the carrier only when required for printing. This method considerably reduced the complexity of equipment compared to the continuous ink flow technology. In these early printers the ink drops were shot by piezoelectric converters which supplied the necessary pressure by mechanical motion. (Fig. 6.13).

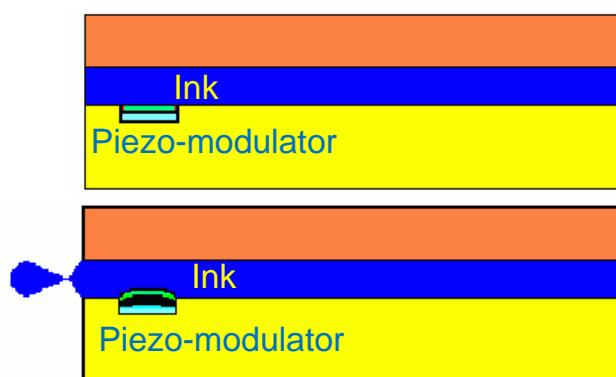


Figure. 6.13. Inkjet printer with a piezoelectric converter

At the same time **Canon** developed the “**bubble jet**” technology. By this method a small electroresistive heater heats the ink to boiling. A bubble is formed which expands and pushes out the ink through the nozzle. Soon after Canon, and independently of them, **Hewlett-Packard** developed a similar technology known as **thermo-inkjet** (Fig. 6.14).

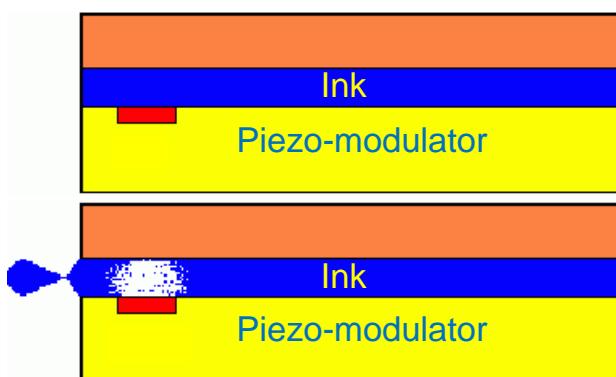


Figure. 6.14. Thermo-inkjet technology

The most popular inkjet printers use consecutive printing (Fig. 6.15). Similarly to matrix printers, serial inkjet printers employ print heads with

nozzles vertically arranged in one or more columns. The process of image synthesis is similar to the process in matrix printers. [45]

Similarly to matrix printers, inkjet models with linear printing are also manufactured to achieve higher print speed. The process of building the image is identical to the one in LED printers, as shown above in Figure 6.10.



Figure. 6.15. Serial printing

Principal advantages of inkjet printer are quiet operation, high quality colour printing and low cost. Their only disadvantage is expensive maintenance. Compared to laser printers, the cost of a page of inkjet copy is higher. There are exceptions but these are industrial printers designed for heavier service loads.

6.5. Colour Printers

In daisy-wheel and matrix printers colour printing is produced by using bands of stripes coloured in the three subtractive primary colours: magenta, cyan and yellow. By mixing equal amounts of two primary colours we can obtain red, green and dark-blue colours. Black can be obtained by mixing all primary colours or by a separately coloured black stripe on the band. So seven colours in all are obtained. In colour printing the primary colours are produced by absorbing the rest of the components of the spectrum, unlike cathode ray tubes where light is emitted and the additive primary colours are red, green and blue.

Usually colour stripes are positioned horizontally on the band. After printing one of the colours, the whole system moves so that the head is over the next colour and the line is repeated. It is possible, however, that the band consists of several differently coloured sections. Then the colour is changed by rewinding the band.

The colour obtained in laser printers is determined by the colour of the toner. Therefore, a coloured copy is produced by running the paper through the printer several times – each time with a different toner. To achieve this aim, a very precise positioning of the sheet is necessary. The size of the paper and other materials used in printing vary little but enough the affect the superpositioning of the image. This problem is overcome by means of special sensors built into the system which control the position of the sheet. The copy is positioned before each run through the printer.

In jet printers colour printing is produced by using three colours of ink: magenta, cyan and yellow. Each dot of the image consists of three different dots, coloured in the primary colours. The effect is similar to the one in colour TV sets where each frame is made up of triades of red, green and blue dots.

7. DEVICES FOR INPUT AND OUTPUT OF 3D INFORMATION

Design in modern mechanical engineering is performed by the so called CAD systems. They comprise not only software and computing environment but input-output devices for 3D graphic information. This sort of design often requires the input of *real 3-dimentional objects* in order to create models, to measure parameters or to inspect. This is done by means of devices called "**3D scanners**".

Devices for outputting 3D graphic information are a special class of machines for rapid prototyping, by means of which it is possible to build a conceptual model of a 3D CAD file within minutes or hours.

7.1. 3D Scanners

Scanning is a process of optical analyzing of two-dimensional images or three-dimensional objects and their digitalization in order to process the data in a computing system. Conventional scanning copies the coloured information from the image. The three-dimensional however, aims at taking the relief of the studied object. The surface of the body is traveled over by a sensor generating a sequence of points located in a virtual 3D raster grid. The coordinates of points in space are obtained by uniting (joining) the relative position of the sensor to its indications along three mutually perpendicular axes. The digitalized data from the scanning process can serve for creating CAD models or to be converted to a polygonal grid. This allows easy and fast manipulations as: editing, scaling, restoring, animating, etc. It facilitates the machining of prototypes and matrices. Digitalization of objects and parts with complex shapes allows measurements to be taken which are difficult to do on the real object, as well as accurate calculation of parameters of various items.

Three-dimensional scanning makes possible the exact comparison of real objects with arbitrary complicated shape with their CAD models. The purpose is to control and monitor the wear, corrosion and deformation of parts and machines in the course of time.

From user's point of view a *three-dimensional scanner* is each device which automatically adds three-dimensional coordinates from the surface of a body in a certain sequence.

To design a device for inputting a real 3D object with satisfactory accuracy and exporting the data to a format convenient to be used by the known CAD/CAM processing systems, it is necessary to solve the following problems:

- development of a kinematic model for 3D positioning system control;
- selection of drive elements and control algorithm;
- design of software interface for connecting to CAD environment

7.1.1. Development of a Kinematic Model for 3D Positioning System Control

To unify the kinematic model itself, it is necessary to approximate each part which is to be scanned to a three-dimensional object with a geometrical shape.

Three-dimensional bodies in Cartesian space R^3 are divided into two large groups:

- bodies formed by rotating a curve with arbitrary shape around one of the axes (rotational bodies);

- bodies formed by linear motion of a curve with arbitrary shape in succession along the three axes.

Employing the methods of describing these bodies, two approaches to scanning have been taken: rotational and planar.

In rotational scanning the rotation around Z axis is constant. The process of measurement takes place in mutually parallel planes, perpendicular to this axis (Fig. 7.1).

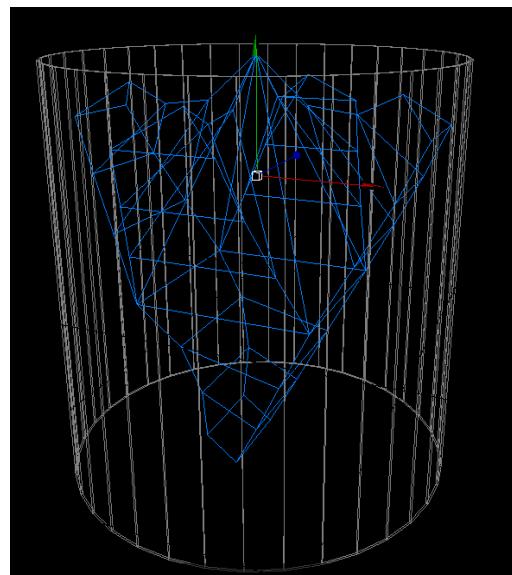


Figure 7.1. Rotational scanning

The visual shell of a three-dimensional body is the closest approximation, obtained by the method of section. Therefore, the visual shell contains all parts of the scanned object which can be inscribed in a cylinder with a certain diameter and height.

In planar scanning the object is scanned perpendicularly, side by side (Fig. 7.2).

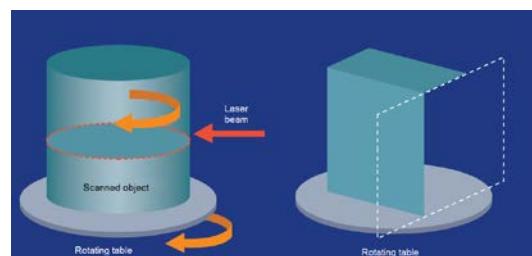


Figure 7.2. Methods of scanning

The model covering these primitives should possess at least 5 degrees of freedom: 3 linear and 2 rotational motions.

The described methods of scanning are static. The trajectory of the measuring tool is preset and cannot be altered between scannings.

Figure 7.3 shows a model of positioning system of portal type. The basic elements configuring this diagram, are numbered as positions.

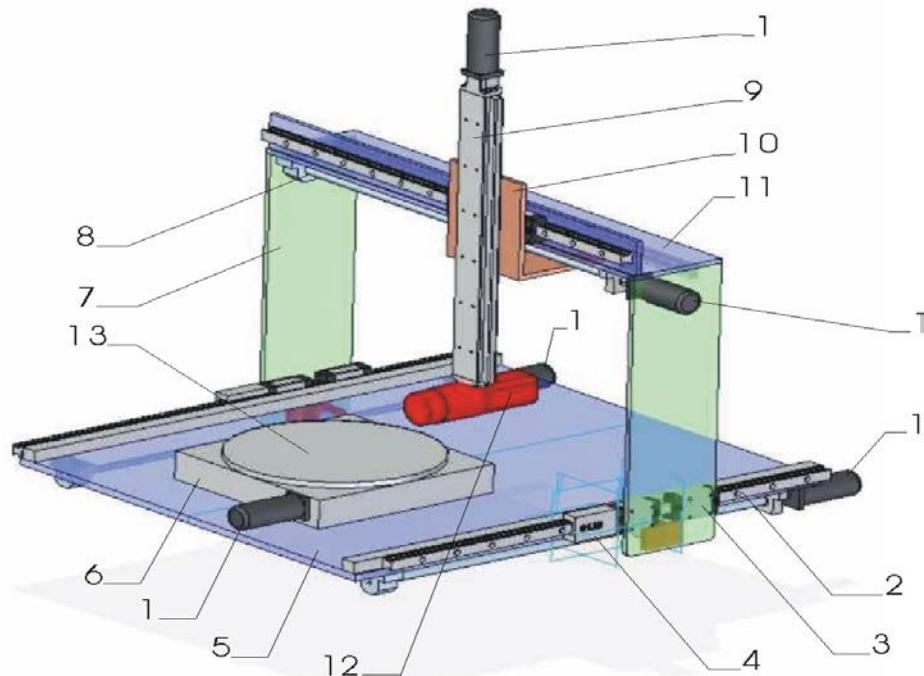


Figure 7.3. Positioning system of portal type

1. Servomotors for drives along the respective axes.
2. Guiding rail along X axis..
3. Runner blocks, providing the motion of the portal part along X axis.
4. Incremental reading head measuring the displacement along X axis.
5. Base.
6. Rotating table providing the rotation of the object to be scanned.
7. Vertical support.
8. Ball-screw pair providing the displacement along Y axis.
9. Ball-screw pair with a slide positioning the measuring head along Z axis.
10. Holder.
11. Horizontal support.
12. Measuring head (or measuring sensor);
13. Surface on which is placed the object to be scanned.

The measuring sensor can be contact. In this case, when it touches the object and when certain pressure is reached, the chosen coordinate point is measured. This approach is known in measuring devices which sense the part roughly and measure its surfaces. When it is necessary to measure the graphic object more quickly, non-contact measuring techniques are used. Technologically, ultrasonic and light converters are applied most frequently. The first type of converters is considerably cheaper but less accurate and focusing on small non-planar areas is more difficult as a method of measurement. Light converters use laser pulse modulation which is focused as a spot with a diameter much smaller than one micrometer. The methods of

measuring distance are triangulation and counting the time for the signal propagation from the emitter to the surface and back (this time is directly proportional to the traveled distance).

In 3D scanners a laser measuring head with the following parameters is usually employed:

- operating distance 140 mm;
- operating tolerance 50 mm;
- resolution 0,005 mm;
- accuracy 0,025 mm;
- interface for connection – specialized PCI controller;
- compact dimensions;
- minimum weight.

This type of laser heads operate on the principle of rotational triangulation which allows scanning of semi-transparent bodies, mirror surfaces and surfaces with steep slope (up to 60°).

The proposed kinematic diagram is comparatively light, without high inertia moments. The portal type provides stability of the system with respect to vibrations when measuring along Z axis.

7.1.2. Selection of Drive Elements and Control Algorithm

Selection of Drive Elements

Two types of motors are used as drive components: direct current and alternating current induction motors. The second type is preferred because it can be controlled by an ordinary frequency inverter. They have flexible characteristics, easy maintenance (in practice they are permanent) and conventional control. The block diagram of 3D scanning device control is shown in Figure 7.4.

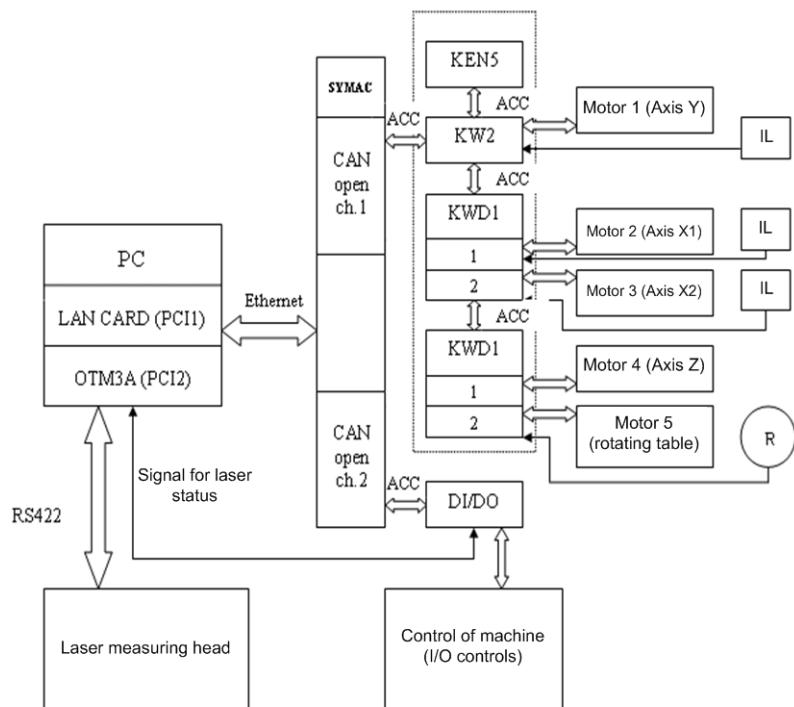


Figure 7.4. Block diagram of 3D scanning device control

The total number of motors is 5. Two of them are used for driving the portal part along X axis, one – along Y, one along Z and the fifth motor is for angular displacement of the rotating table on which the object to be scanned is placed. Each of them communicates with a KW module which ensures precision of control (high accuracy and dynamics). Power is supplied by KEN block.

A personal computer controls the device. Its functions are to acquire data received from scanning, to process information about the positions of the motors and to show the course of the process visually. The industrial multi-axial controller SYMAC communicates, on one hand, with the computing environment through LAN CARD (PCI controller), coupled at the computer slot using Ethernet interface, and on the other hand, with the motors by means of servo-inverters (KW modules). It has scalable computing power in the flexible system, by means of which it ensures the fast processes of PLC program and the processing of a large amount of data. SYMAC guarantees operation in real time with high precision of synchronization and maximum deflection of 1 μ s. It allows addressing external input-output modules connected to ACC (AMK CAN COMMUNICATION) BUS.

The position along X and Y coordinates is read by means of incremental lines (IL). The positioning angle of the rotating table is generated by a photo-raster transmitter (R). In the general configuration of the system limit switches are included, which define the physical format. They are read by the controller by means of the digital data input-output block (DI/DO).

The laser head is a precise measuring instrument which provides non-contact measurement of the scanned object. It is controlled by a program driver and communicates with the computing environment by PCI card. The communication channel is serial interface with differential exchange - RS 422.

Control Algorithm

System software (SS) consists of 2 organizational blocks (OB) and 16 functional blocks (FB). It comprises three basic operating modes of the scanner. (Fig.7.5).

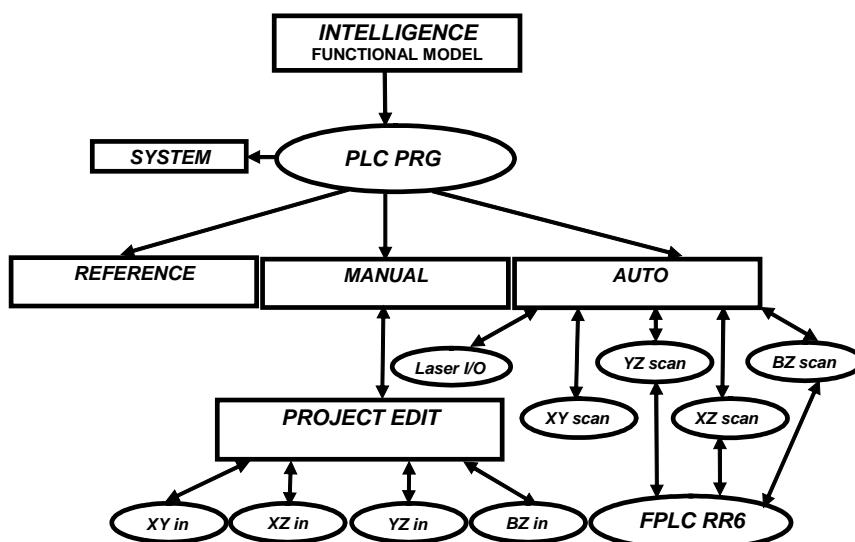


Figure 7.5. Functional model of 3D scanning device

- **FB: Referencing.** It is always activated when the system is switched on. Before performing other procedures it is necessary to orientate the scanner (to find the origin of the coordinates) since all the other distances are measured as regards these coordinates. If it is not referenced, operation in another mode is not possible.

- **FB: Settings.** It is activated for creating or editing projects for the scanned objects. Manual move along all axes is possible here, as well as visualization of the actual position and the distance measured by the laser head. When a new project is started, the following surfaces for scanning are used: XY, XZ, YZ, BZ (B is axis of rotation). Each object subject to scanning is described by n in number surfaces which account for its complexity. The parameters set for each new surface are: initial point, final point, measurement range and scanning step.

- **FB: Auto.** It scans objects automatically by previously designed projects. It is started with the first surface written in the project, the initial and final points are controlled, the axes from the scanned surface are cyclically moved by a specified step. The measuring head range is controlled and the distance towards the object to be scanned is adjusted. Axes positioning and laser head control is carried out by organization block synchronously satisfiable which determines execution time of 1 ms. The operating mode is consecutive and it is divided into two clock cycles. The first clock cycle is positioning to the next point with a set step. The second clock cycle is laser head control: start of measurement, validity of measured data and recording of coordinates. After scanning the first surface, a buffer is loaded with the recorded coordinates along the three axes participating in scanning and it is sent to the personal computer. Then the following surface written in the project is scanned, if there is one, if there is none, scanning is over.

Organization Blocks of SS

- After starting the controller, the first organization block (**PLC_PRG**) is executed. The three modes of operation (FB:Referencing, FB:Settings, FB:Auto) are selected and called here. The block monitors and controls the state of the servo-system and laser head (FB:System).

- The second organization block (**FPLC_PRG**) is cyclically satisfiable for a fixed time (1ms). It performs major control functions when operating in automatic mode: position control of servo-axes, laser head control and scanning by a set surface (FB:XY_scan, FB:XZ_scan, FB:YZ_scan, FB:BZ_scan).

Functional Blocks of SS

- *System.* It performs the function of monitoring and diagnostics of the servo-system and laser head. It visualizes the errors and monitors the limit switches. It switches on and off the ready state of the servo-system and laser head.

- *Referencing.* It performs the first basic mode - scanner referencing. Manual or automatic referencing is possible. In manual referencing only one axis is orientated, and in automatic – all axes in sequence: Z, X, Y and B,

where B is the axis of the rotating table which rotates the object to be scanned.

- *Settings.* This block performs the second basic mode. The functional blocks for manual move along axes (FB:Manual_Move) are called by it, the values measured by the laser head are visualized (FB:Laser_IO) and projects are edited (FB:Project_Edit).
- *Auto.* It performs functions concerning automatic scanning control. It starts and stops scanning by a set project. The basic functions of the automatic mode are fulfilled by FPLC_PRG.
- *Manual_Move.* It performs functions concerning manual move of axes in „+“ or „-“ direction. Buttons are used for control.
- *Project>Edit.* It performs functions concerning adding or editing surfaces for scanning the object (geometric decomposition).
- *XY_in.* It inputs initial and final coordinates for scanning on XY surface; it inputs range and step of measurement.
- *XZ_in, YZ_in, BZ_in.* It performs the same functions as *XY_in* for XZ, YZ and BZ surfaces.
- *Laser_IO.* It performs functions concerning visualization of the values measured by the laser head – in settings mode.
- *XY_scan.* It performs functions concerning monitoring and positioning axes in automatic scanning mode on XY surface. The scanning process takes place in three cycles:
 - axes positioning and adjustment of scanning range;
 - starting and synchronization of measurement by the laser head;
 - recording of coordinates.
- *XZ_scan, YZ_scan, BZ_scan.* It fulfills the same functions as *XY_scan* for XZ, YZ and BZ surfaces.

7.1.3. Design of Program Interface for Connection with CAD Environment

The program interface is designed in the following sequence:

- the scanner reads 3D points;
- figures of a set of points are built;
- discrete data about the figure are created;
- the object is exported in a popular CAD format.

Reading 3D Points by the Scanner

The information about the points to be scanned is read by program interface or output file. The program interface is described by means of a counter. The counter has two functions: Get Points Count (as a result returns the number of scanned points) and Get Point (as a result returns the structure containing the three coordinates x, y, z). This function receives as input parameter index $i \in (0; n)$, where $n = \text{Get Points Count}$. The output file contains information which is served by the program interface. The difference is that the file can be portable, even if 3D scanner is switched off or if it has lost the information about scanning. The data format in the output file is:

x_1, y_1, z_1
 x_2, y_2, z_2
 .
 .
 $x_n, y_n, z_n.$

In each line there are three coordinates of points in space separated by a spacing. The number of point is determined by the number of lines in the text output file.

Building a Figure out of a Set of Points.

Building a figure out of a set of points is an open problem in computing geometry because it is not possible to obtain an optimum surface on the basis of input data. Various algorithms are available to approximate the figure to a certain percentage. As input data, the algorithm employs a set of points in space, and as a result of its performance, it generates a set of triangles in space. One of the methods for calculating the figure comprises Voronoi diagram, Delaunay triangulation and the shell of the figure (finding a convex polyhedron).

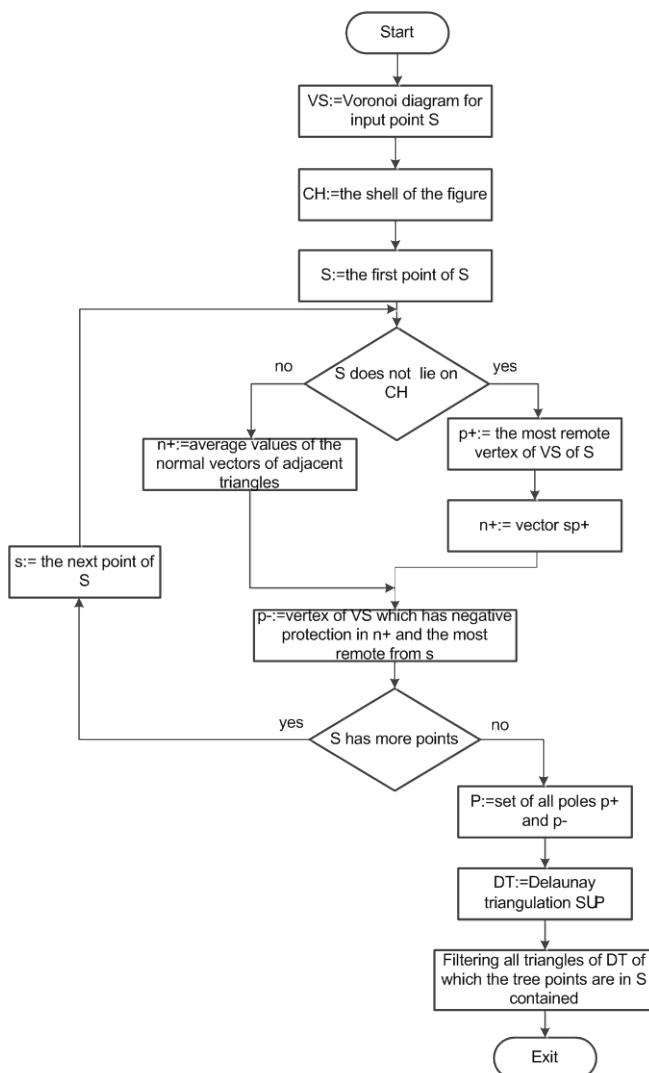


Figure 7.6. Algorithm for building a figure of points

Voronoi diagram

Voronoi diagram, also referred to as Voronoi tessellation or Voronoi decomposition is a special type of decomposition of metric space represented as distances of a certain discrete set of objects in space – e.g. a discrete set of points. In the simplest and most general case, in the plane for a given set of points S , Voronoi diagram for S is a separation of planes which associate area V_P at each point p from S in such a way that all points in V_P to be closer to p than to any other point from S .

Delaunay triangulation

Delaunay triangulation for a set of points P in the plane is a triangulation $DT(P)$, which does not contain points from P , internal for the circle circumscribed around each triangle in $DT(P)$. The triangulation increases the minimum angle of all triangles circumscribing the figure.

A shell (envelope) of a figure

In mathematics the convex shell (envelope) of a set of points or the convex boundary of an area for a set of points X in a true vector space V , is such a set of points X' , which contains minimum number of points enveloping the set X .

The algorithm for building a figure out of points is illustrated in Fig. 7.6.

Creating Discrete Data about the Figure

The algorithm for building the figure employs a set of points as an input and it generates a set of triangles as an output. Two approaches to output data storage are described below (discrete data about the figure).

Approach 1

The structure of triangles is stored and each of them carries information about three points. A shortcoming of the approach is that data about one point can be repeated because the point can be found in a number of triangles.

Approach 2

A table of points is stored. Each point from the table is not repeated, and it is identified by an index. Each triangle defined by three points, refers to definite indices in the table.

The second approach to data storage is more complicated to carry out but it is optimal in relation to storage space used.

Object Export in Popular CAD format

Data export in popular CAD format is represented by a simplified algorithm with the following steps:

- formation of data file flow;
- saving the data flow about the heading block of data for the exported format;
- scanning all triangles and their conversion into a respective CAD format;
- saving a final block of CAD format in the output flow;
- closing the output flow;
- exit.

The algorithm for export to CAD format is illustrated in Fig. 7.7.

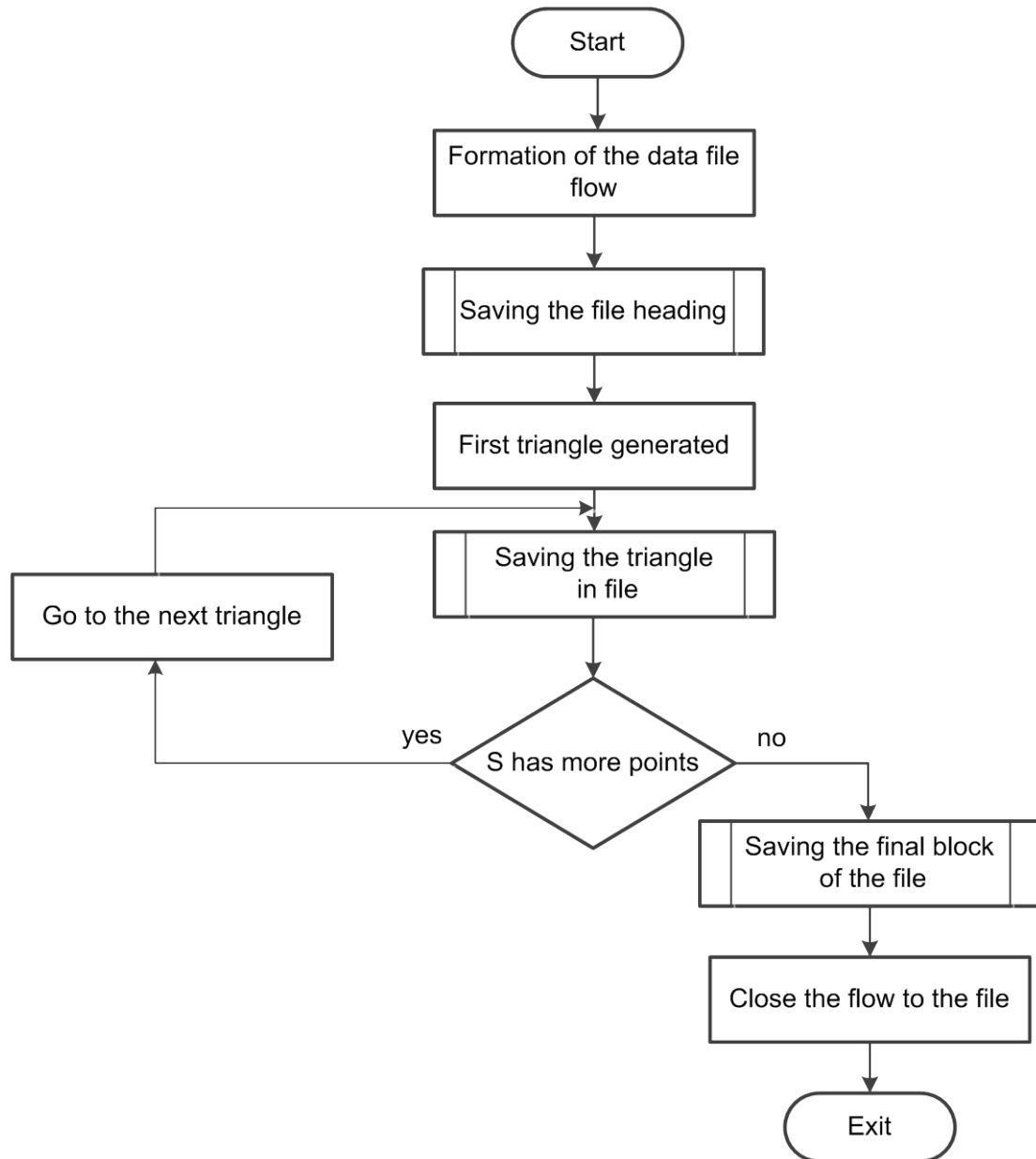


Figure 7.7. Algorithm for data export in CAD format

The general view of a 3D scanner is shown in Figure 7.8. Each user can work with it using the so called user's interface (Fig. 7.9) It allows the initial coordinates to be determined, the mode and displacement speed to be selected, to describe the surfaces to be scanned and scanning to be started.

The proposed approach to the design of a 3D scanning device is for working with CAD systems. Accuracy is of the order of 0,025 mm. In various devices the physical format is determined by the maximum runs along axes X, Y and Z. Rotating motion of the measured element is provided. Minimum step of positioning is around 5 μ m on the average.

The file generated for the scanned object is an arranged sequence of 3D points. It is converted into one of the known **DXF** or **STL** formats to be read by the CAD system.

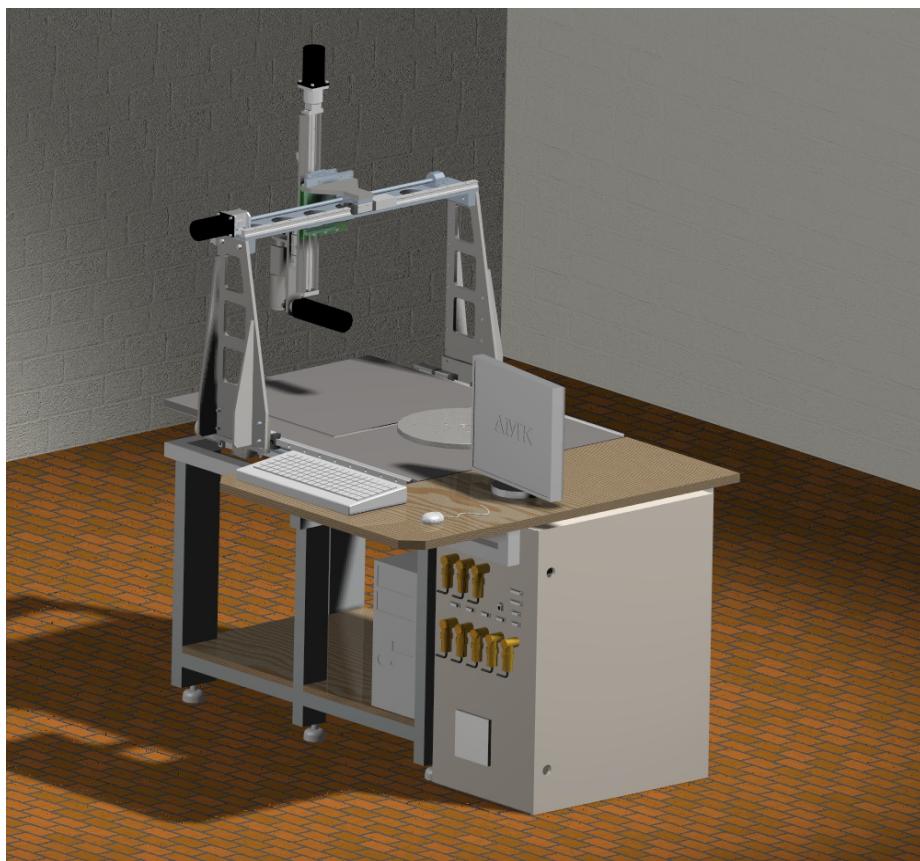


Figure 7.8. General view of 3D scanning device



Figure 7.9. User's interface for working with a 3D scanning device

7.2. Devices for 3D Information Output

The image in 3D devices for output is built by representing the object as a sequence of horizontal thin layers whose subsequent superposition creates a precise model of the 3D file (usually a STL file). These technologies became accessible in the late 1980s.

In the process of superpositioning the machine reads information from the CAD file and deposits layers of liquid, powder or sheets and in this way builds the model as a series of horizontal cuts. The layers are joined to form the final form of the prototype. Some techniques for „rapid prototyping” comprise two materials during the design. The first is the basic material building the object, and the second is the supporting material (when there are cavities and suspended structures in the model). Later the supporting material is removed.

A number of „rapid prototyping” technologies are available nowadays.

Selective Laser Sintering

This is a technique with additive rapid manufacturing which employs a powerful laser (CO_2) for sintering tiny particles of plastic, metal or ceramic powders into a mass which is the desired 3D object. The laser sinters the powders selectively by scanning horizontal cuts of a 3D CAD file on a surface covered with the powder. After the current working surface has already been scanned, the next layer of powder is deposited on it which is scanned by a laser and this process is repeated until the object is completed. Compared to the other „rapid prototyping” methods, this approach uses a wide range of accessible materials, e.g. polymers (polystyrene), metals (steel, titanium and various alloys), green sand (a mixture of sand, bentonite clay, pulverized coal and water). The process takes place by complete melting, partial melting or liquid sintering. Depending on the material, up to 100% of density can be reached with properties close to the ones reached by conventional methods of treatment (milling, casting, etc.). This technology was invented and patented by Dr. Carl Decard at Texas University in Austin.

Thermo-laying of Material

This technology is based on consecutive laying of material, layer after layer, until the process is complete. Plastic fiber or metal wire is unwound and thus material is supplied to a nozzle which is turned on and off when necessary. The nozzle is heated until the material melts and it can be moved both horizontally (X, Y) and vertically by a CNC (Computer Numerical Control) controller, controlled by CAD software packet (Fig. 7.10).

In a way, similar to stereolithography, the model is formed by layers. Plastics set after extrusion from the nozzle. Several types of materials are accessible with different compromises between endurance and temperature, e.g. polycarbonate, polycaprolactone, polyphenolsulphone and wax. Water-soluble materials can be used in cavities and suspended structures in the part, if necessary.

One of the deficiencies of this technology is the comparatively low discreteness of items compared to the output file.

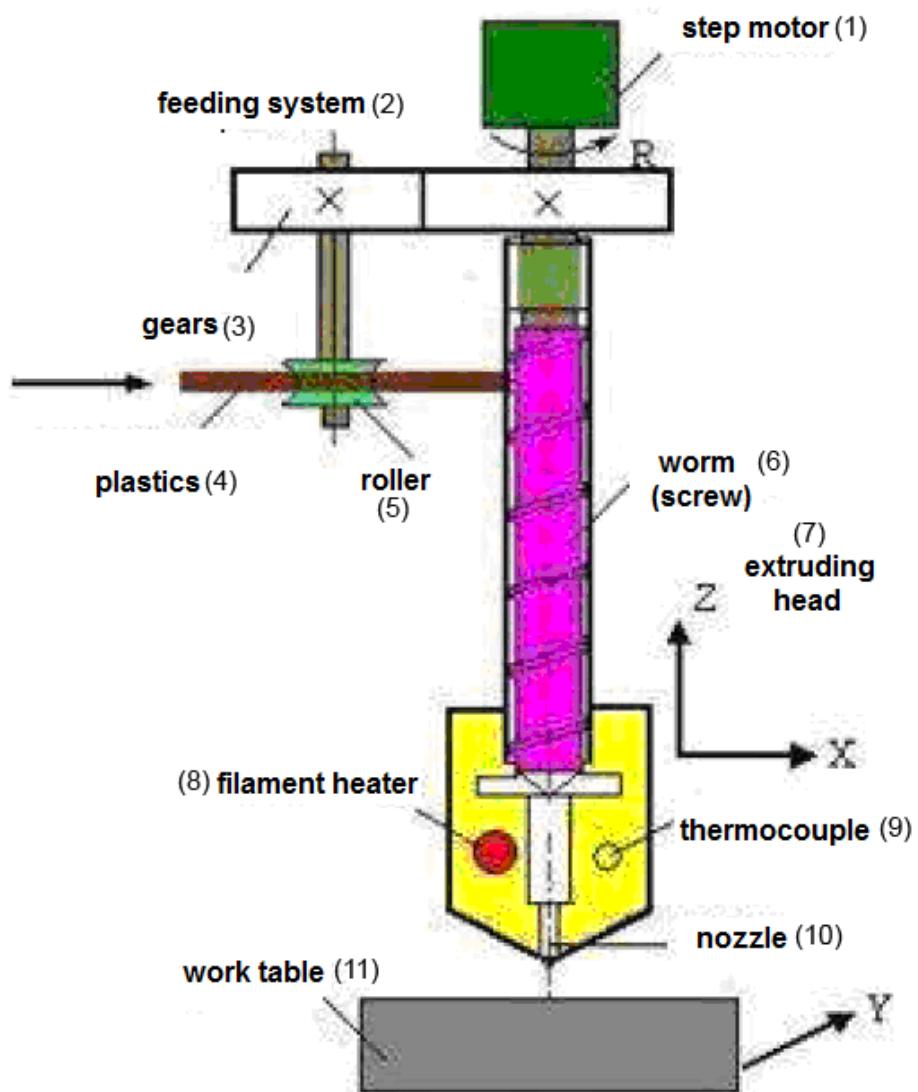


Figure 7.10. Building a 3D prototype by thermo-laying of material

Stereolithography

Stereolithography is a method which also builds items layer after layer (Fig. 7.11). In the process of building a container of liquid active photopolymer resin is used and a laser by means of which the layers are built. For each layer the laser traces the raster (from the horizontal section of the part) on the surface of the liquefied resin. Under the effect of the laser, the illuminated path of the resin cures and sinters with the layer underneath. After curing the worktable lowers by one thickness and the next layer is applied. This process is repeated until the part is completed. Finally the part is cleaned by dipping in a chemical bath and baking. Stereolithographic technology requires the usage of supporting structures to fix the parts on the work-plot and to prevent transposition of parts caused by gravitation and the tool for applying the new layer. Supports are generated automatically with creating the CAD file but they can be altered manually for convenience. After finishing the ready product, these supports are removed. Objects produced by means

of stereolithography typically have greater discreteness and accuracy compared to objects produced by the method of thermo-laying.

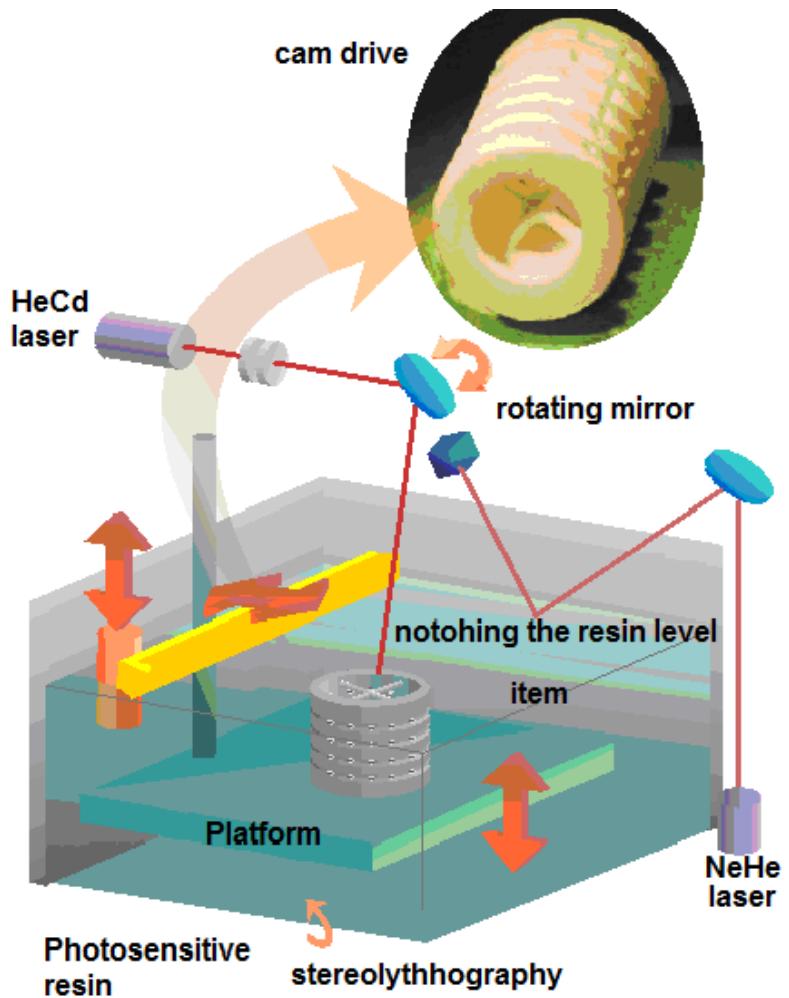


Figure 7.11. An example of construction of a machine working by stereolithography technology

Melting by an Electron Beam

This technique uses metal powder as material. Prototypes made in this manner are entirely compact and have the properties of the material used. The device reads information from CAD file and lays layers of metal powder which subsequently sinters by a computer controlled electron beam. The material to be melted is of pure alloy in powder form. Therefore the prototype does not need further heat treatment to obtain its mechanical properties. The minimum layer thickness is 0,05mm, which allows very high accuracy to be achieved when making and storing the parts from CAD file. Fig. 7.12 shows examples of finished products.



Figure 7.12. Examples of finished products made by melting by an electron beam technology

3D Printing

3D printing is a method of converting a virtual 3D model into a physical one. 3D printing is also a kind of prototyping and as such it shares the common characteristics of the methods mentioned above. The model is formed by consecutive depositing of layers one on top of the other, until a finished product is obtained. The method of 3D printing usually is more accessible and faster than the remaining technologies. One of the variants of 3D printing is printing by injecting. The raw material (usually in the form of powder) is laid on the last layer and the printer head passes and injects glue and thus the material is soldered. Then a new layer of powder is laid according to the sections obtained from the 3D model and so on, until the physical model is finished. This is the only technology which allows full-colour printing.

Another method of printing is by employing a photopolymer as a raw material which subsequently is hardened by an ultraviolet lamp built in the printer head. A deficiency of 3D printing is the low accuracy (...) of models. Resolution along Z is 0,1mm, and along XY-plane it is commensurable with the resolution of laser printers (450dpi). Powder particles have a diameter of approximately 0,05 to 0,1mm.

3D printing is a summing up of the basic methods. It is optimized by speed, low cost, easy service and it does not need a lot of post-processing.

REFERENCES

1. Valchev, T., Hr. Hristov Input-output Devices for Computer Graphics S., Tehnika, 1987. (in Bulgarian)
2. Institute of Mechatronics - Gabrovo. 413.044.007 ПО-13 Roll Plotters 297×420 Micronics ПР 297. Operator's Manual, Gabrovo, 1987. (in Bulgarian)
3. Karakehayov, Z., S. Grigorov Single-chip Microcomputers S., Tehnika, 1992. (in Bulgarian)
4. Atanasov, A. Application of Charge Transfer Structures S., Tehnika, 1988. (in Bulgarian)
5. Vittih, V. Et all. Image Processing in Automated Systems for Scientific Studies M., Nauka, 1982. (in Russian)
6. Gordan, I., M. Luka Machine Graphics and Automation of Design M., Mir, 1987. (in Russian)
7. Gruver, M., E. Zimmers Cad Systems and Automation of Manufacturing M., Mir, 1987. (in Russian)
8. Zhuk, M. et all. Computer Hardware and Operating Systems M., Vissaya shkola , 1986. (in Russian)
9. Ivanov, E. L., I. M. Stepanov, K. S. Homiakov Computer Peripherals and Systems M., Vissaya shkola, 1987. (in Russian)
10. Myachev, A. A. Reference Guide – Interface Devices in Computers M., Radio i svyaz , 1993. (in Russian)
11. HELP series Laser Printer – this is very simple!! M., 1992. (in Russian)
12. Tompkins, W., J. Webster Interfacing Sensors to the IBM PC Prentice-Hall, Inc., 1988.
13. Reference Book on Automated Design Systems in Radioelectronics Edited by Prof. Norenkov, I. P., M., Radio I swyaz, 1986. (in Russian)
14. Tormishev, Y. Hardware for Machine Graphics M., Nauka I tehnika, 1987. (in Russian)
15. Foley, J., A. van Dam Fundamentals of Interactive Computer Graphics Addison-Wesley Publishing Company, Inc., 1982.
16. Horowitz, P., W. Hill The Art of Electronics Cambridge, CAMBRIDGE UNIVERSITY PRESS, 1980.
17. Baron, W., et all. Development of a High-Performance, Low-Mass, Low-Inertia Plotting Technology, Hewlett Packard Journal, october 1981.
18. Wilkinson, B., D. Horrocks Computer Peripherals second edition, HODDER AND STOUGHTON, 1987.
19. Hnatek, E. R. A User's Handbook of D/A and A/D Converters John Wiley & Sons, 1976.
20. Saveta, N. N. Computer Peripherals M., Mashinostroenie, 1987. (in Russian)
21. Ivanov, R., O. Asenov Architecture and System Programming for IBM XT/AT Габрово, 1993. (in Bulgarian)
22. Mladenov, M. K., N. Kirov Peripheral Devices for Personal Computers S., Tehnika, 1987. (in Bulgarian)
23. Angelov, A., P. Petrov Microprocessors in Radio-technical Systems S., Tehnika, 1989. (in Bulgarian)

24. Ilarionov, R., R. Ivanov An Approach to Designing Large Displays for Alphanumeric Information, Collection of Papers from the Fourth International Scientific Conference "Electronic Engineering '95" organized on the occasion of the 50th anniversary of the Technical University of Sofia - 27 - 29 .09. 1995, Sozopol. (in Bulgarian)
25. Hristov, P. I., H. G. Pandauzov, P. N. Nikolov Methods of Position System Control of Plotters Journal of "Electropromishlenost i Priborostroene" 8 1986. (in Bulgarian)
26. Ivanov, R., G. Mihov Electronic Digital Devices and Systems S., Tehnika, 1990. (in Bulgarian)
27. Ilarionov, R. Et all. 3D Object Input into Computing Environment. Collection of Papers from International Scientific Conference "Uniteh 07", Volume I, 23-24 November 2007, Gabrovo. (in Bulgarian)
28. DXF Reference Guide, 2001 Autodesk, Inc.
29. http://en.wikipedia.org/wiki/STL_file_format
30. <http://www.ennex.com/~fabbers/StL.asp>
31. Craig Peacock, EPP, <http://www.beyondlogic.org/epp/epp.htm>, 2007.
32. Craig Peacock, ECP, <http://www.beyondlogic.org/ecp/ecp.htm>, 2007.
33. Winn L. Rosch, Winn L. Rosch Hardware Bible, Sixth Edition, 2003, Que Publishing.
34. <http://en.wikipedia.org/wiki/Irda>
35. <http://comexgroup.com/comm/bluetooth.htm>
36. http://comexgroup.com/video/displays/_lcd_ring_2003.htm
37. <http://bg.wikipedia.orgg/wiki/%D0%9C%D0%BE%D0%BD%D0%B8%D1%82%D0%BE%D1%80>
38. http://en.wikipedia.org/wiki/Video_projector
39. <http://www.christiedigital.com/AMEN/technologyExplained/theProjector/projectorAnatomy/TheNervousSystem.htm>
40. http://en.wikipedia.org/wiki/Lazer_mouse#Laser_mice
41. <http://www.chuckgroot.com/1Introductionandruleofthirds.html>
42. <http://bg.wikipedia.org/wiki/CD>
43. <http://bg.wikipedia.org/wiki/dVD>
44. <http://www.dvd-burner.ca/DVD-Tutorials/Understanding-Dual-Layer-DVD-Recording.html>
45. <http://mimech.com/> Printers technology from Micromechanic.htm

Table of Contents

Preface to the Second Edition	3
Preface to the First Edition	4
1. CAD/CAM SYSTEMS AND THE CONCEPT OF INTERACTIVE COMPUTER GRAPHICS.	
TOOLS. TOOL REQUIREMENTS	5
1.1. CAD/CAM Systems and the Concept of Interactive Computer Graphics.....	5
1.2. Technical Tools of CAD Systems	7
1.3. Requirements for Graphic Peripherals	9
2. INTERFACES	15
2.1. Parallel Interface	15
2.2. Centronics Interface.....	16
2.2.1. IEEE 1284 Standard for Parallel Interface (Centronics)	19
2.3. Interface IEEE 488	32
2.4. Serial Interface	34
2.5. Standards for Serial Transmission (RS232, V24, RS423)	37
2.6. Current Loop Interface	38
2.7. Standards for Differential Input and Output (RS-422 and RS-485).....	39
2.8. Universal Serial Bus (USB).....	40
2.9. Interface IEEE 1394.....	43
2.10. Infrared Data Association (IrDA)	46
2.11. Bluetooth	52
3. DISPLAYS	59
3.1. Classification of Displays	59
3.2. Video block	61
3.3. Colour video monitor. The principle of producing colour images. Colour CRT.Block diagram of colour monitor	70
3.4. Vector displays	75
3.5. Graphic terminals and stations	76
3.6. Displays based on flat panels	79
3.7. Other display technologies	82
3.8. Advertising displays	83
3.9. Multimedia projectors	89
4. INFORMATION INPUT DEVICES	94
4.1. Interactive Manipulators	94
4.2. Digitizers	104
4.2.1. Model of an Inductive Digitizer	106
4.2.2. Formal Model of Digitizer Functions	114
4.3. Scanners and Digital Cameras	118
4.4. Bar Codes	127
4.5. Optical Carriers (CD)	130
5. PLOTTERS.....	135
5.1. Classification of Plotters	135
5.2. Vector Plotters	136
5.2.1. Model of an Inductive Digitizer	136
5.2.2. Model of an Inductive Digitizer	137
5.2.3. Model of an Inductive Digitizer	141
5.3. Raster Plotters	160

6. PRINTERS	163
6.1. Daisy-wheel and Matrix Printers	163
6.2. Line printers	165
6.3. Laser Printers	168
6.4. Inkjet Printers	170
6.5. Colour Printers	173
7. DEVICES FOR INPUT AND OUTPUT OF 3D INFORMATION	174
7.1. 3D Scanners	174
7.1.1. Development of a Kinematic Model for 3D Positioning System Control	174
7.1.2. Selection of Drive Elements and Control Algorithm	177
7.1.3. Design of Program Interface for Connection with CAD Environment	180
7.2. Devices for 3D Information Output.....	185
REFERENCES	189
TABLE OF CONTENTS	191