GERBNER— Content Analysis, Wiley & Sons, 1969

16 VID-R and SCAN: Tools and Methods for the Automated Analysis of Visual Records[1]

Paul Ekman
Wallace V. Friesen
and Thomas G. Taussig

University of California

The ViD-R (Visual Information Display and Retrieval) system was designed for situations where permanent visual records of some phenomena are required, and where the analysis of such records must proceed on the basis of complex decisions made by a human observer. Let us outline some of the reasons why permanent records might be necessary and why analysis would depend on an observer's judgments rather than on some optical scanning procedure.

Permanent visual records, whether still photographs, motion picture film, or videotape, may be necessary if the phenomenon occurs rarely or is nonrecurrent. Permanent records are also necessary if the phenomenon is recurrent but generally inaccessible. A number of recurrent and available phenomena also may require permanent records. For example, when the investigator is uncertain about the unit of measurement and must apply different measurement techniques to the same record, repeated viewing is necessary. Or, the salient phenomena may be so rapid that they cannot be observed at real time but, instead, require slow motion observation. Or, many events occurring simultaneously may not be measureable in one viewing or pass but, instead, require multiple viewing of the phenomena.

Some permanent records can be subjected to automated analysis through optical scanning procedures. But such processing or conversion of a record into digital data is not possible if the events to be measured are complex patterns that are either difficult to define without previous study or are so complex or varied that they are prohibitively expensive to program for an optical scanner. Another, and perhaps more common problem, which prevents the application of optical scanning, is that the record was obtained in an environment where scanning procedures would not be applicable because of the difficulty in discriminating the particular event from the background, or from other irrelevant events. Another obstacle, again a common one, is that the investigation is exploratory; the investigator has not determined a priori units of measurement, but will develop them from inspection.

In the behavioral sciences, visual records of human behavior are often not amenable to optical scanning but instead must be viewed to be analyzed. Too frequently, however, the investigator is seduced by the allure of capturing the phenomenon on his film and neglects consideration of how to convert that record into data until he is confronted with a nearly overwhelming mass of motion picture film or videotapes. Although permanent records have the virtue of slow motion and repeated viewing, they are, one soon discovers, nearly as complex as the original phenomenon recorded; simply viewing them once will not often immediately reveal the proper units of analysis; if the phenomenon was baffling when it occurred, it will remain baffling when it is viewed on film. Moreover, it takes at least as much time to view a permanent record once as the phenomena originally required, and records are not data. While records may be the raw input for intriguing ideas and discovery, they must be converted into some digital form in order to be analyzed. It is no surprise, then, that rather large collections of motion picture films or videotapes of human interactions, stored in laboratories around the United States, not only have not been analyzed but may not even once have been viewed by the investigator. Imagine how long it would take you to look at the record of a complete psychoanalysis of 300 hours, if you watched it at real time only, including the time necessary for getting the film in and out of the projector, and making a few notes about what you observe in each reel!

Usually, when confronted with such a supply of records, the investigator inquires about hardware that might be available to help him wade through this mass. Unfortunately, he has not had much help. There are motion picture projectors that permit excellent viewing at slowed motion and real time and even reasonable accurate frame counts, but that do not allow viewing at faster than real time, have no capacity for fast search, and no automated retrieval. Editing equipment is excellent for fast search, although search cannot be done automatically, slow-motion viewing is not efficient, and image size is small. All motion picture devices suffer from the problems that rearrangement of phenomena for better inspection or specific analysis tasks, requiring cutting and splicing which is costly and time consuming, and, there is no ability to store "on-line" with the visual record any indexing or measurements which could be used for retrieval of specific events for further analysis. Videotape recorders have fast search capacity, usually up to twenty times real time, but location of specific events means reliance on counters that are not indexed in enough detail for most purposes, and storage of indexing materials and measurements must be off-line.

Before describing the hardware system we have developed, let us first describe some of the functional specifications that would be desirable for any human viewing and analysis of permanent visual records:

1. *Viewing speeds.* The operator should be able to view his record at various slowed and accelerated speeds. Slow motion is necessary to insure reliable location and description of the event in some cases and, in others, even to see the event. Fast motion is useful both for rapid scanning to find an event, or for amplifying small slow movements that are difficult to perceive at real time.

2. Digital addresses for search and retrieval should ideally be stored for the smallest unit which a recording unit can differentiate (24 separate addresses per second with sound motion picture film, 30 per second for videotape). Search and retrieval should be possible by these addresses. The system should be capable of writing, storing, reading, and retrieving index information or measurement codes. Whether the index describes the activity shown in the record (for example, one person or dyadic interaction, basket weaving, or child training) or codes the observers' informal estimate of the value of a portion of the record, the ability to store such indexes and then retrieve the relevant portion of the visual record by means of them is crucial if the observer is to profit from the coding based on past viewing or the viewing of other observers. The ability to store measurements or codes and retrieve by means of them makes it possible to cull from the record all events coded in a particular way, and view them again, either to check the reliability of the code, to recode if reliability turned out to be low, or finally to aid the coding of additional events.

3. *Search and retrieval.* The operator should be able to find automatically any event within his record, retrieving this event by requesting an address or by using an index or measurement code. Search should then be reasonably fast (at least 5 times real time) and exact in terms of stopping at the precise location requested. Such search and retrieval is necessary for a variety of reasons: to review previously measured phenomena; to compare events where initial observers coded them in some way; and to eliminate the necessity to view irrelevant events while waiting for the critical event.

4. *Temporal reorganization of the record.* Any series of events located in the record should be capable of being reorganized into a new record, in any predetermined sequence, for further viewing. This capacity allows the operator to sift from a record particular incidents or specific types of events, and gather them onto a single record for more rapid access. The production of such a reorganized selected record should, of course, be automated, based on operator instructions supplied in terms of code indexes, time code addresses, or coded descriptive measureeents.

5. *Access to a visual library.* The observer should have rapid access to a visual example which defines the meaning or the boundaries of any of the investigator's codes or measurement procedures. Often the coding of human behavior involves the development of classes of behavior or units that are difficult to describe verbally or to remember when described verbally. A visual dictionary or library, with visual examples of the criterion for each type of event or code, can be the most economical way to define a measurement procedure. If the operator has immediate access to any entry in the library, he can retrieve that entry and visually compare the library definition with the event he is attempting to code.

VID-R was designed to accept either 16mm film or videotape as the original recording medium, although the analysis is performed on a videotape version of the record. We felt it important to design the system, to accept either 16mm film or videotape as the original recording medium, for two reasons: (1) so that already existing film records and archives could be employed (and there are many large archives on film of phenomena which would be difficult or expensive to record), and (2) so that the high resolution possible with film records could be preserved if the other conditions of recording allowed full exploitation of this film potential. Actually, the choice of whether to record on film or videotape must be guided by a number of considerations. Although videotape as the original record sacrifices some of the resolution possible with film when lighting is bright, under the more usual conditions of lighting, necessary to record the subject without his knowledge or without his continual cognizance of the lights, videotape may produce as good or better a record as film. Videotape can allow recording continuously for 1½ hours, while film is usually limited to 30 minutes. Because of lower costs, videotape can be used with multiple cameras; with two videotape records, either from different angles or one with a zoom, for less than the costs of one film record. And, of course, videotape can be erased or rerecorded; thus the camera can be left to capture unpredictable events and the waste time reused again. A final consideration is feedback to the investigator; with videotape, one can learn immediately whether the phenomenon being recorded is actually being recorded in a satisfactory way, but the use of film imposes at least one day's delay, while the film developer works, before you know whether you recorded what you think you did. Regardless of how the original record was obtained (film or tape), we think it must be transferred to tape in order to perform analysis in a way that meets the five demands we just outlined.

The components of the VID-R system are:

One film-to-television chain which allows the transfer of 16mm optical sound movies onto the video and audio channels of a videotape recorder.

Two Sony PV 120U videotape recorders with complete remote control of the functions; playback, record, fast forward, rewind, variable slowed-motion, stop motion, variable high-speed playback, and stop.

One video-disc recorder capable of recording at least 20 seconds of video information and playback at high-resolution slowed and stop motion.

Three high resolution television monitors.

One Teletype ASR 33, keyboard-printer with papertape punch and reader. This is the means of operator-system communication.

One Digital Equipment Corporation PDP-8 programmed data processor. This low-cost computer provides the logic for the operations to be described.

One video and audio interface to perform data transfers between the computer and recorders.

Three videotape recorder controllers capable of performing the instructions of the computer, to place the recorder into the proper motion to perform the task.

One or two high resolution Vidicon cameras for field recording.

The Film-to-Videotape Transfer

When 16mm film has served as the original medium for the research record, the film-to-television chain would be employed. Our chain uses an L-W Athena, 16mm Analyst sound projector modified for TV frame synchronization. This projector is capable of moving the film forward or reverse, at slowed motion of 1 to 12 frames per second, holding indefinitely on any single frame, or operating at the normal sound projection speed of 24 frames per second. This projector also has a frame counter that can be used to locate specific places on the film when only partial record transfers are required.

The COHU 8507 high resolution Vidicon camera, employed on the chain, is mounted on a motor-driven carriage to allow zooming. At one extremity of the Vidicon camera track, a one-to-one copy of the 16mm image is converted into a television signal. At the other extreme, the image is enlarged 6 times, and by up-down and side-to-side movement of the lens between the projector and camera, any portion of the filmed image may be focused on the Vidicon. Optics are not employed on either the projector or camera, but there is a single field lens in between projector and Vidicon, which can be adjusted for focus. The signal generated by the Vidicon camera may be displayed on the monitors and/or recorded onto videotape.

This film chain's unique features of slowed and stop motion projection and the capacity to zoom the image to six times normal size provide a flexibility for film analysis that will become more apparent as the applications of the system are elaborated. However, with the film chain alone, extensive editing of film can be made through this system. The control of finding and holding a position on the film and flexibility of projection speeds allows the investigator to produce tapes containing only salient events that he wishes to analyze, recorded at several speeds to aid his viewing.

Writing Digital Codes on the Videotape

As a video transfer is made of a film, the computer controls the recording function of the videotape recorder and generates sequential binary codes that are inserted outside the normal viewing portion of the picture on each video frame. In the data write mode, codes are loaded from the computer into a register and then shifted out and recorded serially in the horizontal scan of the video. Twelve bits of data per scan line are written on the videotape commencing after the horizontal synchronizing pulse. Information is recorded with a redundant code so that the information can be checked and corrected for error during later "reading." Thus when numbering frames, 30 video frames per second are labeled with a discriminable code of 6 numerical digits, so as to allow the retrieval of any one of the 162,000 frames on a 90 minute reel of videotape. Locations of the beginning and end of visual behavior may therefore be established to the smallest measurable unit: the video frame.

Frame coding is not restricted to videotapes made during a film-to-television transfer. Videotape records made in the "field" and duplicate videotapes from a tape master may also be "numbered" during the recording of the video information. Furthermore, the process of generating sequential frame numbers may be interrupted. The computer is capable of "reading" the last locational code of the material presently recorded on a videotape and then will continue numbering all additional frames from that point of reference.

The computer and video interface equipment is not restricted to "writing" codes on any particular part of the video frame. The computer may instruct the interface so that code information continues to be inserted on each succeeding scan line of a video field until all information is recorded. Thus entire video fields may be used to record compiled locational indexes, digitally coded descriptions of the recorded behavior, or any other relevant information that can be coded digitally. The capacity of the computer-controlled-interface components to "write" and "read" binary codes recorded on the videotape is basic to the functional specifications described earlier. The components and procedures involved in meeting those specifications will now be described in some detail.

Viewing the Videotape Record

The operator can request, through the computer controller, all of the remotely controlled operations of the video-recorders. Normal playback, stop-motion, and variable slowed motion can be requested in either the forward or reverse mode. Fast forward and fast rewind can also be requested for search operations. An additional feature is that, through computer control of the speed of fast forward and rewind modes, a faster than real time playback can be requested. Finally, because the two video recorders and the video-disc recorder are all connected to the system, episodes on the videotape can be transferred to the disc for high resolution playback at slowed and stop motion speeds.

Frequently used sets of operations may also be requested by designating a "subroutine." For example, the following sequence can be repeated throughout the viewing of a record. The video-recorder is directed to search at fast forward speed for the beginning of a specified behavioral unit; this event is played back at normal viewing speed from beginning to end. The recorder returns to the beginning of the unit and repeats the playback at normal slowed or fast play speed with the simple designation of the speed. The recorder then can be directed to return to the beginning of the event to transfer to the video-disc if there is a need for higher quality slowed motion, or in careful scrutiny of single frames. When the operator is finished, he pushes the advance key and the recorder is moved to the next event on the tape. When a second video-recorder is used for comparative viewing, the computer-controller may be displaying events from one recorder while searching for another event on the second recorder, or as part of the "subroutine" may be making duplicate recordings between the display of events.

If there is a written record of the location of several incidents, the beginning and end frame numbers for a large group of units may be requested when the operator begins his coding operations. The computer would store the requested group and search, retrieve, and display each item in turn. This method, however, would make use of "off-line" records of the locations, codings and descriptions of incidents that have already been isolated. The VID-R system precludes the necessity for excessive "off-line" bookkeeping. Instead of requesting the display of a series of units by their previously located position in the tape, the operator may request the display of all incidents that have been described in a particular manner. The computer stores this request, refers to the summary index recorded at the beginning of the tape, locates and displays each incident meeting the description, and advances to the next event only after the operator has inserted additional codes or corrections and has otherwise completed his inspection. These locational procedures will be discussed in more detail.

High-Speed Search and Retrieval

Under the control of the computer, the videotape recorders are operated at the most economical speed to any requested position on the videotape record. When receiving a request to move the videotape to a particular location, the computer must "read" the position of the videotape from the frame locational code, compare this location with the one being requested, and determine the direction and speed of movement that is required. For long distances, the videotape recorder is switched into the high-speed forward or rewind mode and through the use of dead-reckoning techniques, moved to the approximate location. The recorder is switched to a slower speed and its exact location determined. As necessary, the computer returns the recorder to the high-speed mode, goes into the play mode, or reverses the direction of tape movement. When the requested position of the tape is approached the recorder is switched into the play mode and stopped at the exact position as desired. Using these techniques, the maximum time required to search from one end to the other of a ninety minute reel of videotape is approximately six minutes or 15 times real time. For short search distance the normal play speeds, slowed motion or controlled high-speed play modes are used to locate most economically the desired frame, so as to not overshoot the target location. During all locational operations, the video switching does not allow noise and irrelevant visual material to be displayed on the monitors, thus reducing confusions and distractions for the operator.

Temporal Reorganization of the Visual Record

Having determined the locations, coding, and descriptions of a set of events, the operator may wish to review visually all events given the same code. This can be done by requesting a high-speed search, then display of each of these events. However, this system is also designed to allow the operator to describe the set of events by code or location and to request the automatic and continuous retrieval and rerecording of each occasion onto a second videotape. In this operation, each set of events where comparison is desired is clustered together. When played back, this "edited" videotape allows the operator to view all events given the same code without any interruptions from irrelevant material or search time. The production of an "edited" record from a single 90 minute videotape requires no operator functions after the selected sequences and their ordering have been requested. What has always been a tedious and expensive editing task with movie film (or could have been accomplished through tedious film-to-videotape transfers on the chain) can be completed automatically and at nominal expense. The advantage of this is that, for the first time, it is feasible to check reliability and validity of coding visually without frustrating interruptions and time delays.

Off-Line Records and Data Analysis

Since all communication in and out of the VID-R system is through the teletype, a typed or papertape record can be maintained of all located events, their codes, locations, and descriptions. Also, a videotape can be played back and the computer can direct a teletype record of any coded information. The duration of each event can be easily determined as well. When the computer is not needed for control of the rest of the system for a period of time, events can be collated by their location, mean durations, and frequencies; and the conditions under which particular events occur most frequently can be calculated by the computer. When the on-line operations of the computer are required for control functions, then the papertape can be read by an off-line system and even more complicated computations based on the codings, locations, and descriptions can be done. The unique capability of the VID-R system, in regard to off-line operations, is that records do not have to be kept in file but are collated from the videotapes only when needed.

The Visual Library

In the analysis of visually recorded information it is often desirable to define items in visual rather than verbal terms. If the operator who is coding a visual record has access to a library that contains visual examples of each item he is to locate or code, he can make a more precise judgment, through comparison of the library example with the record to be coded, than if he must translate a verbal description into visual terms that he then compares with the visual record. Three problems that have previously interfered with the use of such a library are overcome by the VID-R system: fast access to and retrieval from the library; compiling the library; updating or reorganizing the library.

In the VID-R system, the operator describes by means of a code relevant features of the visual event he is about to analyze. His coding descriptions are compared by the computer with the coding descriptions of the library entries; the computer then finds and displays on a monitor any library entries that have similar codes, and the operator can compare them with the event he is attempting to classify. If the computer finds no applicable entries in the library it informs the coder, and he can activate a subroutine that transfers the new event and its codes into the library, as a new entry defining a new class of phenomena.

The original library is compiled by transferring specific examples from other videotape records, with descriptive codes of each entry so that later retrieval is possible. When the events to be entered into the library have been determined, the operator can instruct the computer where to find them in the record, and the production of a library tape is then performed automatically. Adding additional entries to the library, as new events or classes of phenomena are encountered, is a simple matter, following the same procedures.

Up-dating the library, either by changing the specific examples in library definition, or by summarizing definitions into a broader definition, or by refining a single definition into two or more definitions, can be accomplished with precise location of the entry and precise insertions or deletions by computer control.

Although the use of the library procedure may sound complicated, its feasibility in time and expense is best seen by comparison with creating a library with film. If the library were built from film, there would be no computer-directed way of finding entries in the library, unless an off-line device were used. The creation of the library would be costly in time and materials, involving duplicate printing and compiling by splicing or optically printing the entries onto a reel. Even more terrible to contemplate is the cost of tearing such a library film apart to update, reorganize, or substitute examples.

APPLICATIONS OF VID-R

Now let us discuss some of the applications of VID-R. We shall illustrate the use of VID-R by describing, first, procedures we have developed for the analysis of nonverbal behavior shown by mental patients, then a technique for the complete analysis and classification of movements, and then a technique for the more selected analysis of specific types of movements. Finally, we shall briefly discuss how VID-R might be used to index film archives, and how it might be used in a completely different application: programmed learning.

Systematic Classification and Analysis of Nonverbal Behavior (SCAN)

We have developed the SCAN procedures for classifying all observable body and facial movements without reference to an a priori theory about classes or types of movements. The unit of movement, *the act,* is defined as either (1) the start and stop of motion in a given body area, or (2) the change of visual configuration in a continuous movement. SCAN is applied separately to the following body areas: head, hands-arms, shoulder-arms, feet-legs, knees-legs, and the face. These areas are divided in terms of possibilities for independent action, and are separated for analysis so that the coder can concentrate his visual attention with minimal distraction from other areas of the body. Let us explain SCAN in terms of the coder operations for one body area: the hands-arms.

The coder begins his work with a tape showing only this area of the body. During his first viewing, the coder locates the beginning and end frames of each visually distinctive movement. He instructs the computer to have the videotape recorder play back the record at real time. At the moment he observes movement in either hand, he presses a teletype key that reverses the recorder a few frames and plays back the approach to the movement at slowed motion. When he observes the first frame of movement he pushes the "start" key on the teletype. The computer reads the number of the frame that the operator has indicated, punches the number on papertape, and stores it in its core memory. The coder continues through the movement until the hand is motionless or the visual configuration of the movement changes and presses the "stop" key. The operator then describes the movement in terms of a code for the visual configurations of hand movements. The location of the end and the visual configuration code are also stored in the computer memory. At this point the computer moves the tape back to the beginning of the act and plays it back while recording it on a second videotape with information identifying it, for example, original identifying information codes and any code information. This second tape, which is on the second videotape recorder, wil be referred to as the "working tape." The coder then proceeds to the next point at which either hand moves, and codes that movement, which is in turn transferred to the working tape. If this first step of coding has been done accurately, then the working tape should contain all hand movements shown on the original record.

Two verification steps usually follow. While reading begin and end locations from the working tape, the original record is played back at normal viewing speed. Each time the beginning and end of a movement is reached in the record the coder is notified ~~audibly or by~~ visual marker. Any movements that he has missed on the first viewing, or incorrectly located or coded, can be recorded at the end of the working tape.

The original locations and descriptions of the acts incorrectly identified on the first viewing are listed by the teletype, and erased from the working tape at the end of this verification phase. The second verification step is to look at the working tape and determine that each movement was recorded completely from beginning to end. This verification is usually done simultaneously with the next coding phase.

To continue our example of the coding of the hand movements in a single record, the coder puts the working tape on one recorder, and the library tape on the other. When the working tape approaches the first act, the act's description is transferred automatically to the computer, and the search of the library tape for examples of similar description is initiated automatically. This library search proceeds while the operator is viewing the first act to be coded. The coder compares the act with all similarly described acts in the library and, if any one of the library examples is like the new act, the new act is given the same categorical code number. If no movement on the library tape is like the new act, the new act is assigned a new categorical act code and rerecorded onto the proper location of the library tape with all codes, descriptions and locations. The library tape is returned to the beginning of the section and the coded information is added to the index. When this coding is completed, all codes, descriptions, and act locations are written at the beginning of the working tape and original tape.

The final verification step follows. A second working tape is made; all acts that were coded in the same manner are clustered together, adjacent clusters being determined by similarities in description. This tape is produced completely by the computer-controller system and allows the visual verification of the similarities of acts coded as the same, and of the distinctive features of acts coded as different.

After *all* body areas ~~have been coded~~ and verified in the manner just described for the hands, we would consider that we have ***transformed* the behavioral record into *data* which can then be analyzed.** For any area of the body, we know how many different types of acts occurred, the frequency of each type of act, the duration of each act, its exact temporal location and, if we like, the correlations between acts from different areas of the body. We can at any time visually retrieve any act from any body area for inspection. SCAN does not itself tell us the meaning of the nonverbal behavior; the SCAN output is the beginning point for our analysis of the data; SCAN transforms the film record into data which can then be studied. We have described elsewhere (Ekman, 1965; Ekman and Friesen, 1967a) how we analyze SCAN output. Let us summarize briefly the methods we use so that you will not confuse the SCAN procedure with the methods for analyzing the SCAN output. Let me use, as an example, a hand-rub-hand act.

The SCAN output would tell us how frequently this occurred in a particular film, the duration of each such act, and its exact location. Our first method of analysis involves a search for the common characteristics (demographic or personality) of those individuals who frequently show this act; we would expect, for example, to find a higher frequency of this act in our records of interviews with agitated depressives than with nonagitated depressives or schizophrenics. Our second method of analysis is to determine any similarities in the setting when the hand-rub-hand act occurred. We have filmed psychiatric patients at the beginning, middle, and end of hospitalization; our expectation is that this act occurs more frequently at the beginning or middle of hospitalization than at the time of discharge. Our third method of analysis is a search for any other acts that typically accompany, precede, or follow this act, whenever it is shown. We might find that hand-rub-hand acts often occur when the patient breaks eye contact or turns away from the interviewer. Our fourth method of analysis is to check the verbal behavior emitted whenever this act occurs; some nonverbal acts illustrate the verbal behavior in one of six different ways (see Ekman and Friesen, 1967b), but hand-rub-hand acts are not illustrators, and there is probably no relationship to any verbal content across people who show this act. Our fifth method of analysis is to show this act, isolated from its context, to a group of naive observers, asking them to describe their impressions of the person, and analyzing their descriptions for elements common to this act but not present in their descriptions of a different hand act. We would exect hand-rub-hand acts to convey the message of nervousness and self-soothing.

Analysis of Critical Incidents in Nonverbal Behavior

The use of VID-R in the study of nonverbal behavior is not limited to the type of grass-roots empiricism we have just described, in which every movement is isolated, classified, and counted. VID-R can also help in the study of selected incidents considered of critical importance for a theoretical reason. We are in the midst of such an application of VID-R in the study of hand-to-face movements. These movements are considered critical incidents for three reasons: (1) because the face is the site for sensory inputs, for breathing, eating, and making sounds, what the hand does to the face can reveal very personal psychological information relevant to the person's attempts to aid, interfere or otherwise deal with these activities; (2) because the face is the site for affect displays, when the hand touches the face it can show how the individual plans to cope with any particular emotion; and (3) because the face represents the self for many people, the hand touching the face can be interpreted in terms of what the individual is doing to the self, for example, attacking, supporting, and soothing. (A full description of the theoretical rationale of these self-adaptors appears in Ekman and Friesen, 1967b).

In considering hand-to-face acts, there are four variables to code: *location*, what part of the face is touched; *action*, what the hand does to the location; *hand part*, what part of the hand is involved in the action; and *duration*, how long it takes for the hand to get to the face, how long it is involved in the action, and how long it takes for the hand to leave the face and go elsewhere. Just considering the first three variables, a matrix could be generated consisting of at least 20 locations, 9 actions, and 7 hand parts. But do all of these occur in our records? Do they occur with sufficient frequently that we will be able to check any of our hypotheses about meaning? Here is where VID-R helps us.

The hand-to-face acts are scattered somewhere within the 120, 12-minute films of psychiatric patients that we have. Just to look at each act once, in order to develop a code, to determine how many cells in the action-location-part matrix should be considered, would take many days and, if we needed to look again or to have coders apply a scoring system, it would take just as long for them to find each occurrence, and to get each reel in and out of the projector. If we wanted to avoid such time investments, we would have to cut the films, splicing together onto a single reel(s) all hand-to-face acts; but this would be expensive in materials; with VID-R, we proceed as follows.

1. A videotape copy is made of each of the 120 films. The videotape is viewed on a monitor at 5-times real time, since hand-to-face acts are sufficiently salient to be noticeable even at fast time. Whenever the operator sees a hand-to-face act he activates a subroutine that slows the recorder down and backs it up; he then follows the procedures we described earlier, locating the precise beginning and end of the act and transferring the act onto a second videotape, with identifying information. Thus a working videotape would be compiled of all hand-to-face acts.

2. With such a working tape, or tapes, we could simply sit down and look at all occurrences in a relatively brief time. Or, we could have the operator apply an a priori code to all occurrences, storing the code digitally on the tape with each occurrence. VID-R could then produce a frequency distribution for the coding matrix and calculate operator reliability in using the code: after inspecting this information we could instruct VID-R to find and retrieve all hand-to-face acts with a given set of codes, so that we counld visually check whether we are classifying as the same, acts which actually look alike.

Again it should be clear that VID-R is not analyzing data. It is collecting and organizing events that, when coded, are in a data form which can be analyzed. When the hand-to-face operations are complete we would have a listing of each type of hand-to-face act, the frequency for each patient, its durations, and its exact locations. To determine the meaning of hand-to-face acts, we would then apply to the VID-R output the five methods of analysis described above.

Indexing Film Archives

In our cross-cultural studies of nonverbal behavior we have been recently working with Dr. Carleton Gajdusek of the National Institute of Health, in the analysis of motion picture footage he and E. R. Sorenson have visited New Guinea over the past ten years. Their films are a rich source of information about two very primitive groups of people, the

Fore and the Kukukuku; most investigators would have difficulty recording this invaluable information which in a few years will, through culture contact, be gone. While these films contain a great deal of nonverbal behavior, they also show everything else possible about these cultures, since Gajdusek and Sorenson viewed as their purpose, at least in part, to record events that would be of interest to various scientific disciplines. They have effectively prepared their footage to preserve information about when and how it was taken (Sorenson, 1967). But when one is talking about an archive of over 60,000 feet, there is a need not only for indexing but for the indexes to be search tags for rapid retrieval and visual display.

Let us say we are botanists, and we want not only to know how much footage they have on specific botanical phenomena but also to take a look at such footage and see whether it would be of sufficient pertinence for us to request a copy of those film sequences for further analysis in my laboratory. If their film were indexed (and Gajdusek and Sorenson have begun indexing), we might know where to look in the 60,000 feet of film; but finding each sequence would be enormously time consuming. VID-R could be of considerable help.

Let us suppose that a general index were devised which labeled botanical events, manufacturing, musical episodes, child rearing, child play, and adult conversation. The 60,000 feet of film would be copied onto videotape, and the index stored on the videotapes. Now, as botanists, or any other investigators with specific interests, we may request through the index to view the events of potential interest; and VID-R uses the fast search to find and then display them to us.

Programmed Instruction

Often in programmed instructional materials there is a need to present information visually, as with drawings, photographs, films, or even printed material. VID-R offers the possibility of being able to store both single frame material and sequential moving visual events, indexed digitally on the tape, and thus retrieval by computer control at whatever point they might be necessary for presentation to the student. With videotape as the format, changing the entries can be accomplished with minimal costs. The main disadvantage of VID-R in this application is that search time could take up to 6 minutes; but this could be considerably reduced if material were stored in other than a random fashion, so that the need to search from one end of the tape to the other for the next step in a programmed learning sequence would be infrequent.

[1] Paul Ekman and Wallace Friesen are associated with the Langley Porter Neuropsychiatric Institute, University of California Medical School, San Francisco. Thomas Taussig is affiliated with the University of California, Berkeley, and the Lawrence Radiation Laboratory. We express our appreciation to the persons who aided the development of this system: to Henry Jacobs who offered his experience in the video field to help us evaluate the latest equipment available and to introduce us to other knowledgeable persons in this field; to Jerone Russell for guiding us to creative persons working with computers, video, and interface equipment, and for offering his own knowledge of these fields to evaluate prospective contractors for the system; and, finally, to Charles Steinheimer, Field Representative for Sicular X-Ray, Donald Horstkorta, Regional Representative for Sony Corporation, and Jack Sanford, Senior Field Engineer for COHU Electronics, for sophisticated technical suggestions and guidance which allowed the development of an integrated system. Research and development of the equipment and methods were supported by a research grant (MH 11976-03) and a career development award (1-K3-MH-6092-02).