VIDEC: A universal visual input for digital electronic computers

By D. Muszka and A. Sára

In the investigations concerning pattern recognition, the adoption of electronic computers recently became more and more frequent. For a long time it had been impossible because of the absence of suitable technical aids. Though lateron there came some visual readin instruments connectable with digital electronic computers into being, the most of them, however, were special purpose machines which, according to the demand of the commercial data processing, served in general just as reading-in vouchers, for the identifications of alphanumerical signs.

Another, but much smaller group of visual readin instruments has already been useful in scientific research — primarily in medical technics —, as they have proved to be suitable for analysing more complicated visual information sets. These instruments, however, have never been able to come into general use because of their elaborateness and, so, their high prices.

Among these instruments let us mention just for an example the FIDAC instrument, well-known in England, built by Ledley and his collaboraters, in which the film to be processed is resolved by a cathode ray tube (as light source), controlled by a digital sweep circuit, and the greyness degree of its picture-elements is forwarded in an appropriately encoded form to the computer.

In Hungary, on the request of medical research institutes, the Cybernetical Laboratory is continuing researches to develop a suitable technical appliance. The primary goal of this work consists in developing a relatively not expensive visual readin instrument which, connected to any digital electronic computer, is capable of processing both transilluminatable and light-reflecting pictures. In the following, I should like to give a review on our so-scalled VIDEC instrument, which constitutes the produce of the first stage of our research work.

The pictures to be processed (e.g. *X*-ray photographs, diagrams drawn by EEG, ECG registering devices etc.) can be fixed on the superficies of a cylinder made of glass. Taking into account that most of the pictures to be processed are such that the greyness degree of the picture-elements of the visual patterns on the pictures does not carry any essential new information (for instance this is the case of an ECG diagram) we decided to consider the elements, under an appropriate resolution of the picture, as dark and light points.

Along the generating line of the cylinder there are placed 5—8 point-like sources of white light. Near the inner and outer surface of the picture cylinder, according to the position of the sources of light, there are built in 5—8 diodes of appropriate resolving power. (In case of a transilluminatable picture the inner, in case of a light-reflecting picture, the outer diodes are activated, respectively.) The photodiodes and the sources of light are placed along the generating line proportionally in the sense that if the size of the picture to be processed, taken in the direction of the generating line, is "*b*" and we have "*n*" photodiodes then the distances between the photodiodes equal "*b*/*n*". (This distance can be changed on the instrument, and before starting the visual input one has to take care of its appropriate adjustment.) The number of the photodiodes is determined by how many channels of the computer

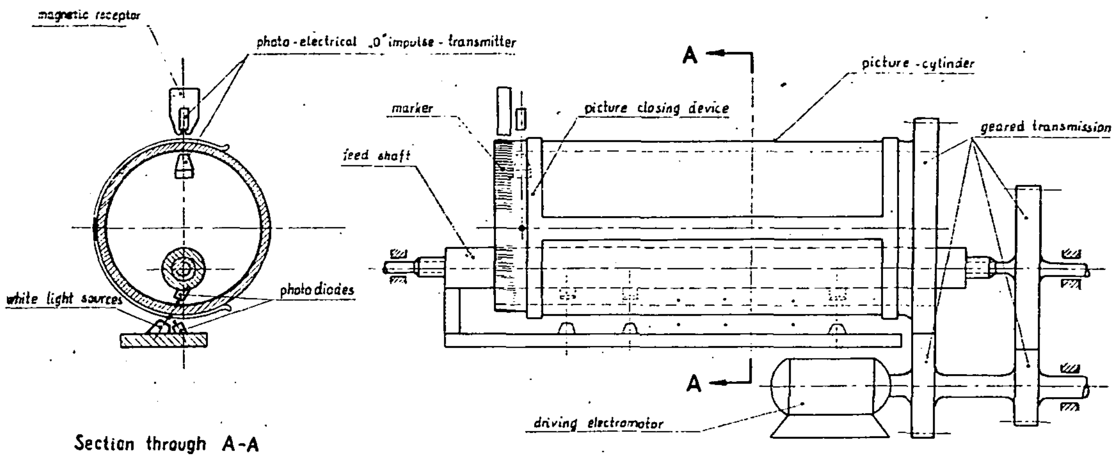


Fig. 1

the instrument is connected to are ready to receive information simultaneously. In case of our instrument, "*n*" cannot exceed 8. The length of generating line of the cylinder is 240 mm, so the size "*b*" of the picture to be processed can be 240 mm at most. The circumference of the picture cylinder is 270 mm, thus taking into account the size of the picture fixing device as well, the size "*a*" of the picture to be processed is 240 mm at most.

On the surface of the picture cylinder, outside the field of picture, there are magnetic Marker signs of density corresponding to the finest resolution applicable. The attainable finest resolution (raster) is 0,5 mm × 0,5 mm, so the number of the Markers is 480. The Markers cannot be found all along the circumference, just on the section of 240 mm. The device fixing the edges of the picture is placed inside the remaining stripe of width 30 mm, on the picture cylinder.

In the following, let "Δe" always denote the resolving fineness in both directions (the possible values of "Δe" are 0,5—1, 0—2,0 mm). The choice of the value "Δe" depends in part on the resolving fineness required by the task in question, in part on the operative memory capacity of the computer we are to work with. In fact, in case of the finest resolution (0,5 mm × 0,5 mm) and of the maximal size of picture (240 mm × 240 mm), 480 × 480 = 230.400 bits as information have to be stored which, in case of an average computer word length (30 bits), means $8K$ word memory capacities.

Here we should point out the fact that the proportional placement of photodiodes mentioned above implies that, after the input, the visual information is stored in the memory unit of the computer in a form unfit for direct processing. In fact, the bits corresponding to neighboring picture-elements are no longer "neighboring" in the memory. Therefore — with the aid of the VIDECORD ordering program[1] to be presented at the end of this talk — the visual information set stored in the memory before processing has to be rearranged into a form suitable for processing. This problem of rearrangement could be eliminated if the photodiodes were placed directly beside each other at a distance corresponding to the resolving finess. However, the construction of photo-perceptors of a sufficiently small size gives a grave technical problem as well as an additional expenditure. Therefore, the divided proportional placement of the photodiodes which made it possible to use any photo-perceptor and optics seemed to us more practical. Furthermore, the running time of the VIDECORD ordering program in most of the cases of the pattern recognition programs is much less than that of an effective pattern recognition program.

The operation unit of the instrument receives the starting impulse from the computer. Then the actuating engine of the instrument starts working and rotates the picture cylinder with 750 revolutions per minute. Thus, the time needed for the input of one picture varies between 5 and 8 second. (In case of a five channel and eight channel input system this is 8 and 5 seconds, respectively.)

The maximal speed of the information transport (in case of the finest resolution and 8 channels) is 54.000 bits per second. Just for comparison: the maximal information transport speed of a modern punched tape reader, RC-2000 of GIER Co., is 16.000 bits per second.

The actuating engine of the instrument, rotating the picture cylinder simultaneously, moves the photodiodes and sources of light with the aid of a suitable cogwheel transmission and pulling spindle in axial direction. The cog-wheel is chosen in such a way that during a full revolution of the picture cylinder the photodiodes and sources of light cover the distance "Δe". This cog-wheel transmission has, of course, to be retooled, according the choice of the resolving fineness.

The operation unit counts the number of revolutions of the picture cylinder (with the aid of the photo-electric "0"-impulses), and after "$b/(n \cdot \Delta e)$" revolutions it automatically changes the direction of rotation of the actuating engine, not allowing any further information flow, and it gives a signal to the electronic computer to begin the VIDECORD-program. Now if the photodiodes and sources of light moving backwards come again into ground position (the first diode ought to be just at the edge of the picture) the operation unit stops the actuating engine and it changes the sense of rotation again.

During the working of the instrument, the signs of the photodiodes, arriving continuously, get on Schmitt-triggers. The output levels of the Schmitt-triggers control gate stages. The signs gating the gate stages are furnished by Marker signs. The Marker signs are emitted by a magnetic perceptive head. At the value "Δe" $=0{,}5$ mm every Marker sign is effective. At "Δe" $=1{,}0$ mm every second, at "Δe" $=2{,}0$ mm every fourth Marker sign is effectual, respectively.

The output impulses of the gate stages are directly connected to the input unit of the computer. The effectual Marker signs also serve to synchronize the com-

[1] The ordering program is maked by P. Hunya.

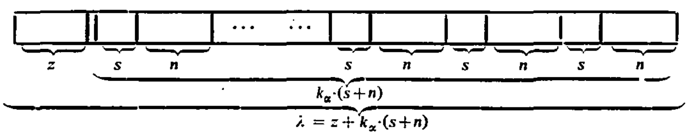
puter. Thus, the picture is transmitted to the computer in a form of parallel-series information. Each photodiode scans just a given zone of the picture, but their joint information gives the entire field of picture. In the further development of the instrument we have set ourselves the task of making the instrument capable of distinguishing the nuances retraceable in the picture. This, on the one hand, raises the requisite concerning the memory capacity of the computer, on the other hand, increases the electronics of the instrument. In fact, in this case, the signs of the photodiodes are conducted onto analogue-digital converters, instead of Schmitt-triggers, and according to the greyness degree of the picture-element, at least 3 bit code per picture element has to be transmitted to the computer.

To carry out pattern recognition experiments (human chromosome analysis) the instrument is connected to a MINSK-22 electronic computer working in our Laboratory.

* * *

The VIDECORD ordering program consists of two parts. The first part of it is suitable for horizontal investigations, and it converts the information into a so-called row-continuous form. If vertical investigations are needed, then the second part of the program, which arranges the information, previously brought into a row-continuous form, is also be used. The entire VIDECORD program leaves the original information set fixed, therefore for the rearranged information as large additional memory capacity is needed as for storing the original information set.

Let the length of word of the computer adapted to the VIDEC instrument be "λ" bits in general. This word plotted against the input visual information is of the following form:



where z = the number of possibly unused bits (e.g. sign-bit);

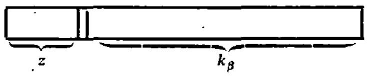
n = the number of elementary informations read-in an effectual Marker sign. It is identical with the number of the input channels (this can be 5, 6, 8);

s = the number of unused bits depending on the input system (ordinarily 0 or 1);

k_α = the number of picture characters of "n" bits stored in one word.

Thus the memory capacity needed to store a picture consisting of $e \cdot e = e^2$ picture-elements is $C = \frac{e^2}{k_\alpha \cdot n}$ words. (The possible values of "e" are 120, 240 and 480.) After each k_α-th effectual Marker sign the VIDEC instrument gives the computer a distinct operating sign to write the information, read in up to that time, in the memory and to increase automatically the adress. k_α has to be chosen such that "e/k_α" is an integer, and that the inequality $k_\alpha(s+n) \leqq \lambda < (k_\alpha + 1) \cdot (s+n)$ holds. The first condition is necessitated by the requirement that, after a full revolution of the picture cylinder, the computer should start storing into a new word by all means, the second one aims the best exploitation of the memory. The values of k_α on the VIDEC instrument can be adjusted 3, 4, 5, 6 and 8.

The placement of the information brought into a row-continuous or column-continuous form in one word is the following:



$$\gamma = z + k_\beta$$

where k_β = the number of the bits used in one word, whether after a row-continuous or a column-continuous rearrangement.

After the rearrangement the memory capacity needed for storing is $C' = \dfrac{e^2}{k_\beta}$

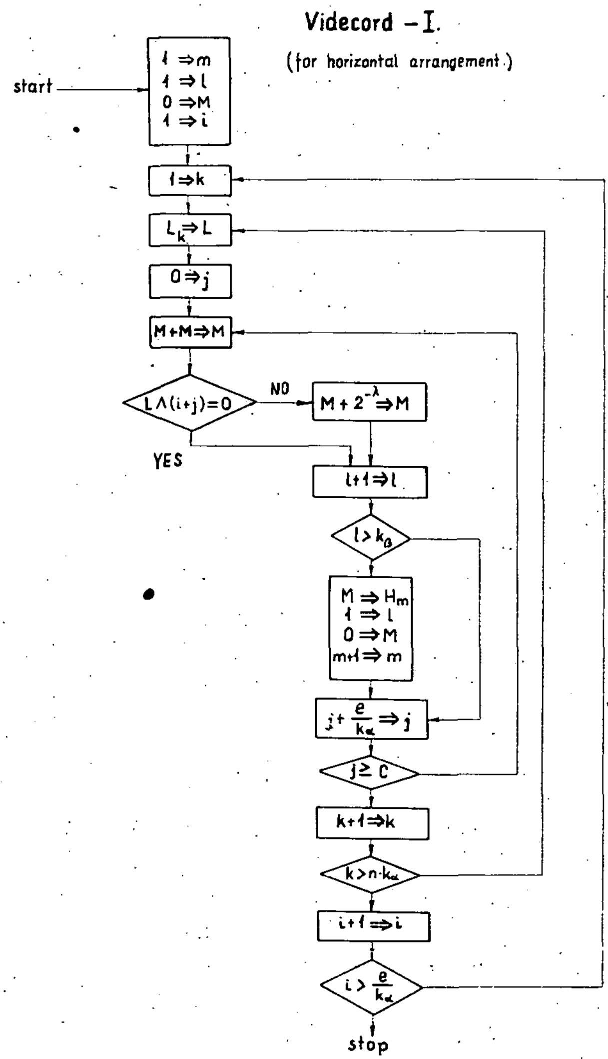


Fig. 2

word. k_β is to be chosen such that "e/k_β" is an integer and that the inequality $k_\beta \leqq \lambda$ holds.

Before the run of the VIDECORD program the following parameters are to be made known to the computer:

e^2 (the number of the picture-elements);
n (the number of the input channels);
k_α (the number of picture characters stored in one word);
k_β (the number of bits used in one word after the rearrangement.)

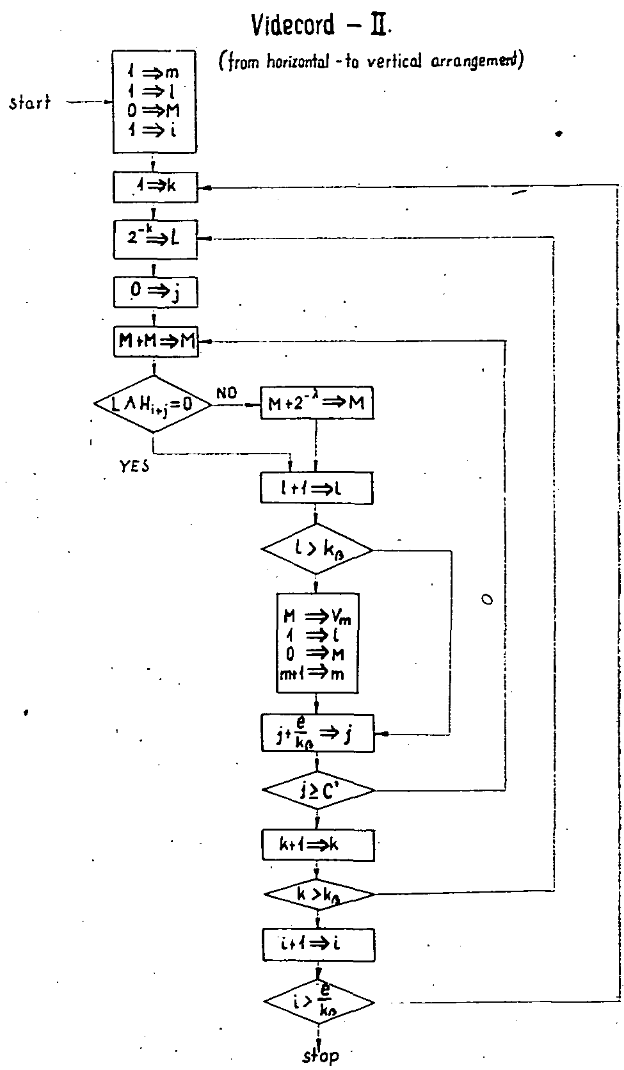


Fig. 3

In the block diagram of the VIDECORD program we use, in addition, the following notation:

H_m = the m-th word of the information brought into a row-continuous form;

V_m = the m-th word of the information brought into a column-continuous form;

i, j, k, l, m = indices;

M = work compartement in the operative memory;

L_k = logical constants depending on the input system separating the corresponding picture-elements ($k = 1, 2, 3, \ldots, n \cdot k_\alpha$).

CYBERNETICAL LABORATORY OF THE
JÓZSEF ATTILA UNIVERSITY,
ARADI VÉRTANÚK TERE 1,
SZEGED, HUNGARY

(Received Dec. 22, 1969)