(2)

AAMRL-SR-90-500

AN INVESTIGATION INTO TECHNIQUES FOR LANDMARK IDENTIFICATION ON 3D IMAGES OF HUMAN SUBJECTS

RANDY BETH POLLOCK, PH.D.

WRIGHT STATE UNIVERSITY

AUGUST 1989

PHASE I REPORT FOR JUNE 1989 TO AUGUST 1989

Approved for public release; distribution is unlimited.

ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY
HUMAN SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433-6573

NOTICES

TECHNICAL REVIEW AND APPROVAL

AAMRL-SR-90-500

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

CHARLES BATES, JR.
Director, Human Engineering Division
Armstrong Aerospace Medical Research Laboratory

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AAMRL-SR-90-500
6a. NAME OF PERFORMING ORGANIZATION Universal Energy Systems Inc 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Harry G. Armstrong Aerospace Medical Research Laboratory
6c. ADDRESS (City, State, and ZIP Code) 4401 Dayton-Xenia Road Beavercreek, Ohio 45432	7b. ADDRESS (City, State, and ZIP Code) Wright-Patterson Air Force Base OH 45433-6573
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620-88-C-0053
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO: 62202F PROJECT NO.: 7184 TASK NO: WORK UNIT ACCESSION NO.:

11. TITLE (Include Security Classification)

An Investigation into Techniques for Landmark Identification on 3D Images of Human Subjects (U)

12. PERSONAL AUTHOR(S)

Pollock, Randy Beth Ph.D.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Phase I	FROM 15Jun88 TO 30Aug89	1989 August	19

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Several approaches to landmark identification using artificial intelligence methodologies were investigated. A blackboard architecture was identified as the most promising approach. Several prototype modules were developed to generate and evaluate hypotheses regarding the positions of the landmarks on 3-D digital images. These were tested on a sample of 20 subjects and the results were sufficiently encouraging to warrant further work. The design of a blackboard system was begun, including components utilizing other technologies such as neural networks and constraint networks.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION	
[X] UNCLASSIFIED/UNLIMITED [ ] SAME AS RPT [ ] DTIC USERS	UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
KATHLEEN M. ROBINETTE	(513) 255-8810	AAMRL/HEG

DD Form 1473, JUN 86 — Previous editions are obsolete. — SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

Acknowledgements

I would like to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this work. I would also like to thank Universal Energy Systems for their efficient and concerned administration of the program.

I was fortunate to work with a group of scientists that made this a stimulating and enjoyable experience. Overall direction was ably provided by Kathleen Robinette and Gregory Zehner, who also gave me the opportunity to see the tribulations and rewards of applying anthropological principles and data in practical applications. Dr. Robert Beecher (Beecher Research) and Joyce Robinson (SRL) gave freely of their time and expertise to help with technical matters. Sherri Blackwell and Phil Walker (Anthropology Research Projects) made the lab a congenial environment in which to work and gave me insight into demanding nature of gathering and manipulating the data. I would like to thank all of these people for their help and support throughout this project.

I. INTRODUCTION

Anthropometric data have been used by the Air Force for many years to help in the design of clothing, equipment, cockpits, etc. The use of anthropometric statistics enables designers to provide a better-fitting product and ensure that the people who will be using the equipment are physically able to do so.

Traditionally, the tools of anthropometry have been physical measuring devices such as calipers and tape measures. Recently, the Human Engineering Division at Wright-Patterson Air Force Base has been exploring the use of 3-dimensional scanning technology to produce a digital representation of the surface of the human head. A vertical stripe of laser light is projected onto a stationary subject as the scanner rotates 360 degrees around him or her, taking video recordings which are converted into digital format. The digitized image is stored as a set of points representing the distance from the center of rotation to the surface of the subject at 256 points along each longitude. This technique has been used to collect data on both men and women to be used in the design of such equipment as night vision goggles for Air Force pilots.

Although the data collected in this way are much more complete than was possible using traditional methods, techniques have not yet been developed to easily analyze and use this type of data. Due to the large size and noisiness of the data set, and the difficulty of recognizing facial features which vary greatly in size and shape among individuals, the use of artificial intelligence methodologies seems appropriate to the problem.

My academic background incorporates two specialties that are directly relevant to the problem described above. My Ph.D. is in physical anthropology, which is the academic discipline that includes the study of anthropometry. More recently, I received a M.S. degree in computer science, with a specialty in artificial intelligence. I have been involved in AI research for the past several years. These factors led to my assignment to the AAMRL/Human Engineering Division for the summer research program.

II. OBJECTIVES OF THE RESEARCH EFFORT

One major problem already encountered in the use of 3-D data is the identification of facial landmarks on the image. The landmarks are used to define boundaries between different regions of the head, and to serve as endpoints for distance measurements. There are approximately 40 landmarks that must be recorded for each image and stored as a set of longitude/latitude coordinates.

Currently, about half the landmarks are being marked on the subjects prior to scanning using small felt patches which do not reflect the laser light. The remaining landmarks are those which can be visually identified on the image, such as the tip of the nose, or those which cannot be marked with patches, such as the pupil or the stomion (center of the mouth along the meeting line between the two lips). Someone must then record the longitude and latitude of each landmark in a separate file associated with the image file. This is done by using the cursor to locate each landmark on a graphic representation of the head and instructing the program to record the latitude and longitude of the point identified by the cursor.

A means of having the landmarks automatically identified by an AI program would save a great deal of time for both the technicians who currently identify the landmarks manually, and for the subjects being scanned. The goal would be to eliminate the need for marking subjects with patches, or to reduce those to as few as possible, and to have a computer program identify and record the coordinates of each landmark. In addition to saving time, computerizing this process would ensure consistent landmark identification across subjects, which is especially difficult when more than one person does the marking.

My goals for the 1989 Summer Faculty Research Program were (1) to survey the scientific literature for techniques that might be applicable to the problem described above, and (2) to test the most promising approaches, and (3) if any approach appeared to be successful, to propose a general design for a system that would do the described tasks.

III.

A survey of the computing literature revealed several possible approaches to the problem. There is a large literature on computer vision, but most work in that field is inappropriate to this domain for one of the following reasons: (1) It deals with issues of creating 3-dimensional models from 2-dimensional images. In this research, the 3-D image is the initial form of the data and therefore those issues are not relevant. (2) Much of the computer vision work takes a highly mathematical approach to analyzing curves, surfaces, etc. For the landmark identification problem, a knowledge-based approach using contextual information about the face and relevant features appears to be more useful. This approach is found in the work by Kjeldsen, Califano and Bolle (1989) on object recognition in a 3-D framework.

The most useful areas in the scientific literature are in the categories of pattern recognition and artificial intelligence. The neural network approach, as described by Pao (1989), has potential for contributing to the landmark identification problem, but there are some problems to be overcome. First, the large size of the data set for even one image (a 512 X 256 array, or 131,072 points) requires an extremely large network of neurons for processing, and therefore would require large measures of both time and memory. Second, the problem at hand would seem to require two levels of network processing: a first pass might identify facial features, and a second pass could be used to locate landmarks in relation to those features. Third, it difficult to tell whether a neural network approach is capable of solving this problem without building the entire system.

The approach that appears to be the most promising is a knowledge-based system using a blackboard architecture. This type of system has been applied to symbol-processing problems that are in some ways analogous to landmark identification. One major advantage of the blackboard approach is the ease with which various types of data may be integrated into the system, and the use of neural network technology within a blackboard system is a viable alternative.

Blackboard systems are characterized by more sophisticated control structures and data representation facilities than are found in more traditional expert systems. They usually consider a problem from multiple levels of abstraction, with modules at one level evaluating hypotheses generated by modules at other levels and generating their own hypotheses in turn. Blackboard systems usually use 'opportunistic' reasoning, which gives them the ability to vary the problem solving algorithm according to the way in which the solution is developing. They are highly flexible systems and have been used to solve such problems as speech recognition (Erman, et.al. 1980), sonar signal analysis, (Nii, et.al. 1982), and analysis of aerial photographs (Nagao, et.al. 1979). A good overview of the blackboard architecture and applications may be found in Engelmore and Morgan (1988).

IV.

A blackboard system will work only if there are modules capable of generating reasonable hypotheses about the data. The feasibility of the blackboard approach by was tested this summer by the development of several such modules.

The first module looks at points on individual longitudes in an attempt to locate facial features. These points are the radii (distance from the center of rotation) that have been measured at each latitude. Traversing a single longitude, one can discern local maxima and minima in the radii which represent protrusions and indendations respectively. For this purpose, a local minimum is a point whose neighbors both have larger values; a local maximum is a point whose neighbors both have smaller values.

Facial features that show up along a longitude as local maxima include the chin, nose, upper and lower lips, etc. Features that show up as minima include the mouth (which appears as an indentation between the two lips), the eyes, and the center of the ear. In each of these cases, the local maximum or minimum occurs across a number of longitudes at approximately the same latitude.

However, due to errors introduced by the imaging and digitization process, the maximum or minimum typically does not show up on all stripes. An additional complication is that most features are not horizontal lines but have a curved or angled shape and show up at gradually changing latitudes. For example, following the line of the chin from the center to either side will take you to from lower to higher latitudes.

These problems were successfully solved in Module A which is capable of recognizing a feature as a pattern of local maxima or minima that can be followed across a set of adjacent longitudes. The features discovered by this module have no semantic information associated with them. That is, Module A does not recognize that Feature X is a nose or a chin. Many of the features found in this way do not, in fact, represent facial features that are being sought. The ear, for example, has a large number of maxima and minima that this approach identifies as features, while they are really part of one larger feature. The fact that some meaningful features are recognizable, however, is a major step toward landmark identification. One way to locate the promenton (most protrusive point on the chin) is to identify the chin as a gross feature and take the center point. This technique applies to a number of other landmarks as well. The difficult part of the process is identifying which min/max feature is actually the chin.

Module B was developed to look at possible features identified by

Module A and generate hypotheses regarding which facial features they might represent, and to also generate hypotheses about the positions of landmarks on those features. In a full blackboard system, Module B would have access to information about the positions of human facial features relative to one another. In this prototype, that information is supplied by the user. When Module B generates 2 or 3 hypotheses about a particular landmark, the user chooses the correct hypothesis (if a correct hypothesis is generated), and that information is used by Module B in the generation of hypotheses for other landmarks. The results obtained using this prototype are given in Table 1.

Some landmarks are not associated with features that show up as local maxima or minima. For example, the sellion is the point on the central profile between the nose and the forehead where the radius is at a minimum. In many people, however, this is not a local minimum, but a flat-bottomed valley. It frequently does not show up as a minimum using the algorithm described above.

Another way to approach the data is to look at a the points on a longitude as a set of slopes computed as the differences between adjacent radii. One can then identify points at which the slope changes, which may help to identify landmarks. The sellion is an example of a landmark that is more easily identified in this way. A pattern can be seen in which sellion is clearly discernable as the slope changes between the nose and the forehead. Changes in

	Pro-nasale	Sub-nasale	Pro-menton	Menton	Right Chelion	Stomion	Left Chelion
s21	292/129	292/117	292/82	295/74	309/100	293/103	275/100
	293/129	291/117	294/83	294/73	314/101	292/104	269/102
s24	284/167	284/156	286/124	285/113	307/142	285/145	263/142
	284/167	284/157	286/125	286/115	322/136	286/145	250/137
s29	298/138	298/122	297/84	295/75	324/105	298/112	270/106
	300/138	300/122	294/85	294/75	335/104	300/110	265/107
s30	286/172	286/161	288/123	289/110	305/143	289/146	272/142
	286/171	287/162	288/123	288/113	315/142	287/148	259/141
s50	278/151	277/140	276/105	277/97	300/121	277/124	254/122
	278/151	277/142	274/105	274/95	304/121	275/126	247/122
s54	298/155	297/143	295/108	296/102	315/130	296/132	276/130
	298/157	302/143	298/112	298/102	337/126	302/132	267/127
s60	296/171	296/159	293/121	293/112	322/139	295/143	266/139
	296/172	295/161	287/123	287/113	334/135	295/144	256/138
s66	305/158	306/148	305/114	304/105	332/133	305/137	276/133
	307/159	307/149	304/114	304/104	339/131	303/139	268/133
s67	301/168	301/156	297/117	296/107	318/133	299/141	276/133
	301/168	300/160	296/119	296/109	325/126	297/141	269/135
s84	300/148	300/139	301/96	304/88	323/117	301/121	284/117
	300/148	300/138	302/97	302/87	328/115	300/121	272/113
s90	278/151	278/138	275/97	275/91	299/118	276/122	253/118
	278/150	276/135	273/99	273/89	317/115	276/122	236/115
s91	298/152	298/143	295/105	294/95	310/123	296/127	284/125
	298/152	296/144	294/105	294/95	318/120	296/127	274/123
s92	288/141	288/134	288/92	288/83	313/113	287/117	266/115
	288/142	287/118	288/79	288/69	323/111	287/118	251/113
s93	301/134	302/125	299/88	299/79	319/106	302/108	287/103
	301/134	300/125	300/89	300/79	322/105	300/109	279/106
s95	309/140	311/132	307/95	307/86	324/111	308/115	295/111
	310/141	310/132	309/95	309/85	329/108	308/115	287/111

Table 1. Recorded (top row) and computed (bottom row) landmarks for randomly chosen male subjects.

	Pro-nasale	Sub-nasale	Pro-menton	Menton	Right Chelion	Stomion	Left Chelion
s99	281/139	282/130	281/95	280/84	304/108	281/113	257/108
	281/139	282/128	275/92	275/82	320/104	284/113	248/106
s133	277/102	277/98	278/61	277/53	295/82	277/85	263/82
	277/106	277/135	277/63	277/53	299/80	277/85	256/82
s160	298/130	298/122	296/85	297/76	321/101	297/105	275/103
	298/130	298/123	293/87	293/77	324/99	298/105	273/102
s176	306/126	306/119	306/84	306/74	322/99	306/104	292/100
	306/127	306/116	307/85	307/75	324/97	306/102	288/99
s199	300/134	300/122	301/78	302/72	317/102	300/106	289/103
	299/133	299/126	303/83	303/73	318/104	299/108	281/104

Table 1. (continued)

slope are useful in identifying a number of other landmarks as well, such as the promenton, subnasale and pronasale.

Module C tests this approach to analyzing the data. Given the longitude representing a central profile, it is able to identify the sellion, pronasale, subnasale and promenton. Since it is possible to generate hypotheses regarding the latter three landmarks by both modules B and C, a module operating at a higher level of analysis in a complete blackboard system would be able to check for consistency and therefore evaluate the hypotheses with more confidence.

V. RECOMMENDATIONS

The prototyping effort was judged to be successful in demonstrating that modules can be developed to generate reasonable hypotheses regarding the positions of a select group of landmarks on a sample of 20 subjects in one pose. Although these results are narrow in scope, they are sufficiently positive to justify further work on this project.

The design of a blackboard system should be undertaken to implement the ideas described in this report. In addition to the modules implemented in this work, other techniques that can be integrated into the blackboard system include the following:

a. Knowing the approximate longitude of the center profile improves the performance of the existing modules. A module should be developed to identify this region with some certainty before attempting to find any landmarks. This module may be implemented as a neural network system which has been trained on the pattern of slopes on central profile longitudes. The longitudes of a new image can then be run through the network, and the one most resembling the previously seen central profiles can be identified. It will be necessary to test this technique in order to determine if it alone is sufficient to identify the central profile, or if additional analysis is necessary to confirm its results.

b. Analysis of the relationships among the landmarks should be undertaken to determine if there are landmarks whose positions can be identified by reference to other landmarks. For example, in the sample of subjects used in the prototype testing, the program was able to locate the menton to within 3 latitudes simply by subtracting 10 from the latitude of the promenton. If this relationship holds for a larger sample, it will greatly simplify the process of locating the menton. Testing should be done to ascertain if similar relationships exist among other sets of landmarks.

c. A constraint network should be incorporated into the system which contains the positional relationships among facial features and landmarks. This can be used to identify landmark locations as described in (b). However, it will also be valuable in the evaluation of hypotheses. If, for example, the central profile longitude has been identified with high certainty, a hypothesis which places the ectocanthus (outer corner of the eye) along that longitude should be quickly rejected because its position does not satisfy the constraints of these relationships.

d. A method for handling uncertain knowledge should be incorporated into the system. This should probably take the form of certainty or confidence factors as used in many expert systems. However, other methods for handling uncertainty may be investigated.

e. For those landmarks whose positions can be determined only by palpation and marked with a felt patch, a module should be implemented to recognize the voids in the data set which represents those landmarks.

REFERENCES

Engelmore, R. and T. Morgan, Blackboard Systems. Addison-Wesley Publishing, 1988,

Erman, L.D., et.al., THe Hearsay-II Speech Understanding System: Integrating Knowledge to Resolve Uncertainty, ACM Computing Surveys 12:213-253, 1980.

Kjeldsen, R., A. Califano and R. Bolle, Evidence Integration for 3D Object Recordnition: A Connectionist Framework, Proceedings of the 5th Conference on AI Applications, pp.55-63, 1989.

Nagao, M., T. Matsuyama and H. Mori, Structural Analysis of Complex Aerial Photographs, Proceedings of the IJCAI pp.610-616, 1979.

Nii, H.P., et.al., Signal-to-Symbol Transformation: HASP/SIAP Case Study, AI Magazine 3:23-35, 1982.

Pao, Y., Adaptive Pattern Recognition and Neural Networks. Addison-Wesley Publishing, 1989.