Use of computers in dysmorphology

JOHN H DILIBERTI*

From the Departments of Pediatrics and Medical Genetics, Emanuel Hospital, Department of Pediatrics, Oregon Health Sciences University, and Department of Computer Science, Oregon Graduate Center, Portland, Oregon, USA.

SUMMARY As a consequence of the increasing power and decreasing cost of digital computers, dysmorphologists have begun to explore a wide variety of computerised applications in clinical genetics. Of considerable interest are developments in the areas of syndrome databases, expert systems, literature searches, image processing, and pattern recognition. Each of these areas is reviewed from the perspective of the underlying computer principles, existing applications, and the potential for future developments. Particular emphasis is placed on the analysis of the tasks performed by the dysmorphologist and the design of appropriate tools to facilitate these tasks. In this context the computer and associated software are considered paradigmatically as tools for the dysmorphologist and should be designed accordingly. Continuing improvements in the ability of computers to manipulate vast amounts of data rapidly makes the development of increasingly powerful tools for the dysmorphologist highly probable.

During the past two decades digital computer technology (DCT) has evolved to the point where some form of computer usage may be found in virtually every professional and business environment. This evolution has been especially evident in the field of medicine because of the need to manipulate and control large amounts of information. In presenting the role of computers in dysmorphology, I shall attempt to discuss only those areas of DCT which manipulate data or information usually managed directly by the dysmorphologist. (Obviously this will result in an emphasis on diagnostic issues.) Thus, although computerised tomography has become essential in the evaluation of many aspects of morphogenesis, the algorithms used to reconstruct the two dimensional image from the raw intensity data will not be considered germane to this review. The use of computers to extract and manipulate information in the 2D image, or even reconstruction of a 3D image from 2D information, will be shown as relevant, however.

In a review article of this length, detailed discussion of the principles of computer science and characteristics of computer hardware is not feasible. The reader is referred elsewhere for such information.^{1,2} Knowledge necessary to understand the principles illustrated in this article will be developed as necessary.

The computer as a tool

Before examining computer applications specific to dysmorphology, one important concept must be developed: the paradigm of the computer as a tool. The rapid development of DCT has led to many misconceptions and misgivings, ranging from science fiction models of artificially intelligent computers largely eliminating the need for human thought to equally naive views of the computer as a mechanistic automaton lacking a proper role in a humanistic society. Regarding the computer as a tool to assist, rather than replace, a human being is a perspective likely to lead simultaneously to better design of computer applications and better acceptance in the workplace. More optimal design of computer applications for dysmorphology should result from thoughtful analysis of the tasks performed by the dysmorphologist and the consequent design of a tool to assist in the performance of that task. The analysis of the tasks performed by the dysmorphologist will be emphasised in this review of the functions of DCT in dysmorphology. Failure to assess the need for a specific computer tool has often led to the development of sophisticated computer technology which was not useful because it did not emulate an appropriate tool.³

Task analyses for the dysmorphologist

One of the most frequent tasks the dysmorphologist

*Present address: Department of Pediatrics, St Francis Hospital and Medical Center, 114 Woodland Street, Hartford, Connecticut 06105, USA.

Received for publication 16 November 1987.

Accepted for publication 17 November 1987.

is called upon to perform is to make a diagnosis for a patient or family. On some occasions a mere glance at the patient is sufficient to make a presumptive diagnosis, but at other times the dysmorphologist completes the history and physical examination without reaching even a tentative conclusion. Faced with this situation, the dysmorphologist must perform a series of tasks to achieve the goal of reaching a diagnosis. In order to pursue further the paradigm of the computer as a tool, it is necessary to create a model of these tasks in order to create a context in which to examine actual and possible computer tools. The proposed model is to be taken as a plausible approach to diagnosis and the reader should not conclude that it is the only reasonable one.

In the absence of DCT the first step taken by most dysmorphologists when faced with a puzzling patient would generally be to examine reference texts using one of several strategies. One possible strategy is to select a single physical feature found during the examination of the patient and search through the reference texts for syndromes or diseases with that feature. Obviously, if the feature selected is found in a large number of syndromes the dysmorphologist will be forced to examine many disorders of little or no interest to the case at hand. If the most unusual feature is selected, however, the search should be considerably truncated. In either event, the result of this task is a short list of candidate syndromes to examine further.

The preliminary list can be approached by comparing the list of features for each syndrome in the list with those features found during examination of the patient. This comparison may also include the use of a photographic atlas. The result of this task will be a shorter list, perhaps as short as one syndrome.

After further contemplation, and probably also the examination of original articles in published journals, the dysmorphologist who is still perplexed may call upon the knowledge of an expert for the syndrome under consideration. Based on the information gathered by the dysmorphologist, the expert might render an opinion regarding the patient.

The tasks described above may be outlined as follows:

- (1) Select a feature and search the index of a book of syndromes to create a short list of candidate syndromes.
- (2) Compare the rest of the patient's features with those found in each of the candidate syndromes, shortening the list.
- (3) Using published references and photographs, select a presumptive diagnosis.

- (4) Obtain an opinion from a dysmorphologist expert on the syndrome(s) in question.
- (5) If no candidate diagnosis has been found, enquire about other cases seen by colleagues, and perhaps identify a new syndrome.
- (6) Perform additional tests to assist in the diagnosis.

For each of these tasks, actual or theoretical computer tools will be considered and critiqued. In addition, the application of DCT to the study of two and three dimensional biological shapes will be discussed.

Specific roles for computers in dysmorphology

Perhaps the most visible role for DCT in relation to dysmorphology has been what might be called the syndrome database (SD) system. Implementation of SD systems was attempted at least as far back as the early 1970s, but required expensive mainframe hardware, resulting in limited accessibility. Subsequent microcomputer based SD systems have become feasible and fairly widespread, since current desk top computers are generally quite affordable, yet at the same time more powerful than their earlier mainframe ancestors.

Most modern database systems are set up using the relational model, which is relatively simple conceptually.⁴ The data in a relational database system are set up as a set of one or more tables, which may be visualised as accounting spreadsheets (table 1). In the SD database each row has a label corresponding to a syndrome and each column is labelled with a feature. If the syndrome in a given row has the feature under a column label then a mark is made in that column (table 2). In a somewhat more complicated representation, the

TABLE 1 Sample relational database: employment by city for health care.

Job title	City				
	ABC	EFG	HIJ	LMN	XYZ
Physician	10	5	17	21	58
RN	21	12	37	40	126
Lab tech	4	2	7	8	17

TABLE 2 Sample SD relational database (non-quantitative).

Syndrome	Feature			
	MR	CHD	SS	
Down	+	+		+
Turner	-	+		+
Achondroplasia	-	-		+
Klinefelter	-	-		-

mark could have a qualitative or quantitative meaning, such as "In syndrome X feature Y is present Z percent of the time" (table 3).

The actual use of a printed table of syndromes would be somewhat impractical except, perhaps, as a quick reference guide to a limited number of syndromes. The clinician would read across the top of the table until the column containing the feature of interest was identified and then would read down that column, noting or marking all of the rows (syndromes) which contain that feature. The subset of syndromes identified in this fashion could be further reduced by selecting another feature and examining the appropriate column for this feature, in only those rows which have already been selected for the first feature. With patience, a large table could be searched and the syndrome(s) with matching features extracted. This repetitive searching of a large table is obviously a much better task for a computer than for a human. Several computerised implementations of this method have been produced and are currently available for small computers. One example, the London Dysmorphology Database (LDD), will be analysed.⁵

The LDD consists of a program written in the relational database language, dBase (versions II or III), along with a set of large files which contain the data, and some additional files which speed the process of finding the data.⁶ Of the several options which the program offers to the user, the 'Code entry/review/search' function is most relevant to this discussion, since it is the tool which corresponds most closely to the task outlined above. The user enters the identified clinical features by following a series of prompts provided by the program (table 4a, b, c). For example, if the patient has been found to have a cleft lip, the user first selects the code for Mouth, code 11, the most general category. A second prompt then allows a more refined description, Upper Lip, general abnormalities, 04, and finally the third level of prompt gives the most fine grained code, Cleft Lip (non-midline), 01. The resulting numerical code, 11.04.01, is used to search the syndrome database. Each additional feature is encoded in the same fashion resulting in what should be a reasonably complete clinical description of the patient according to the available LDD nomenclature.

Note that it is possible to preserve generality, to some extent, by selecting the codes, 00 or 01, which usually refer to 'general abnormalities' at the second or third levels. If the database is searched using the codes 'Colonic tumours' (18.03.04) and 'Muscle weakness/myopathy' (30.01.07), the LDD identifies a single candidate syndrome, the Ruvalcaba-Myhre syndrome. If, on the other hand, the more general categories 'Abdomen, general abnormalities' (18.01) and 'Muscles, general abnormalities' (30.01) are used for the search, a dozen candidate syndromes are identified, the Ruvalcaba-Myhre syndrome among them.

The next phase of LDD usage involves the identification of a search strategy. As noted above,

TABLE 4a LDD menu. Code selection.

Select a code:	
(01) Build	(18) Abdomen
(02) Stature	(19) Pelvis
(03) Cranium	(20) Genitalia
(04) Hair	(21) Urinary
(05) Forehead	(22) Upper limbs
(06) Ears	(23) Hands
(07) Eyes (globes)	(24) Nails
(08) Eyes (general)	(25) Lower limbs
(09) Nose	(26) Feet
(10) Face	(27) Blood vessels
(11) Mouth	(28) Endocrine
(12) Oral region	(29) Haematology/immunology
(13) Teeth	(30) Muscles
(14) Voice	(31) Joints
(15) Neck	(32) Neuro
(16) Back and spine	(33) Skeletal
(17) Thorax	(34) Skin
Enter code: 11: : :	

TABLE 4b LDD sub-menu. Code selection.

Select a code:	
(00) Mouth	
(01) Mouth, general abnormalities	
(02) Lower lip, general abnormalities	
(03) Philtrum, general abnormalities	
(04) Upper lip, general abnormalities	
Enter code: 11:04: :	

TABLE 4c LDD sub-menu. Code selection.

Select a code:	
(00) Upper lip, general abnormalities	
(01) Cleft upper lip (non-midline)	
(02) Midline cleft upper lip	
(03) Prominent upper lip	
(04) Thick upper lip	
(05) Thin upper lip	
Enter code: 11:04:01:	

TABLE 3 Sample SD relational database (quantitative).

Syndrome	Feature (%)			
	MR	CHD	Y	SS
X	M	N	Z	Q
ABC	100	40	5	100
QRS	3	10	20	97

selecting the most unusual feature is often an efficient approach. As an example, consider the CHARGE association. Table 5a lists the number of candidate syndromes found by the LDD for inputs of single features. Obviously, the more unusual the feature selected, the fewer syndromes will be found. Since the LDD allows the user to specify the use of several features in the search of the database, additional specificity is possible. This is accomplished by input of the form '[1] and [2] and not [3]', which means search for all syndromes which include features 1 and 2 and not feature 3. In addition to the connectives 'and' and 'and not', 'or' may also be used, in essence allowing the user to combine two searches simultaneously. As a general rule, the greater the number of features used in a search the smaller the number of potentially matching syndromes will be identified. As with searches using single features, the more unusual the features entered in a search using multiple features, the fewer syndromes will be identified. These principles are illustrated in table 5b. The reader should note that using the most unusual feature first improves the efficiency of the search if only a few of the features are used. If, however, the entire set of patient features is used, the same candidate syndromes will be identified, independent of order. The user must design a strategy which identifies a reasonable number of candidate syndromes, being neither too exclusive nor too general. By selecting a small number of the most important features (as judged by the dysmorphologist) for a given patient,

TABLE 5a *Number of syndromes found in the LDD for each feature.*

Mental retardation	533
Short stature, proportionate	309
Small mandible	195
Congenital heart disease	104
Coloboma of iris	39
Choanal atresia/stenosis	19

TABLE 5b *Number of syndromes found in the LDD for multiple features.*

Two features	
Mental retardation/short stature	183
Mental retardation/micrognathia	86
Mental retardation/coloboma	21
Short stature/micrognathia	50
Short stature/congenital heart disease	23
Micrognathia/congenital heart disease	24
Coloboma/congenital heart disease	7
Coloboma/choanal atresia	4
Three features	
Mental retardation/short stature/micrognathia	31
Mental retardation/short stature/coloboma	4
Mental retardation/micrognathia/choanal atresia	1

a reasonable search will generally result. The usefulness of the approach used by the LDD is supported by recent data from clinical genetics services.

Additional features of the LDD include the ability to create the user's own patient database, a set of relevant references for each of the syndromes in the database, and a list of the features which have been found for each of the syndromes in the database. The patient database can be searched at the same time as the LDD database enabling the user to identify and/or diagnose a syndrome seen previously and also allowing different users to pool data (anonymously). Several previously unreported syndromes have apparently been synthesised as a result of this latter capability.⁷ Future versions will apparently include databases which contain chromosomal disorders (R M Winter, 1988, personal communication).

POSSUM is another SD which includes some of the same types of facilities as the LDD, including the recent addition of some of the same database, but has added the capability of showing full colour images obtained taken from photographs on a video monitor.⁸ The images are stored on an optical disk storage medium and scanned by semiconductor laser to recover the data. The images and data are prerecorded and updated regularly, but the user cannot add his/her own patient photographs to the system at this time.

Since the LDD accepts only a limited set of codes, each corresponding to a fixed nomenclature, a standardisation between the search criteria and the tabular data has been achieved. This standard comes at the cost of forcing users of the LDD to conform to a rigid, albeit reasonable, hierarchical nomenclature which represents the authors' well reasoned approach to this problem. Unfortunately, the lack of international agreement on nomenclature and coding of morphological abnormalities makes it unlikely that the LDD, or any of the other similar programs will, in their current forms, be more than interim steps to a long term solution.

Another set of problems which the current SD faces has to do with quantification and imprecision. As noted above, the accountant's spreadsheet or table usually contains more than 'yes/no' in an individual entry. For example, feature X is generally thought to occur in syndrome ABC about Z% of the time. The LDD and other SDs do not make use of this knowledge in any logical way and in essence record the fact 'present'. Conversely, the obscure anomaly associated infrequently with a given syndrome may be excluded from the database, because when using the all or none approach, inclusion could be misleading. A branch of logic called 'fuzzy logic'

has developed to deal with this kind of situation. Although there is still some disagreement about whether fuzzy logic is fundamentally different from using traditional logic and adding statistical methods, in either case the difficulty of dealing with imprecision and uncertainty must be addressed. This can be done to a certain extent, empirically, using existing data for the frequency of occurrence of clinical features for many syndromes. Thus, if a given feature is found in 90% of reported cases of the XYZ syndrome, the absence of that feature in a given patient would diminish, but not obliterate, the belief that the patient had that syndrome. On the other hand, if a patient was found to have a number of features which were frequently found in a disorder, the confidence in that diagnosis would be significantly enhanced. Note that this system still allows for a 100% entry which would be taken to mean that the absence of that feature excludes the syndrome. A more refined approach would use statistical correlations for each feature found in each syndrome. This would, however, be very time consuming because of the need to develop the necessary data, but obviously should lead to more accurate assessments of diagnostic probabilities. A more feasible solution to the problem of uncertainty and imprecision is the use of the artificial intelligence or expert system methods described below. The marriage of traditional database techniques with the more modern expert system methodology is a promising technology sometimes known as the intelligent database interface.

The SD of the (near) future should also include a method to translate the user's non-standard terminology into the standard used by the system. Ideally this would allow input such as:

"Find all syndromes with macrocephaly and intestinal polyps."

"Which craniosynostosis syndromes do not have hand anomalies?"

"What is the most likely syndrome with choanal atresia and coloboma?"

The SD would translate these input queries into the appropriate standard coding nomenclature, search the database using an indexing system, analyse the data using an expert system-like approach, and finally provide a natural language response, if appropriate, or a table or graphic response if better suited to the query.

Searching medical literature

When the dysmorphologist has identified one or more candidate diagnoses, or has been unable to find a satisfactory one, examining appropriate articles published in medical journals may be quite

rewarding. Computerised medical literature searching has been available for many years, most notably in the form of the Medline search of the National Library of Medicine (NLM) database. Although this approach has been of enormous value in the identification of relevant medical publications, the deficiencies for dysmorphology are substantial. The Medline search is typically done using medical subject headings (MESH) which are descriptive categories assigned to each article by professional staff at NLM. A problem with this approach is the lack of correspondence between the MESH terms and the types of categorisations generally used by dysmorphologists. This leads to a failure to identify relevant articles and the inclusion of irrelevant articles. Searches done using keywords rather than MESH terms sometimes work better in this regard for syndromes relevant to dysmorphology, but this approach still leaves something to be desired, since the keywords accessible to the dysmorphologist are those identified by the author. Typically this will be a limited subset of the actual words used in the article, limiting the breadth of search capability. As improvements in computer and data storage techniques continue, it will become feasible to have the entire text of articles available for search. This has already happened on a limited basis.

A better system for searching medical literature would include the facility to accept natural language queries, have some understanding of medical semantics, and also have access to the entire text of articles, abstracts, and books. Such a system would accept a query such as "Find all articles which describe patients having preaxial limb anomalies, chromosome deletions, and congenital heart disease" and return a substantially complete set of relevant articles. This kind of functionality would result from the application of artificial intelligence techniques to extensive literature databases. The query would be parsed by the system into an appropriate format and then each phrase would be expanded into a set of alternative terminologies, which would then be used to search the extensively indexed text. This would be a highly appropriate use of the expert system DCT discussed below. The availability of this kind of tool, along with the SD systems described above, should lead to more rapid and accurate diagnoses of patients with morphological anomalies.

Expert systems

Expert systems (sometimes also called knowledge or rule based systems) are attempts to capture the knowledge of an expert in a specific domain by analysing the expert's approach to problem solving in that domain.⁹ Although expert systems are

generally considered to be a part of the area of computer science known as artificial intelligence, the approach to the development of expert systems has become rather straightforward, and they are in practical use in fields ranging from pulmonary function to credit card authorisation.¹⁰

To gain some idea of how expert systems work, consider the simple problem of classifying animals. The 'expert' in this case may be a biologist and the users of the system government workers preparing inventories of natural regions. The expert can walk through the fields and identify animals by sight but can only be in one place at a time. His less well trained assistants can cover a larger geographical area but do not have his knowledge. Although the expert can identify animals by sight, the decisions made must have some formal basis. In this example, it is obvious that a set of rules can be formulated to codify the decision process. For example: If the animal has fur or hair then it is a mammal. If the animal has feathers then it is a bird. If a bird hunts animals at night and can turn its head 180° in either direction then it is an owl. If an owl is white and lives in the Arctic above the tree line then it is a snowy owl. These rules may or may not be immediately evident to the expert who sees a white bird in the Arctic and immediately knows that it is a snowy owl rather than an ivory gull. If there were no general agreement on the differentiation between the two birds then an expert system would be of little help to the novice, an important issue for expert systems. Expert systems work best in domains where there is general agreement about knowledge in the domain.

In this particular problem, the rules used by the expert in making a decision must be formalised. The role of analysing the expert's decision processes and developing formal rules is played by the 'knowledge engineer'. In some situations the rules may be immediately apparent to both the expert and the knowledge engineer, but in others the formal processes of decision making may be difficult to establish. Once the rules have been identified, the knowledge engineer uses these data to create a computer program which embodies them. The expert then assists the knowledge engineer in testing the program to establish logical consistency and identify any changes needed in the formal rules. This process may be iterated many times until the rules, system performance, and expert performance are self-consistent. During the development phase the expert may discover or gain insight into new or previously unrecognised relationships in his/her knowledge base.

A major problem in the design of expert systems is the way to manage uncertainty. What if the rule is true only 80% of the time? If the assumption is

made that a given feature must be present for the rule to be applicable, we will be missing the correct classification 20% of the time. When multiple rules are used this effect may be exacerbated. By adding a probability to each rule, this problem can be circumvented to a considerable extent. Several different approaches have been attempted, such as certainty factors and fuzzy logic.^{9 11} In many cases these apparently disparate methods can be shown to be mathematically equivalent.

The alert reader may notice some similarity between the information contained in the relational database and the expert system. A rule in the expert system has a correspondence with a row in the database. For this reason, the marriage of database systems with expert systems has become an area of considerable interest in the past few years.¹²

Some important considerations in the design of expert systems are related to the type of knowledge needed to solve problems in a given domain. Obviously the expert must, with the help of the knowledge engineer, be able to explain the nature of the 'rules' used in the decision making process. If the knowledge base in the expert domain is largely ad hoc, as may be found in newer areas of medicine, the possibility of generating a self-consistent set of rules is greatly diminished. Note the difference between a self-consistent set of rules with uncertainty and a set of inconsistent rules. In the former case the logic of making decisions is reasonably clear but, as in most of medicine, biological variability causes uncertainty. Thus, in the case of traffic law, it is possible to say that all vehicles travelling faster than 100 km per hour are violating the law (in the USA) with 100% certainty, whereas in medicine it is not possible to say that all persons with blood glucose levels higher than 150 mg/dl have diabetes mellitus. A rule of the form 'adults with fasting blood glucose levels higher than X mg/dl two or more hours after last eating have a Y% chance of having type 2 diabetes mellitus' would be likely to be quite satisfactory in an expert system dealing with glucose metabolism and its disorders. When the knowledge base is such that different experts give a wide range of answers to the same question, however, the possibility of creating a satisfactory rule based system is greatly diminished.

The facility with which expert systems generally manage pattern classification problems may make them highly desirable tools in the field of dysmorphology, where syndrome identification is a major task and clearly represents a highly specialised form of pattern classification.¹³ The amount of time expert dysmorphologists spend assisting others in classifying syndromes is substantial and a continual commitment. In some cases creation of an expert

system may result in a very effective problem solving tool. Since the expert system has such a rich body of data contained in the rule structure, it has been possible to use the program for teaching purposes by adding an explanation module. Computer generated graphics and prerecorded images may make this kind of system particularly effective.

The development of expert systems may have another useful fringe benefit. During the development of a working expert system the knowledge engineer forces the expert to articulate a set of rules for a given knowledge domain. This experience may be painful for both the expert and the engineer, because the expert frequently makes decisions using rules which are not used consciously. By forcing the expert to develop a conscious representation of the decision making process, the knowledge engineer may cause the expert to develop additional insight into the domain of knowledge. This has been observed in our laboratory during the recent development of an expert system for neurofibromatosis (V Riccardi, 1987, personal communication).

When properly conceived and executed, expert systems can be highly efficient tools which improve productivity and accuracy of tasks done by the users. Indiscriminate use of this technology will result in considerable waste of resources and unwarranted dissatisfaction with this promising application of DCT.

Image enhancement and reconstruction

An area that has just begun to be applied to clinical medicine is digital image enhancement. An image is converted to a digital array which can be modified by the computer to improve the ability of the observer to identify features of interest.¹⁴ No new data are actually created, but the alteration of image intensities makes it easier for the human eye to pick out boundaries between internal structures. Systems which allow cytogeneticists to enhance chromosome images digitally are currently in use, although secondarily related to dysmorphology. Similar techniques have been applied to radiography on a research basis and would appear to offer some promise of improving morphological analysis for the skeleton, particularly in the fetus and younger child because of the low levels of mineralisation.

In addition to enhancing images, DCT offers the capability of generating representations of images which were previously unavailable. The most interesting current application is the reconstruction of 3D images from 2D cross sections found in CT and MRI scans.¹⁵ In the initial employment of this technique, the 2D contour of interest was manually outlined for each section using a graphics digitising tablet. After the successive sections were all input to

the computer a 3D image, which could be manipulated or transformed or both, was reconstructed using geometrical algorithms. Recently the contouring of the individual sections has been automated to some extent and the entire process can be performed by the computer. Obviously, the ability to examine the shape of internal structures in a 3D representation has considerable interest for the dysmorphologist. Unfortunately this type of reconstruction is computationally intensive and still largely relegated to research laboratories. Reconstructions done from user derived tracings from CT and MRI scans are much simpler to process and within the reach of most clinical genetics units from the standpoint of computational requirements. Although the tracing of individual contours is rather tedious and time consuming, for selected cases the ability to see a more intuitive 3D representation on a monitor would appear greatly to enhance the ability of the dysmorphologist to compare morphological relationships between features in different malformation syndromes. This is especially true if the image is of sufficiently high resolution to look realistic, and also if it can be shifted, rotated, and expanded on the screen.

Computer analysis of biological shape

The publication of D'Arcy Thompson's classic text *On growth and form* marks the beginning of what might be termed the formal analysis of biological shape.¹⁶ Although Thompson considered shape from what was essentially an evolutionary perspective, his methods have been adapted for application to dysmorphology. Before discussing the formal analysis of biological shape, examination of motivation is important. Viewed from the perspective of the evolutionist, the formal analysis of shape is important in understanding how the shape of an ancestral species was transformed into the shape of a more modern one. The emphasis would appear to be on understanding how the transformation took place. Similarly the embryologist might also benefit from formal methods of shape analysis which emphasise transformations. Used prospectively, transformation analysis is also applicable in the field of reconstructive surgery, since the ability to study the expected outcome of surgical procedures in a model is obviously advantageous. Note that the dysmorphologist has a different motivation, in general, for needing formal shape analysis. It is not the transformation of one shape to another which must be understood, but rather the difference between the two shapes. This point may appear to be a subtle, or merely semantic, difference on first impression, but is an important distinction when considering the design of a DCT tool for dysmorphology. The

reconstructive surgeon, evolutionary biologist, teratologist, or embryologist can benefit from a tool which clarifies the physical transformations necessary to move from one shape to another. The emphasis is on the nature of the transformation, not necessarily on the beginning and ending shapes. The dysmorphologist, on the other hand, is most concerned from a diagnostic standpoint with being able to classify the shapes, not transform one into the other. As a result of this difference in approaches it is likely that the most effective methods of solving these problems may also differ considerably.

Consider, for example, a lateral view of the nose. In some persons, the 2D curve along the superior aspect from near the root to near the tip is essentially a straight line. The mathematical representation of the 'curve' is therefore $y=mx+b$, the formula for a straight line in Cartesian coordinates. In some syndromes this same region has a scooped out appearance or, in more mathematical terms, an upward convexity. Since this is a relatively short segment it is likely that an analytical formula might be found to represent the curve. This might be possible with a quadratic curve, a portion of a conic section such as a circle, ellipse, hyperbola, or parabola, but certainly can be represented by a cubic curve to any required degree of accuracy. The transformations necessary to change the straight line representation of the nose into the quadratic or cubic form are beyond the scope of this paper, but well defined and applicable to segments of real biological structures in two and three dimensions.¹⁷ Bookstein^{18,19} has written extensively on analysis of shape from a transformational viewpoint. Starting from the work of Thompson he has added modern statistical and geometrical methods which appear to provide a firm theoretical basis for analysing shape change. The application of these principles to reconstructive surgery has been successfully demonstrated.

When functioning in a diagnostic role the dysmorphologist often is called upon to compare biological shapes. This comparison usually involves a living patient and a memory of what the shape of a corresponding structure was like in one or more patients with a specific diagnosis. It may be further complicated by the need to compare the 3D image in vivo and the 2D image (from which the dysmorphologist must partially visualise a nearly 3D image) in a photograph. In this setting, whether the shapes are compared using analytical mathematical formulae or descriptive approaches also amenable to computerisation would appear to matter little to the dysmorphologist, but the latter method may lead to a more satisfactory application of DCT.

We have recently begun exploring the possibility of creating a tool to be used by the dysmorphologist

in the clinical evaluation of biological shape using the principles outlined above.^{20,21} It is possible to obtain three dimensional digital representations of the head using stereo photography. By segmenting the complete 3D contours into regions of biological interest such as nose, ear, etc, and then subsegmenting one further level, for example, the tip of the nose, it should prove feasible to subdivide the initially overwhelming computational task into smaller, more tractable comparisons. Adding the rule based approach of the expert system along with a library of standard shapes from control populations and patients with known syndromes would result in a tool with which the dysmorphologist would use a high resolution video display to compare the actual shape of the bridge of the nose in a new patient with the corresponding shape in a patient with a known syndrome. No longer would the decision regarding a flat nasal bridge be subjective; the diagnostic decision would be completely objective and, to some extent, quantifiable.

Dysmorphology/clinical genetics office management system

Integration of the various tasks performed by the dysmorphologist and supporting staff could be accomplished through the use of a clinical genetics office management system (CGOMS). The CGOMS would operate on a multi-user computer system accessible from directly connected terminals and also from remote terminals communicating over telephone lines. The essential components of such a system are outlined in table 6. As can be seen in the table, all data regarding dysmorphology or clinical genetics patients would be managed using DCT resulting in a 'paper-free' environment. (For a variety of reasons, including legal requirements, paper records would continue to be available as an adjunct back-up system.) In hospitals and clinics which already have some degree of computerisation, the CGOMS could perhaps be implemented on existing hardware and overlapping data such as name, address, and other demographic and billing

TABLE 6 *CGOMS facilities.*

Patient schedule module
Follow up reminder module
Patient database
(a) History
(b) Physical examination
(c) Reports
(d) Correspondence
(e) Laboratory values
(f) Images 2D/3D
Syndrome database module
Literature search dysmorphology expert system
Telecommunications capability
Electronic mail
Word processing capability

data centralised. The sharing of data could also be accomplished in a networked system with CGOMS residing on its own computer, which was then connected to the main hospital computer.

CGOMS would have a complete record of each patient including demographic and billing data, the full clinical report, correspondence, scheduling and follow-up capabilities, an SD including photographs and perhaps karyotypes and radiographs, electronic mail, and word processing facilities. In addition to being able to search the SD using indexed features as described above, CGOMS should also allow the user to access and search full text descriptions of patients. In this way more subtle or less easily described features, which would not ordinarily be indexed in the SD, would be retrievable (although probably at a much slower rate because the full text might not be indexed the way the SD was).

It is important to note that CGOMS would benefit greatly from nomenclature and coding methodology which is far more standardised than that found in LDD or POSSUM. Recently, a variation of the ICD 9 has been proposed by the British Paediatric Association and modified slightly by the Center for Disease Control, Atlanta, Georgia.²² Adherence to such a standard, or a similar one, would greatly facilitate the use and interchange of dysmorphology data, and presumably result in more rapid identification of previously unreported syndromes and faster diagnosis of patients with very rare disorders. The existence of a standard on the level of data interchange would not preclude individual centres from developing additional levels of data collection protocols which could be layered on top of a lower level standard. If, at some future time, a representation of 3D contours of the face were deemed important enough to save in a database, it would be feasible to experiment with a new set of ideas about storing these data. A well designed system would permit the addition of a resulting new 3D data standard gracefully, without the need to make major changes because of the low level standards.

Conclusions

The amount of power available on small computer systems should continue to increase on an accelerated basis for the next decade or two. During that time DCT has the capability to change the practice of medicine, and particularly dysmorphology, to an enormous extent. Proper use of DCT in almost any domain, however, demands considerable imagination, careful analysis of any proposed applications, and adherence to the basic principles of computer science. Currently, little interchange between dysmorphologists and computer science departments appears to be taking place. If dysmorphologists

become convinced that DCT has more to offer than word processing and routine database applications, and can convince the computer scientists that an exciting set of solvable problems is available, a fruitful cross-fertilisation will take place, resulting in the creation of powerful tools which will benefit clinicians, researchers, and, most importantly, patients.

References

- Sanders DH. *Computers today*. 2nd ed. New York: McGraw-Hill, 1985.
- Brightman RW, Dimsdale JM. *Using computers in an information age*. Albany, NY: Delmar, 1986.
- Rubinstein R, Hersch HM. *The human factor: designing computer systems for people*. Bedford, MA: Digital Press, 1984.
- Martin D. *Advanced database techniques*. Cambridge, MA: The MIT Press, 1986.
- Patton MA. A computerised approach to dysmorphology. *MD Computing* 1987;4:33-9.
- Castro L, Hanson J, Rettig T. *Advanced programmers guide: featuring dBASE III and dBASE II*. Culver City, CA: Ashton-Tate, 1985.
- Flannery DB, Peterson SL. Practice analysis of dysmorphology diagnostic database. *Am J Hum Genet* 1987;41:59A.
- POSSUM: pictures of standard syndromes and undiagnosed malformations. Computer Power Pty Ltd. Applied Research and Development, 613 St Kilda Road, Melbourne, Victoria 3004, Australia.
- Buchanan BG, Shortliffe EH. *Rule-based expert systems*. Reading, MA: Addison-Wesley, 1984.
- Hayes-Roth F, Waterman DA, Lenat DB, eds. *Building expert systems*. Reading, MA: Addison-Wesley, 1983.
- Negoita CV. *Expert systems and fuzzy systems*. Menlo Park, CA: Benjamin/Cummings, 1985.
- Kerschberg L, ed. *Expert database systems*. Menlo Park, CA: Benjamin/Cummings, 1986.
- Rich E. *Artificial intelligence*. New York: McGraw-Hill, 1983: 368-74.
- Andrews HC, Hunt BR. *Digital image restoration*. Englewood Cliffs, NJ: Prentice-Hall, 1977.
- Davis RS, Jack DB. A surgical cut into computers. *MD Computing* 1986;3:21-8.
- Thompson DW. *On growth and form*. Cambridge: Cambridge University Press, 1961.
- Foley JD, Van Dam A. *Fundamentals of computer graphics*. Reading, MA: Addison-Wesley, 1982.
- Bookstein FL. *The measurement of biological shape and shape change*. Lecture Notes in Biomathematics Vol 24. Berlin: Springer-Verlag, 1978.
- Bookstein FL. The geometry of craniofacial growth invariants. *Am J Orthod* 1983;3:221-34.
- DiLiberti JH. Computerized analysis of biological shape: applications to dysmorphology. David W Smith Morphogenesis Workshop, Burlington, Vermont, July 1986. *Proc Greenwood Genetic Center* 1987;6:132-3.
- DiLiberti JH, Storck L. Three-dimensional contours of the human face obtained using a microcomputer-based stereometric imaging system. David W Smith Morphogenesis Workshop, Greenville, South Carolina, August 1987. *Proc Greenwood Genetic Center* (in press).
- Birth Defects Branch Six Digit Code. Based on BPA Classification of Diseases (1979) and WHO ICD (1979). Birth Defects Branch, Center for Environmental Health, Centers for Disease Control, Atlanta, Georgia 30333, USA.

Correspondence and requests for reprints to Dr J H DiLiberti, Department of Pediatrics, Emanuel Hospital and Health Center, 2801 N Gantenbein Avenue, Portland, Oregon 97227, USA.