Biostereometric analysis for breast cancer detection

F. Proietti-Orlandi*, R.S. Varga[§], D.B. Sheffer, T.E. Price[†], and C.W. Loughry[‡]

Institute for Biomedical Engineering Research; [†]and Department of Mathematical Sciences, University of Akron, Ohio, USA; [‡]and Department of Surgery, Akron City Hospital, Akron, Ohio, USA: [§]Department of Mathematical Sciences, Kent State University, Kent, Ohio, USA

Received July 1987, accepted August 1987

ABSTRACT

A measurement technique has been developed for noninvasive breast cancer detection. The process involves the use of close-range stereophotogrammetry as a data acquisition device for the determination of breast surface concavities. We report the methodology used to detect these surface depressions, the rationale for the study, and our preliminary findings.

Keywords: Biostereometrics, stereophotogrammetry, breast cancer detection, convex hull

NOMENCLATURE

D	Digitized breast
$P_j^{(k)} := (X_j^{(k)}, Y_j^{(k)}, X_j^{(k)})$	Coordinates of D
N	Number of contour levels of D
N_k	Number of points for the kth contour of D
w_j	Points of an arbitrary polygon γ in the xy-plane
$\gamma, \tilde{\gamma}, \gamma_k$	Arbitrary polygons in the xy-plane, the kth contour polygon
ζ	Arbitrary set in the xy-plane
$\theta_j, \theta_j^{(1)}, \ldots, \theta_j^{(t)}$	Interior angle of γ formed by approximate line segments
π	Greek letter pi which represents the area of a circle of radius 1
$I(\gamma)$	Interior of γ
$E(\gamma)$	Exterior of γ
$C(\gamma)$	Convex hull of γ
$\cup$	Set union
$T := (t_1, \ldots, t_{n-1})^t$	A Boolean vector
Σ	Summation symbol
M, m_k	Centre of mass of a concave area of a polygon
r	Number of vertices in a concave area of a polygon
L_k	Number of concave areas for the kth contour of D
A	Set of all concave areas of D
$d(\cdot,\cdot)$	Distance function for the xy-plane
τ	Greek letter tau used to represent a threshold value
$d_1(\cdot), d_2(\cdot)$	Special distance function using contours
ϵ	An element of
L, ℓ, ℓ_p	Used to donate line segments
$\subset$	Set inclusion

INTRODUCTION

One of the major objectives of breast cancer screening is the development of a detection system which is both rapid with respect to patient time and provides little or no biological hazard to the individual; these objectives have been addressed in our research by use of biostereometric measurements. Biostereometrics, defined as a three-dimensional spatial analysis of biological form and function, encompasses a broad assortment of sensing or data collection equipment[1]. For the purpose of the research described in this paper, the stereometric data acquisition methodology was accomplished using close-range stereophotogrammetry[2], a technique that has been successfully used for many years in the aerial mapping industry and

*Present address: TTC and M Network Department, International Telecommunications Satellite Organization, Washington, DC, USA. Correspondence and reprint requests to Dr Daniel B. Sheffer, Akron City Hospital, 525 E. Market St, Akron, OH 44309, USA

0141-5425/88/030237-09 $03.00

has been shown to be efficacious in the documentation and quantification of complex topographic structures[3]. Sheffer *et al.*[4,5] have described and discussed the instrumentation and sensitivity of this particular application of close-range stereophotogrammetry.

Stereophotogrammetry was used in this study to obtain three-dimensional Cartesian coordinates representing the surface of the human breast in an arbitrary but fixed coordinate system. The data acquisition stage of the process permitted two metric quality photographs of the breast to be taken simultaneously. When the photographs were suitably oriented on a stereoplotting instrument, a sufficiently accurate three-dimensional model of the breast was produced. This optical stereomodel was then contoured in a manner which identified numerically data points representing the three-dimensional shape of the breast. A complete explanation of the data acquisition and reduction technique is given in Sheffer *et al.*[4,5].

The rationale for use of this technology for detection of breast cancer is based primarily on literature review and preliminary studies in our laboratory[1,4-8], which suggest that tumours within the breast may cause slight to moderate dimpling of the breast surface. Such concavities, if detected in a rapid and noninvasive manner, would certainly reduce the screening or pre-screening effort necessary in order to permit early treatment of the disease. Loughry and associates in 1980[6] reported the use of biostereometric analysis in a preliminary study of breast cancer. They visually examined the stereophotogrammetrically obtained contour map of the breast (*Figure 1*) for contours containing concave regions. Using this method to indicate potential underlying breast pathology, they were able to detect the location of lesions in eight of the ten preliminary cases studied; the tumours represented in this study ranged in size from 1 to 4.5 cm in diameter. Five of the eight cancers detected were not visible during the normal physical examination of the breast. Although it was recognized that the visualization of the contour maps was subjective, the results suggested that the localization of breast surface depressions could be a valuable aid in the detection of breast cancer.

Because the contour maps are essentially a geometric representation of a numerical data base, we deemed it more appropriate to analyse them objectively using mathematically based computer algorithms. This paper is designed to describe the algorithm and its mathematical basis, and to demonstrate the use of this method in a clinical research project. We include a discussion of a procedure to discriminate between those concavities which may truly represent a lesion in the breast and those representing data reduction 'noise'.

In the first section we introduce the concept of a digitized breast, i.e. a numerical data base modelling the subject's breast. The geometric interpretation of a digitized breast is a contour map which depicts, graphically, the shape and form of the breast. The effects of breast surface depressions on both the digitized breast and the level curves of the

Figure 1 Contour map of the breast.

contour map are presented. The main thrust of this section is the description of the geometry of a single contour and the location of depressions on contours not caused by noise. The remainder of this section is devoted to describing and identifying such concavities.

The second section is a description of the algorithm used to determine which concavities on consecutive contours form a pattern possibly caused by a suspect surface depression. This section also contains a description of the program used to execute the algorithms. Finally, we discuss the potential efficacy of the method as a prescreening device for breast cancer.

CONTOUR MAPS AND CONCAVITIES

Mathematically, a digitized breast is defined to be a set of Cartesian coordinates:

$$D := \{P_j^{(k)} := (X_j^{(k)},\ Y_j^{(k)},\ Z^{(k)}) :$$

$$j = 1,2,\ldots N_k,\ k = 0,1,\ldots,N\} \qquad (1)$$

where N, N_0, N_1, ..., N_k are positive integers ($N_0 = 1$). The points $\{P_j^{(k)}\}_{j=1}^{N_k}$ are the coordinates of a set of points lying in the plane $Z \equiv Z^{(k)}$. These coordinates are given with respect to a three-dimensional Cartesian coordinate system, with the units for each direction being 1 cm in length. The origin is symmetrically located between the subject's breasts, with the xy-plane parallel to the subject's

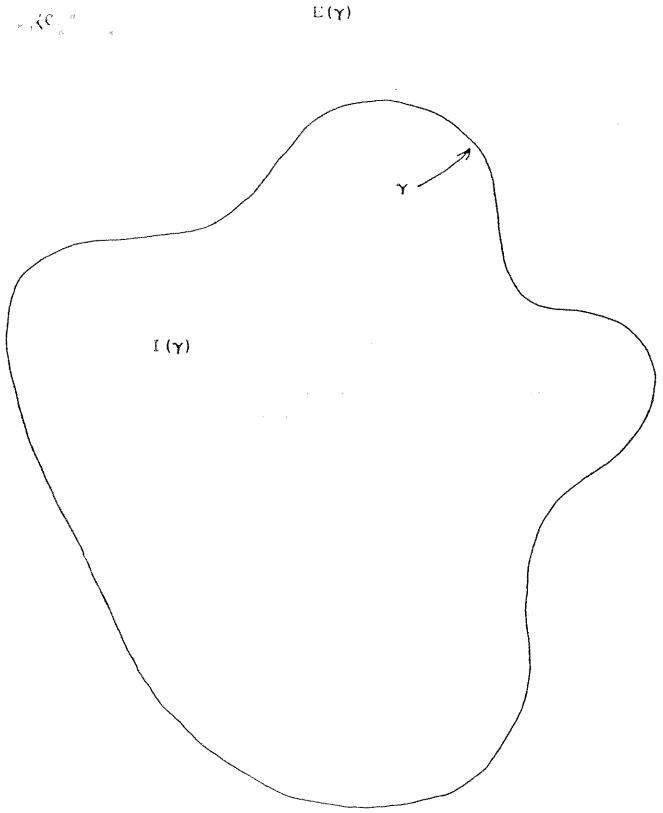

Figure 2 Simple closed curve. γ indicates the curve, $E(\gamma)$ is the region outside the curve and $I(\gamma)$ is the region inside the curve.

frontal plane. The positive Z-axis is oriented in an anterior (outward pointing) direction. The point $P_1^{(0)}$ corresponds to the most anterior point on the surface of the breast. For $k = 1,2,\ldots N$, the points $\{P_j\}_1^{N_k}$ correspond to points which lie on the surface of the breast intersected with a plane $Z \equiv Z^{(k)}$ parallel to the subject's frontal plane. From a stereoplotter operator's perspective, and hence a mathematical perspective, such an intersection would form a simple closed curve or contour (*Figure 2*). Recall that a simple closed curve is a continuous path which begins and ends at the same point, but does not otherwise cross itself. (For this reason it is convenient to assume that $P_{N_k}^{(k)} = P_1^{(k)}$, $1 \leq k \leq N$.) It can be shown[11] that such a curve, say γ, divides the plane into two connected components: the interior of γ, denoted by $I(\gamma)$, and the exterior of γ, denoted by $E(\gamma)$ (see *Figure 2*).

As previously mentioned, $P_1^{(0)}$ represents the most anterior part of the breast and usually corresponds to a point in the nipple-areolar region. The value $Z^{(1)}$ is chosen so that $Z^{(0)} - Z^{(1)} \leq 0.25$ cm and $Z^{(1)}$ is an integer multiple of 0.25 cm. The remaining Z-levels satisfy the relationship:

$$Z^{(k+1)} = Z^{(k)} - 0.25 \text{ cm}, \quad k = 1,2,\ldots, n-1 \qquad (2)$$

At each Z-level, the xy-coordinates of $P_j^{(k)}$, $j = 1,2,\ldots, N_k - 1$, are chosen by the plotter operator. (Recall that $P_{N_k}^{(k)}$ is set equal to $P_1^{(k)}$.) These points are obtained by the operator while traversing the optical model of the breast in a clockwise direction. It has been shown that this procedure is accurate and reliable[4,5,10].

As with many mathematical concepts, it is helpful and instructive to think of a digitized breast in graphical terms. The approach we used was to associate a polygon with each Z-level of equation 1 except, of course, for the single point $P_1^{(0)}$. More precisely, for a fixed but arbitrary k, $1 \leq k \leq N$, join the planar points $(X_j^{(k)}, Y_j^{(k)})$ and $(X_{j+1}^{(k)}, X_{j+1}^{(k)})$, $j = 1,2,\ldots, N_k$, obtained by projecting $P_j^{(k)}$, and $P_{j+1}^{(k)}$ onto the xy-plane, by a line segment. Since $P_{N_k}^{(k)} = P_1^{(k)}$, the path determined by the union of these line segments for $j = 1,2,\ldots N_k - 1$ forms a closed curve, denoted by γ_k. This curve is assumed to be simple, so that γ_k is a polygon. If all these curves, including the single point, $(X_1^{(0)}, Y_1^{(0)})$, are plotted in a plane, the resulting graph is composed of level curves of polygonal contours which represent the breast, much the same way a contour map represents a portion of the Earth's surface (see *Figure 1*). Hereafter, the name contour map shall refer to this graph of level curves associated with the digitized breast. For $1 \leq k \leq N$, the curve γ_k shall be called the kth contour of the contour map for the digitized breast D.

Since a digitized breast is a precise model of the photographed breast, aberrations on the breast surface should result in corresponding indentations on various level curves of the contour map. More specifically, a surface concavity should result in two-dimensional concave areas on the corresponding contours of D. Consequently, the first problem mentioned in the introduction to this paper, i.e. detecting surface concavities, reduces to locating the concave areas of the curves γ_k $k = 1,2,\ldots, N_k$. Because the distance between two Z-levels is at most 0.25 cm and the distance between two consecutive plotted points on a contour is even less, no significant concavity could lie between two levels or two plotted points, and escape detection.

Since the computer algorithm used to locate the concavities analyses one contour at a time, its description will be discussed in a general setting. Let $\{w_j\}_1^{n+1}$, $(n > 4)$ be points in the XY-plane such that $w_{n+1} = w_1$. The polygon γ derived from the points $\{w_j\}_1^{n+1}$ shall refer to the path (which is assumed to be simple) obtained by connecting consecutive points with line segments (see *Figure 3a*). Concave areas, if any, of γ are the sets of points which lie in the intersection of $E(\gamma)$ and $C(\gamma)$, where $C(\gamma)$ represents the closed convex hull of γ, that is, the smallest closed convex set containing γ (see *Figure 3d*). (Recall that a set ζ is convex if for any two points a, b in ζ, the line segment joining a and b is a subset of ζ). Locating the concave areas of γ is accomplished by first constructing $C(\gamma)$ and then removing from that set the points in $I(\gamma)$. The algorithm used to construct the convex hull of a polygon depends on the following two theorems, the proofs of which are given in the Appendix. The quotient in equation (3) below refers to complex division.

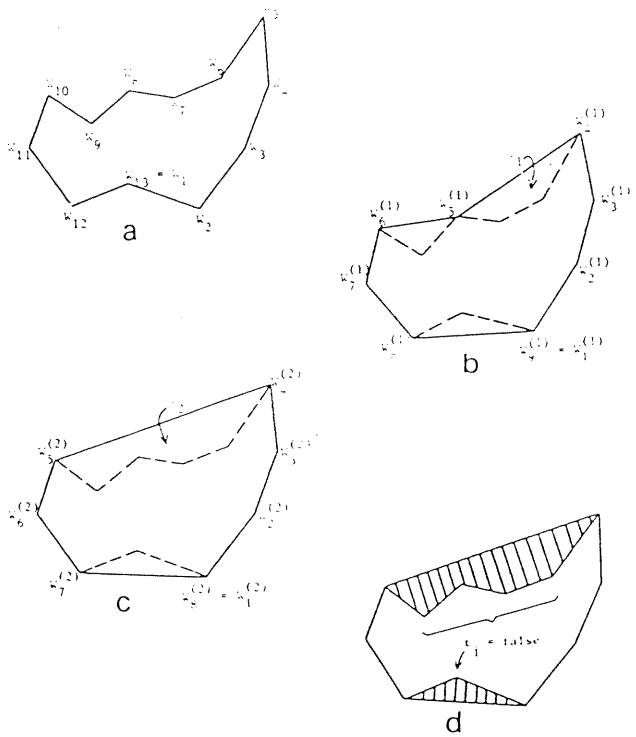

Figure 3 Sequential steps demonstrating the location of the concave area(s) of a polygon. The Boolean variable t_j for j = 1,6,7,8, and 9 indicates those vertices of the polygon which causes it to deviate from the convex hull.

Theorem 1. Let $\{w_j\}_{j=1}^{n}$ and γ be as described in the previous paragraph. Suppose that for each j, $j = 1,2,\ldots, n-1$, the angle:

$$\theta_j := \arg\left(\frac{w_{j+1} - w_j}{w_j - w_{j-1}}\right), \; w_0 := w_{n-1} \tag{3}$$

satisfies $0 < \theta_j < \pi$. Then, $C(\gamma) = \gamma \cup I(\gamma)$ ($\cup$ = set union).

Note that $0 < \theta_j < \pi$ if and only if the second coordinate of $(w_{j+1} - w_j) / (w_j - w_{j-1})$ is positive. Likewise, if this coordinate is negative, then $\theta_j > \pi$.

Theorem 2. Using the notation of Theorem 1, suppose there exists a p, $1 \leq p \leq n-1$, for which $\theta \geq \pi$. Then $C(\tilde{\gamma}) = C(\gamma)$ where $\tilde{\gamma}$ is the polygon derived from the original set of points with exception of w_p, that is, the set $\{w_j: 1 \leq j \leq n-1, j = p\}$. (If $p = 1$, then after removing w_1 from the set, w_n is redefined to be w_2.)

The algorithm used by the authors to construct $C(\gamma)$ can now be described. First, by Theorem 2, all vertices of γ for which $\theta_j \geq \pi$ can be removed leaving say, the points $\{w_j^{(1)}\}_{j=1}^{n_1}$ from which a new polygon $\gamma^{(1)}$ can be formed. (Refer to *Figure 3a,b* for an example using 12 points). Since $\gamma^{(1)}$ may also have vertices whose interior angles measure at least π, this process must be repeated for $\gamma^{(1)}$ using the formula:

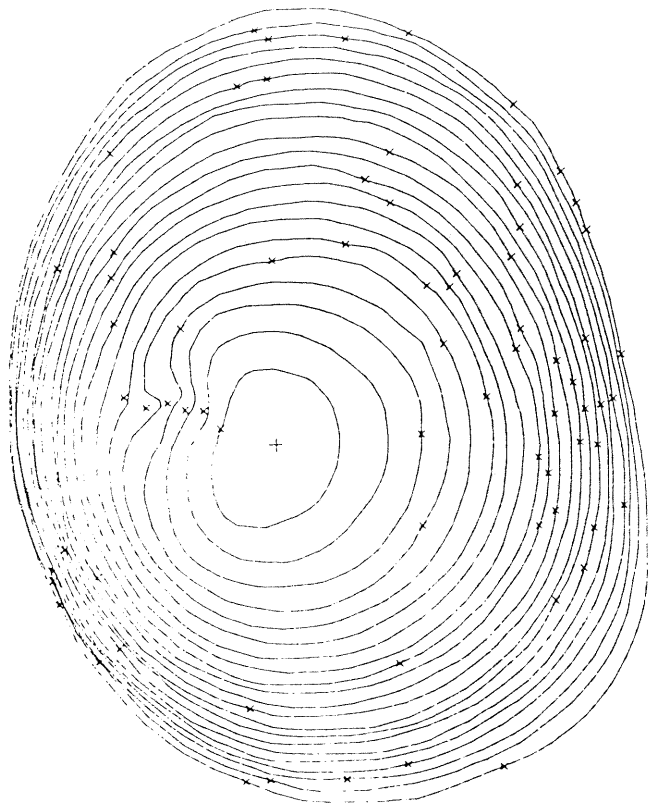

Figure 4 Contour map with concave areas indicated by *x*.

$$\theta_j^{(1)} := \arg\left(\frac{w_{j+1}^{(1)} - w_j^{(1)}}{w_j^{(1)} - w_{j-1}^{(1)}}\right), \; j = 1,2,\ldots, n_1 - 1$$

This results in a set of points $\{w_j^{(2)}\}_{j=1}^{n_2}$ from which $\gamma^{(2)}$ can be derived (see *Figure 3c*). This is repeated until one obtains a set of points $\{w_j^{(\ell)}\}_{j=1}^{n_\ell}$ from which a polygon $\gamma^{(\ell)}$ is derived and for which the interior angles have measured less than π (*Figure 3d*). By Theorem 1, $C(\gamma^{(\ell)}) = \gamma^{(\ell)} \cup I(\gamma^{(\ell)})$ and by Theorem 2 $C(\gamma) = C(\gamma^{(\ell)})$. The concave areas of γ are, then, the connected components of $(\gamma \cup I(\gamma))/C(\gamma^{(\ell)})$ where the solidus denotes set difference (note the two shaded areas of *Figure 3d*).

Since the determination of the concavities depends on both γ and $\tilde{\gamma}$, it is necessary to retain all the points $\{w_j\}_{j=1}^{n}$ in the implementation of this algorithm. This was accomplished by a Boolean vector $\mathbf{T} = (t_1, t_2, \ldots, t_{n-1})'$ where, at first, t_j is set equal to true for $j = 1,2,\ldots, n-1$. Next, the values θ_j are computed and t_j is given the value false if $\theta_j \geq \pi$. The process is then repeated using the set of points $\{w_j: t_j = \text{true}, \; 1 \leq j \leq n-1\}$. This is continued until no change in $\mathbf{T}$ takes place. For example, referring to *Figure 3b*, the vector $\mathbf{T}$ would have values $t_1 = t_6 = t_7 = t_8 =$ false while the other components of $\mathbf{T}$ would remain true. Consequently, w_1, w_6, w_7 and w_9 are 'removed' while w_2, w_3, w_4, w_5, w_8, w_{10}, w_{11}, w_{12} would correspond to $w_1^{(1)}$, $w_2^{(1)}$, $\ldots$, $w_8^{(1)}$, respectively, in *Figure 4b*. Repeating these steps would change t_8 to false, thereby completing the process. The computer implementation of this technique is described in Algorithm A below. This algorithm requires the subroutine *Proceed-follow* (see

Algorithm B). This procedure returns for a Boolean vector T and integer j, the largest integer r such that w_r procedes w_j on γ and t_r = true. It also returns the least integer s such that w_s follows w_j and for which t_s = true. For example, suppose T is the final vector used in the example above so that t_1 = t_6 = t_7 = t_8 = t_9 = false and all other coordinates have value true. Then *Proceed-follow* (t,j,r,s) would return $r = 12$ and $s = 3$ if $j = 2$, and $r = 5$ and $s = 11$ if $j = 10$. Note that in the first case $s < r$ but that w_r $(= w_{12})$ procedes w_j $(= w_2)$ which is followed by w_s $(= w_3)$. For programming purposes, and in the remainder of this paper, this circular or modular way of addressing, which arises from the fact that γ is a closed contour, will be used.

Once the vector T has been divided using Algorithm A, the concave areas of γ are determined using the following scheme. Suppose j_1 and j_2 are integers which satisfy:

$$\begin{aligned} &i)\ t_{j_1} := t_{j_2} = \text{true};\\ &ii)\ t_{j_1+1} = t_{j_1+2} = \ldots = t_{j_2-1} = \text{false};\\ &iii)\ \text{the points } w_{j_1}, w_{j_1+1}, \ldots, w_{j_2} \text{ are consecutive vertices of the polygon } \bar{\gamma}. \end{aligned} \tag{4}$$

Again, note that it is possible that $j_2 < j_1$ because of the circular addressing technique. The region, bounded by the path derived from $\{w_l\}_{l=j_1}^{j_2}$, is a connected component of the concave area of γ (consider the shaded areas of *Figure 3d*). The centre of mass of these concave regions will be used in the next section for computational and reference purposes. The centre of mass for the points described by equation (4) is given by:

$$M := \frac{1}{r} \sum_{j=j_1}^{j_2}{}' w_j \tag{5}$$

where

$$r := \begin{cases} j_2 - j_1 + 1 \text{ if } j_1 < j_2 \\ n + j_2 - j_1, \text{ otherwise.} \end{cases}$$

The value r represents the modular adjustment described above. The prime (') on the summation symbol indicates that the sum should include all vertices on γ encountered as an observer traverses γ from w_{j_1} to w_{j_2} (see the examples below).

Consider again the example illustrated by *Figure 3*. There are two concave areas for γ as indicated by *Figure 3d*. For the top areas $t_5 = t_{10}$ = true while t_6 = t_7 = t_8 = t_9 = false; consequently, $j_1 = 5$ and j_2 = 10 satisfy equation (4). The centre of mass for this area, as determined by equation (5), has the value:

$$M := \frac{1}{6} \sum_{j=5}^{10}{}' w_j \quad (6 = 10 - 5 + 1)$$

For the bottom concave area it is evident that j_1 = 12 and j_2 = 2. Consequently for this region:

$$M := \frac{1}{3} \sum_{j=12}^{2}{}' w_j = \frac{1}{3}(w_{12} + w_1 + w_2)$$

$$(3 = 13 + 2 - 12)$$

A real function subprogram *Centre-of-mass* (t,w,k,j) is needed at each step to compute the centres of mass for the concave areas of each γ_k. The computer implementation of this function follows directly from equation (5), and its description is omitted. *Figure 4* is the contour map given in *Figure 1* with concave areas indicated by an 'X'.

CLASSIFICATION OF CONCAVE AREAS

In this section the algorithm used to classify all the concave areas of the various contours of a contour map D is described. The purpose of the classification is to determine which concave areas on adjacent contours of D form a pattern that represents a single breast surface depression; the initial step in this process is to designate those contours, if any, which have concave areas. This is accomplished by using Boolean variables b_k, $k = 1,2,\ldots, N$, which are assigned the value true if γ_k has at least one concave area, and the value false otherwise. Consequently, if b_k = true for some k, $1 \le k \le n$, then γ_k has, say $L_k \ge 1$ concave areas. Let C_j, $j = 1,2,\ldots, L_k$ denotes these areas where the ordering is determined by traversing γ_k in a clockwise direction. Then the set

$$A := \{C_j^{(k)} : b_k = \text{true}, 1 \le j \le L_k, \quad 1 \le k \le N\} \tag{6}$$

contains the totality of concave areas for all contours of D.

The next step locates patterns of concave areas on a sequence of contours by requiring that the distance between two areas on adjacent contours be within a certain tolerance or threshold value. The algorithm uses the length of the line segment between the two centres of mass of two concave areas as a measurement of this distance. More precisely, for each $C_j \in A$, let $M_j := (p_j, q_j)$ denote its centre of mass according to equation (5). The distance between any two concave areas $C_j^{(k)}$ and $C_s^{(r)}$ in A is defined by the Euclidean distance:

$$d(C_j^{(k)}, C_s^{(r)}) := ((p_j^{(k)} - p_s^{(r)})^2 + (q_j^{(k)} - q_s^{(r)})^2)^{1/2} \tag{7}$$

This is the length of line segment connecting $(p_j^{(k)}, q_j^{(k)})$ and $(p_s^{(r)}, q_s^{(r)})$. Let $\tau > 0$. The quantity τ is called a *threshold* value and will be defined later. Two concave areas of adjacent contours $C_j^{(k)}$ and $C_s^{(k+1)}$ are said to be *consecutive* (with respect to τ) if

$$\begin{aligned} &i)\ d(C_j^{(k)}, C_s^{(k+1)}) \le \tau,\\ &ii)\ d(C_j^{(k)}, C_s^{(k+1)} < d(C_j^{(k)}, C_t^{(k+1)}), \end{aligned} \tag{8}$$

where $t = s$ and $1 \leq t \leq L_{k-1}$. Condition (i) suggests that $C_j^{(k)}$ and $C_s^{(k+1)}$ are related to the same surface concavity (within a threshold value of τ). The second condition ensures that no other concave area of γ_{k+1} is a more likely candidate for a particular surface concavity. A *pattern* is any maximal subset of A of consecutive concave areas. That is, a collection $\{C_{j_k}^{(k)}, C_{j_{k+1}}^{(k+1)}, \ldots, C_{j_l}^{(l)}\}$ of consecutive concave areas forms a pattern if $C_{j_k}^{(k)}$ is not consecutive to any concave area of γ_{k-1} if $k > 1$ and $C_{j_l}^{(l)}$ is not consecutive to any concave area of γ_{l+1} if $s < n$.

The implementation of the algorithm for classifying the concave areas of contours of D into patterns is relatively simple. Let $k_0 \geq 1$ be the smallest integer for which γ_k has a concave area, i.e., $b_{k_0} =$ true. Next, find a concave area, if there is one, at the $k_0 + 1$ level which is consecutive to $C_1^{(k_0)}$. Continue this until a maximal set of consecutive concave areas, i.e. a pattern, is found. This process is repeated until all $C \in A$ belong to some pattern. Clearly, those concave areas which already belong to a pattern cannot be used to start a new pattern and must be removed from the list A of available areas given by equation (6). The pseudo code for this procedure is given as Algorithm C below where the ordered pair (q,s) is used to identify $C_s^{(q)}$. An array *used* stores the identifiers of those areas of A which belong to a pattern. Several arrays P_1, P_2, ... are used to denote patterns; each contains a list of identifiers indicating those areas it contains. A function $Put(q,s,\cdot)$ is used to store the identifier (q,s) into a set or linked list denoted here by "$\cdot$". For example, $Put(q,s,used)$ places the identifier (q,s) into the array *used* while $Put(q,s,P_i)$ places the identifier (q,s) into the pattern P_i. A Boolean function $Locate(M,q,r,s\tau)$ returns *true* if there is a concave area $C_s^{(q+1)}$ which is consecutive to $C_r^{(q)}$ (note the role of s). This function returns *false* if there is no such area: it necessarily returns *false* if $q = n$. A Boolean function $Available(\cdot,q,s)$ returns *true* if the identifier (q,s) is not in the set denoted by "$\cdot$". If, for example, $used = \{(3,1),(4,2),(5,1)\}$ then $Available(used,4,2) =$ *false*, indicating that $C_2^{(4)}$ cannot be used to start a new pattern. However, $Available(used,4,1) =$ *true* and $C_1^{(4)}$ can be used to start a new pattern. Pseudocode for these functions is omitted. The integer np represents the total number of patterns found. Other notations (e.g. b_k, m_k) have previously been defined.

The threshold requirement equation (8i) ensures that two concave areas on consecutive contours are reasonably close to each other. This value τ should, then, depend on the particular contour map D being studied. Larger breasts will require larger values for τ. In this study τ was given the value:

$$\tau := \left(\left(\frac{d_1(x) + d_2(x)}{2} \right)^2 + \left(\frac{d_1(y) + d_2(y)}{2} \right)^2 \right)^{1/2}$$

Here $d_1(x)$ represents the difference of the largest x-values of γ_1 and γ_2 and $d_2(x)$ represents the difference of the smallest x-value of γ_1 and γ_2. A similar statement is true for $d_1(y)$ and $d_2(y)$. *Figure 5*

Figure 5 Contour map in which two distinct patterns of concavities have been identified.

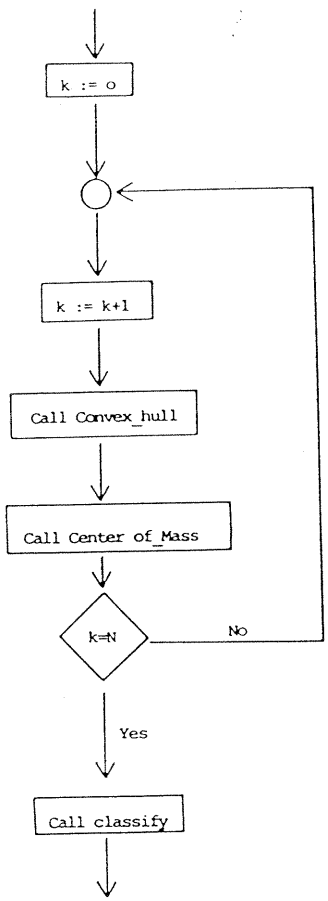

Figure 6 Flow chart of the main program.

is the contour map given in *Figures 1* and *4*. Two patterns were identified. These are indicated by the numbers '1' and '2'.

A brief description of the flow of the main program is in order at this point. First, for a given contour level k the main program calls procedure *Convex-hull* which computes the coordinates of the convex hull and vector T and γ_k. The next step is to calculate the centres of mass for any concave areas of γ_k. Once these computations have been performed for all contour levels, a call to procedure *Classify* completes the algorithm. (A flow chart for the main program is given in *Figure 6*.)

CONCLUSIONS

The development of the algorithms presented in this paper is part of a multifaceted project aimed at providing a non-invasive, rapid and cost effective screening or prescreening methodology for the early detection of breast cancer. The algorithms were applied initially to the same subjects whose breast contour maps were visually inspected by Loughry *et al.*[6,8], with very encouraging results. More recently, this procedure was employed on a data set consisting of 29 subjects (58 breasts) which included 7 cancers, 13 benign tumors and 38 apparently normal breasts[7]. Results of the use of this analysis technique included: the correct identification of the malignant neoplasms in seven out of seven breasts; correct identification of the location of benign masses in 11 of 13 breasts; and correct assessment of no apparent breast pathology in 30 out of 38 breasts. These results demonstrate the biostereometric analysis can detect breast surface concavities which are apparently associated with underlying breast pathologies. Further development of this technique in conjunction with other data analysis procedures, such as computation of breast volume and volume distribution, as well as determining discrepancies between right and left breasts may provide the satisfactory breast cancer detection tool so vitally needed.

ACKNOWLEDGEMENT

The authors acknowledge the Akron City Hospital Foundation and The University of Akron Faculty Summer Research Fellowship Program for the support of this project.

REFERENCES

1. Herron RE. Biostereometric measurement of body form. *Yearb Phys Anthropol* 1972; 16: 80-120.
2. Karara HM (Ed.). *Handbook of Non-Topographic Photogrammetry*, 2nd edn. Falls Church, VA.: ASPRS, 1987.
3. Slama CC (Ed.). *Manual of Photogrammetry*, 4th edn. Falls Church, VA.: Am Society of Photogrammetry, 1980.
4. Sheffer DB, Price TE, Loughry CW. Reliability of a photogrammetric determination of the breast-thorax boundary. *Proceedings of The Second International Technical Symposium on Optical and Electro-optical Science and Engineering (SPIE), Biostereometrics '85:* Cannes, December 2-6, 1985.
5. Sheffer DB, Price TE, Loughry CW, Bolyard BL, Morek WM, Varga RS. Validity and reliability of biostereometric measurement of the human female breast. *Ann Biomed Eng* 1986; 14: 1-14.
6. Loughry CW, Liebelt RA, Herron RE, Hamor RH, Liebelt AG, Cuzzi JR. Detection of breast cancer utilizing biostereometric analysis, a preliminary report. *Breast, Diseases of the Breast* 1980; 6(2): 11-15.
7. Sheffer DB, Herron RE, Morek WM, Proietti-Orlandi F, Loughry CW, Hamor RH, Liebelt RA, Varga RS. Stereophotogrammetric method for breast cancer detection. *Proceedings of the 26th Annual Technical Symposium of Society of Photo-optical Instrumentation Engineers, Biostereometrics '82:* San Diego, California, August 1982.
8. Loughry CW, Sheffer DB, Hamor RH, Liebelt, AG, Proietti-Orlandi F, Varga RS. Breast cancer detection utilizing biostereometric analysis. *Cancer Detec Prev* 1981; 4: 589-94.
9. Markushevich AI. *Theory of Functions of a Complex Variable*, I, II, III. Englewood Cliffs, NJ: Prentice Hall Inc. 1965.
10. Herron RE, Cuzzi JR, Hugg JE, Rauk KR. Stereometric measurement of body and limb volume changes during extended space missions. *Final report NAS9-10567:* 1971.
11. Eves H. *A Survey of Geometry*. Boston: Allyn and Bacon, 1972.

APPENDIX

Proof of Theorem 1. Denote by ℓ_j the line segment connecting the points w_{j-1} and w_j, $j = 1,2,\ldots, n-1$. The θ_j is the angle in $I(\gamma)$ formed by ℓ_j and ℓ_{j+1} (ref. 9). It is known that the sum of the interior angles of a polygon of n sides is equal to $(n-2)\,\pi$ radians[11]. Thus, at most $(n-3)$ interior angles can have measure grater than π radians. Now suppose for the purpose of a contradiction that $\gamma \cup I(\gamma)$ is not convex. Then there are points $a,\ b \in \gamma \cup I(\gamma)$ such that the line segment ℓ joining a and b is not fully contained in this set. Then line segment ℓ must cross γ at least twice. Let $\tilde{\ell}$ be a connected portion of ℓ joining points on γ and having no intersection with $I(\gamma)$. The line segment $\tilde{\ell}$ breaks the polygon γ into two polygonal paths say γ_1 and γ_2. One of these, say γ_1, satisfies the relation $I(\gamma_1) \subset E(\gamma)$. Thus, at most two interior angles of γ_1 have measurement less than π radians (those where ℓ joins γ) since the remainder are exterior angles of γ. This means that all but two angles of γ_1 have measurements larger than π radians. This contradiction completes the proof.

Proof of Theorem 2. Consider the triangle formed by the lines ℓ_p, ℓ_{p-1} *and* ℓ where ℓ is the line segment joining w_{p-1} and w_{p+1}. Let a be any point of ℓ other than the end point w_{p-1} and w_{p+1}. Clearly, $a \in C(\gamma)$ since w_{p-1} and w_{p+1} are in this set. Next, construct the line L which passes through w_p and a. Evidently, as an observer travels on L from a through w_p, he passes from $E(\gamma)$ over w_p into $I(\gamma)$. As he continues on L, then, he must again cross γ, say at b. The point b cannot be on ℓ_p or ℓ_{p+1}. Thus, that part of L connecting a and b is a subset of $C(\tilde{\gamma})$ (because $C(\tilde{\gamma})$ is convex). But w_p is an element of this line segment so $w_p \in C(\tilde{\gamma})$. Consequently, ℓ_p and ℓ_{p-1} and, hence, all of γ, are contained in $C(\tilde{\gamma})$. This means $C(\gamma) \subset C(\tilde{\gamma})$. It is obvious that $C(\tilde{\gamma}) \subset C(\gamma)$ and this completes the proof.

Algorithm A: Procedure *Convex-hull*

```
Procedure Convex-hull (n: integer; w: array [1 . . n] of complex
numbers; var t: array [1 . . n – 1] of Boolean);
/★Locates the vertices of a polygon which determines its convex
hull. A vertex w_j which can be omitted is indicated by setting t_j =
false.★/
var finish: Boolean; r,s,j: integer;
begin
 for j : = 1 to n – 1 do t_j : = true;
 finished : = false;
 while not finished begin
  finished : = true;
  for j : = 1 to n – 1 do begin
   if t_j = true then begin
    call Procede-follow (T,j,r,s);
    if imaginary ((w_r – w_j)/(w_j – w_s)) < 0.0 then begin
     t_j : = false;
     finished : = false
    endif
   endif
  endfor
 endwhile
end
```

Algorithm B: Procedure *Procede-follow*

```
Procedure Procede-follow (t: array [1 . . n – 1] of Boolean; n, j:
integer; var r,s: integer)
/★Returns for a Boolean vector t the largest integer r such that w_r
precedes w_j and the smallest integer s such that w_s follows w_j★/
var finish: Boolean; incr, m,k: integer
begin
 incr = 1;
 for m = 1 to 2 do
  incr : = – incr;
  finished : = false;
  r = s;
  s = j;
  while finish = false do
   s : = s + incr;
   if s < 1 or s ≥ n then s = s – incr* (n – 1);
   if t_s = true then finish : = true
  endwhile
 endfor
end
```

```
Procedure-Classify (N: integer; L,b: array [1 . . N] of Boolean;
var P: array [1 . . 8] of pattern; var np: integer)
/★Classified the concave areas of the contours of D by patterns.★/
var i,j,k,q,r,s: integer;
begin
 i := 1;
 for k := 1 to N do
  if b_k then
   for j := 1 to L_k do
    if available (used, i, j) then begin
     put(k, j, P_i);
     q := k;
     r = j:
     while locate (M, q, r, s, τ) do begin
      q := q + 1;
      put (q, s, used);
      put (q, s, P_i);
      r := s
     endwhile;
     i := i + 1
    endif
   endfor
  endif
 endfor
 np := i - 1   /★ np is the number of patterns found ★/
 nd
```