Shape matching in multimodal medical images using point landmarks with Hopfield net

S. Banerjee*, D. Dutta Majumdar[1]

Electronics and Communication Sciences Unit, Indian Statistical Institute, 203 Barrackpore Trunk Road, Calcutta 700035, India

Received 23 April 1997; accepted 22 March 1999

Abstract

Analysis of concavities within complex patterns provides a promising approach to structural shape descriptions. Such concavities are characterized by entrance, exit and height points, collectively termed as point landmarks, as well as height and width. This paper presents a scheme for using these characteristics with the Hopfield neural network for matching structural shape descriptions corresponding to identical regions of interest in images of the same object obtained from different sensors. Application to biomedical imaging is also discussed. In these applications, both the shape description and the Hopfield network based matching scheme are illustrated for matching neuroanatomical shapes of a human brain obtained from axial slices obtained from CT and MR modalities of the same region.

Keywords: Hopfield net; Point landmarks; Shape matching; Optimization

1. Introduction

One of the primary tasks of Pattern Recognition is the description of an object in terms of its shape. This shape description is especially crucial in the processing of digital binary images and one of the techniques widely adopted is the structural approach. The description of an object in terms of its parts and spatial relations

between these parts constitutes structural object descriptions [14] and such descriptions explicitly represent the spatial organization of the primitives forming these objects. Structural descriptions are rotation and translation invariant and are useful for shape matching [15]. A complex scene can be decomposed into regions with simpler shapes and described in terms of features of these regions and spatial relations within these regions. The primitives making up the object can be low level geometric tokens, like lines, arcs or subparts of different shapes which are characterized by a set of property-value pairs [3]. Real scenes are often extremely complex and cannot always be decomposed into simpler shapes. An alternative approach is to analyze a pattern on the basis of its global shape properties like elongation or roundness and geometric properties like area and moments. Such global features suffer from a high degree of imprecision. Location and analysis of concavities within the complex patterns provides a more promising alternative in structural shape description [4]. These concavities can be used in registration as well as structural shape matching of binary images. Registration is the process of obtaining a point to point correspondence between two images of the same object which might be obtained from two different sensors, or the same sensor at two different times. Such correspondence can be obtained by finding a mapping between structural object descriptions based on the equivalence of the property-value pairs such that the spatial constraints between the primitives in one shape are preserved by the matched primitives in the other shape. Matching an object with an atlas can be conducted similarly. This constitutes a constraint satisfaction problem. To obtain the best possible match, a mapping is to be determined so as to maximize the match between the common portions of the structural descriptions. This is essentially a constraint optimization problem and Hopfield networks [8–10,12] are useful for solving such problems.

In medical imaging applications, one often encounters situations where anatomical regions of interest (ROI) from different images have to be matched. This matching can be between an image cross-section of the ROI obtained from a sensor and a medical atlas, or images obtained from two or more different sensors or images acquired by the same sensor at different times. Object descriptions of the ROI can be obtained from edge images of the ROI. These ROIs have a purely geometric representation and can be extracted from the original image by edge segmentation. The primitives forming such object descriptions constitute lines and arcs. These lines and arcs form concavities in the image, which possess entrance and exit points as well as points of inflexion, which constitute point landmarks or signatures. In addition, concavities are characterized by their height and width. The selection of such point landmarks is based on the geometric invariance properties of the point landmarks within the image [2].

In this paper, we describe a technique for structural shape matching of concavities found in 2D edge images of the same object obtained from different sources. These concavities constitute the primitives comprising the object. The geometrically invariant point landmarks described above constitute the equivalent property – value pairs to be matched in the different images. This is an inexact matching problem where the next step is to find a mapping which maximizes the match between the common portions of the two structural descriptions which are the edge images of the ROI, and therefore, this is a constraint optimization problem. A Hopfield type network similar

to the one described in [3] used to solve this problem has a similar architecture as that used by Lin [11]. The energy function that we have considered caters to problems related to inexact matching of the descriptions.

The paper is organized as follows. Section 2 formulates the problem. The network and the energy function which we have used have been described in Section 3 and results for two case studies for a set of computed tomography (CT) and magnetic resonance (MR) images are covered in Section 4. Discussions and concluding remarks are given in Section 5.

2. The problem

Complex scenes such as CT or MR cross sections of the human body contain many ROI that can be extracted using standard routines. Location of shapes and determination of their characteristics are described in [2], so we briefly mention the steps involved. Canny's edge extraction operator [7] and edge linking is then applied to these ROIs to obtain binary images of the edges of the ROIs and the entire edge point set is traced by an edge following algorithm [1]. These binary images generally contain numerous concavities. For each closed contour, concavities are obtained by substracting the contour from the convex hull [13]. Entrance and exit points are obtained from the bitangent to the concavity and the height point is the point of inflexion and is furthest from this bitangent. The point of intersection of the perpendicular from the concavity height point to the bitangent is the base point. Concavities within the edges constitute the subparts and these are located and labeled, starting with the leftmost/topmost concavity and proceeding in a counter clockwise direction. Each subpart (or primitive) is characterized by its eight attributes, namely, the concavity height point, the entrance, exit and base points, height, width, and distances between the height point and entrance and exit points. Distinction between entrance and exit point is purely arbitrary and we adopt the convention of taking the leftmost or uppermost (in that order) point as the entrance point. Distances between the concavity height points of the different concavities constitute the relational constraints between the primitives. These are illustrated in Fig. 1 with the nodes denoting the attributes (height points) and arcs giving the numerical parameters indicating the distance. e_1 and E_2 are, respectively, the entrance and exit points, h is the height point and b is the intersection of the perpendicular from h to the bitangent of the concavity (see Fig. 2). e_1 and e_2 lie on this bitangent. Numeric parameters corresponding to the distance between the nodes, are assigned to them. There are four numeric parameters:

$r1$ = distance between e_1 and h,
$r2$ = distance between e_2 and h,
$r3$ = distance between e_1 and e_2
$r4$ = length of perpendicular from h to bitangent (b).

Additional numeric parameters are assigned for interconcavity distances. For the case study presented here, these correspond to

$r5$ = distance between h of one concavity and those of adjacent concavities.

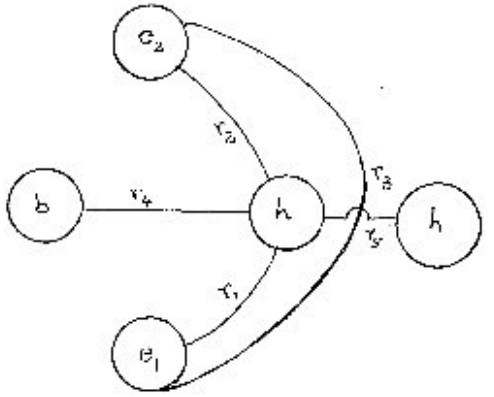

Fig. 1. Relational constraints between primitives and their attributes.

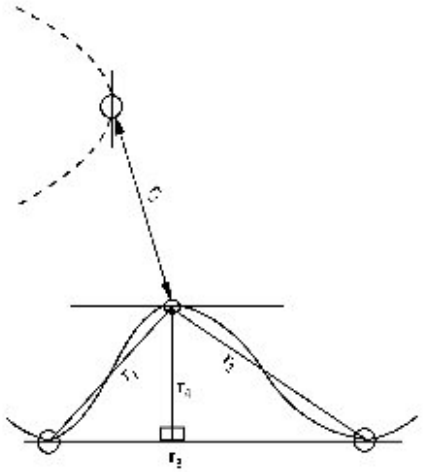

Fig. 2. Concavity and its characteristics.

Distances can be calculated using the Euclidian distance. However, the Euclidian distance demands a lot of resources in terms of time and memory. Also, the points (nodes) are corrupted by noise, and so, an approximation is used. Following Borgefors [5], 'chamfer' distances are used to approximate the distances between the nodes. The underlying idea is to obtain global distances by propagating local distance, i.e. distances defined by a 3×3 pixel neighborhood surrounding the pixel under consideration. The two local distances have been defined to be the distance between horizontal/vertical neighbors, assigned a value of 3 units and the distance between diagonal neighbors, given a value of 4 units. Using these values, the maximum difference with the Euclidian has been found to be 8% [6]. Our results also confirm

that the chamfer distance approximately falls within 10% of the Euclidian distance. In order to compute the distances, the pixels corresponding to the nodes within a primitive are set to zero and other pixels are set to infinity for each primitive. The property-value pairs corresponding to the numeric parameters characterizing each distance is thus computed. The interconcavity distance can be computed by setting inflexion points of two concavities at a time (considered as nodes) to zero and others to infinity. Two passes are made over the scene. The first pass is the forward pass made from left to right and from top to bottom, i.e.,

forward
for $i = 2 \ldots$ rows do
for $j = 2 \ldots$ columns do

$$v^{n}_{i,j} = \text{minimum}(v_{i-1,j-1} + 4;\ v_{i-1,j} + 3;\ v_{i-1,j+1} + 4;\ v_{i,j-1} + 3;\ v^{o}_{i,j}), \tag{1}$$

where the superscript o denotes old, n denotes new and $v_{i,j}$ denotes the distance value of the pixel in position (i, j).

The second pass is the backward pass from right to left and from bottom to top, i.e.,

backward
for $i =$ rows $- 1 \ldots 1$ do
for $j =$ columns $- 1 \ldots 1$ do

$$v^{n}_{i,j} = \text{minimum}(v^{o}_{i,j};\ v_{i,j+1} + 3;\ v_{i+1,j-1} + 4;\ v_{i+1,j} + 3;\ v_{i+1,j+1} + 4). \tag{2}$$

One pass is sufficient to compute these distances. The backward pass is intended as a check.

3. The network

The Hopfield network [8–10] is a fully interconnected recurrent network with weighted links. Individual neuronal units have either binary or continuous valued activation functions, and can receive external bias in addition to inputs from other units. The network response is dynamic. For each input, outputs are calculated and these outputs constitute the new inputs for the succeeding iteration and this process is continued until convergence is reached and outputs become constant. The network is considered stable if the weights are symmetric and there are no self exciting loops. The energy function normally decreases each time the outputs of the neurons change their values, considered as a change of state of the network. The energy function is minimum when the network attains a minimum. The ability of such a network to reach a minimum energy function through iterative updating of its units has been used for obtaining good solutions [3] to certain classes of structural matching problems at various levels of complexity.

We consider two shapes having N primitives each and use a neural network having $N \times N$ neural units. Each unit is a possible match between the primitives of the two shapes. The lowest energy state of the network, in accordance with an appropriately

defined energy function, will represent the best mapping between candidate and prototype shapes satisfying all the constraints of the given problem.

Following [3], we assume energy function to be a sum of four terms, viz.,

$$E = E_1 + E_2 + E_3 + E_4. \tag{3}$$

The first term which is composed of two terms, ensures that only one primitive in the candidate shape can match the corresponding primitive in the prototype, i.e.

$$E_1 = A/2 \sum_p \sum_i \sum_j V_{ip} V_{jp} + B/2 \sum_i \sum_p \sum_{q \neq p} V_{ip} V_{iq} \tag{4}$$

with A and B being positive constants and V_{ij} representing the state of the (i,j)th neuron in the range $[0,1]$.

The first term on the RHS of Eq. (4) ensures one entry in each column whilst the second term ensures one entry in each row.

The second contribution to the energy function guarantees that the total number of primitives matched equals N. This term is written as

$$E_2 = C/2 \left(\sum_i \sum_p V_{ip} - N \right)^2 \tag{5}$$

with C being a positive constant. In the perfect matching case this term vanishes. For shapes with different numbers of primitives $N1$ and $N2$, the parameter N must be set to a minimum of $(N1, N2)$ as the network can find a match for the maximum of $\min(N1, N2)$ primitives. Even in the case of a partial match, a minimum value of this term ensures the maximum possible number of matches between primitives.

Primitives of the candidate and prototype shapes are characterized by a set of possible attributes and values of these attributes which are to be matched. This is represented by

$$E_3 = D/2 \sum_i \sum_p V_{ip} \sum_{a=1}^{A} \sum_{b=1}^{B} W_i(a)(P(i,a,b)(1 - C(p,a,b)) + C(p,a,b)(1 - P(i,a,b))), \tag{6}$$

where $W_i(a)$ is the priority or weight of the attribute a for the primitive i assumed to be known a priori, a is the variable representing individual attributes assumed to be indexed from 1 to A, and b is the index for individual elements of the possible value sets of the attributes. The value sets have been assumed to be finite and discrete. Cardinality of the largest set is assumed to be B.

$P(i,a,b)$ is the selector function which can be either 1 or 0, $P(i,a,b) = 1$ if the primitive i of the prototype possesses attribute a with value b, $P(i,a,b) = 0$ otherwise, $C(p,a,b)$ is the corresponding selector function for the primitives of the candidate shape. The fourth term on the RHS of Eq. (4) ensures the existence of identical spatial constraints between matched primitives of the candidate and prototype

shapes, viz.,

$$E_4 = E/2 \sum_i \sum_j \sum_p \sum_q V_{ip} V_{jq} \sum_{t=1}^{T} W(t)(f(i,j,t)(1 - g(p,q,t)) + g(i,j,t)(1 - f(p,q,t))). \quad (7)$$

where E is a positive constant, t is an index for relational attributes, $W(t)$ is the priority of the tth spatial relation, $f(i,j,t)$ is the selector function which takes values either 0 or 1, $f(i,j,t) = 1$ if the tth relation exists between the ith and jth primitives of the prototypes and 0 otherwise, and $g(p,q,t)$ is the corresponding selector function for the primitives of the corresponding candidate shape.

The connection weights are determined as follows. The coefficients of the quadratic terms correspond to the weights in the connection matrix whereas the coefficients of the linear terms correspond to the input bias of the neurons. The connection weights are thus

$$T_{ijpq} = -A\delta_{pq}(1 - \delta_{ij}) - B\delta_{ij}(1 - \delta_{pq}) - C - E \sum_{t=1}^{T} W(t)((f(i,j,t)(1 - g(p,q,t)) + g(i,j,t)(1 - f(p,q,t)))). \quad (8)$$

with δ_{ij} being the Dirac delta function.

Likewise, the input bias to each neuron is

$$I_{ip} = C_N - D/2 \sum_{a=1}^{A} \sum_{b=1}^{B} W_t(a)(P(i,a,b)(1 - C(p,a,b)) + C(p,a,b)(1 - P(i,a,b))), \quad (9)$$

where C_N is the excitation bias. The attributes of the primitives contribute to the input bias of the neurons and the relational constraints contribute to the inhibitory constraints between the neurons.

4. Methodology and results

In this section we consider two case studies of a contiguous axial CT and MR sections (Figs. 3 and 5) of the same region of the human brain of the same subject. The ventricles are the ROI for our case study, and these are extracted and shown in Figs. 4 and 6; with (a) and (b) denoting CT and MR, respectively In Figs. 4 and 6, the concavities and their characteristics are given. Table 1 lists the numerical parameters (with appropriate scaling) corresponding to the property-value pairs of the attributes as well as the parameters corresponding to the relational constraints between the primitives of the candidate (C) and prototype (P). Quantities within parentheses for $r5$ denote destination concavities. Table 2 gives the percentage agreement of the two concavities on the basis of these parameters. We find that the right and bottom concavities give the best match for Fig. 4 and concavity I of Fig. 6 gives the best

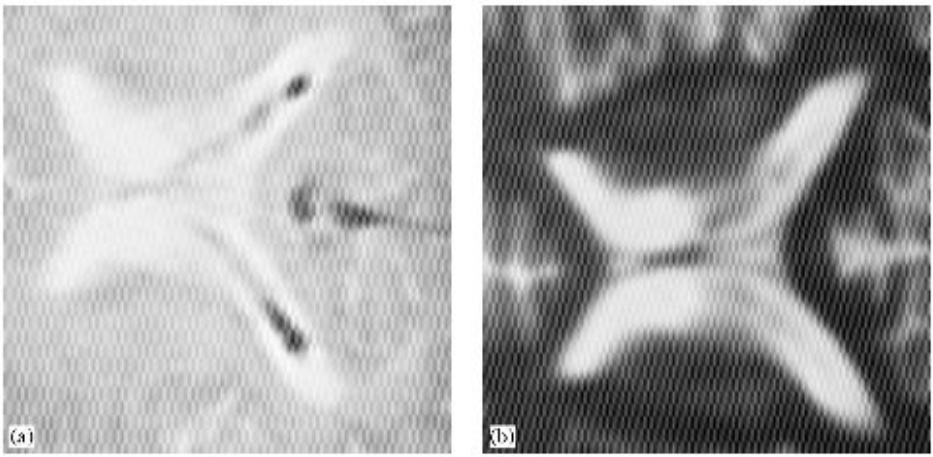

Fig. 3. Ventricle from the same region of the brain of a human subject obtained from (a) CT modality and (b) MR modality.

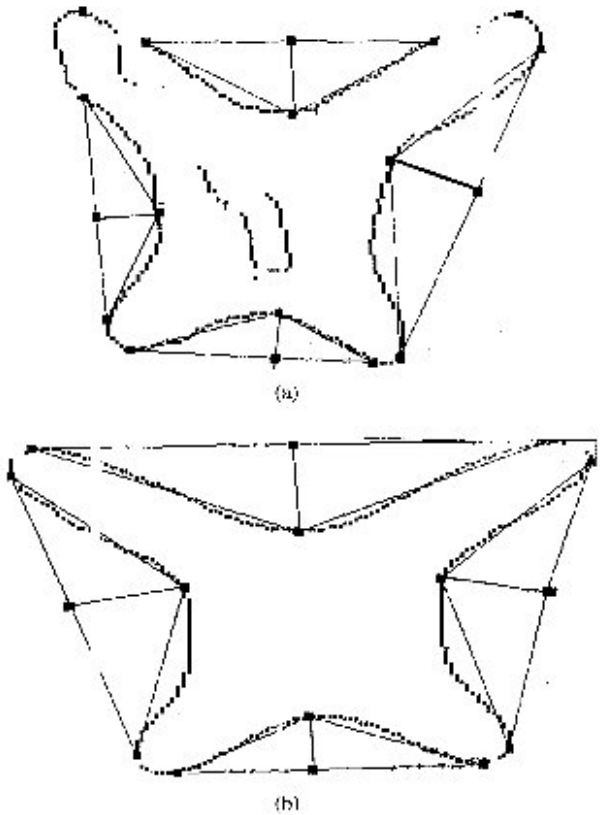

Fig. 4. Concavities and their characteristics extracted from the convex hull of the edge image of (a) CT and (b) MR of Fig. 3.

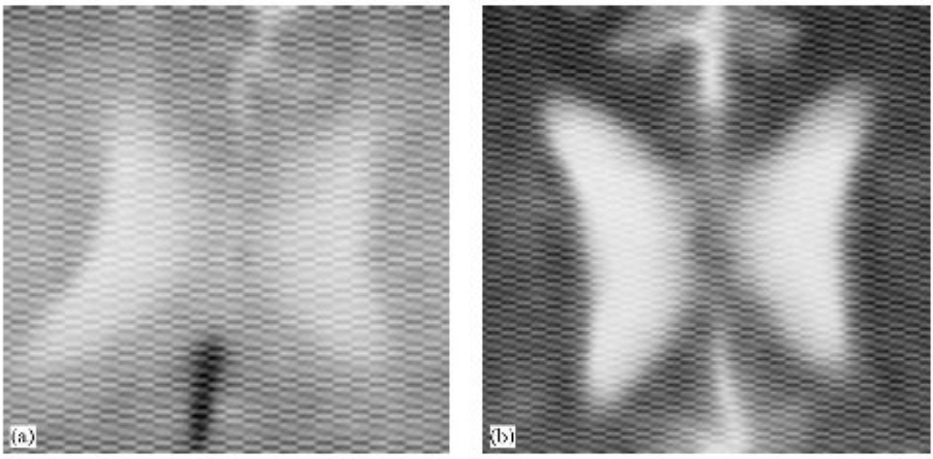

Fig. 5. Ventricular section from a contiguous slice obtained from (a) CT and (b) MR.

possible match. Quantities within parentheses in the columns for r5 denote destination concavities for the first test case. In the second case only one set of concavities are considered corresponding to the inner concavity of the ventricular shape.

The network [3] was simulated on a Silicon Graphics Indy machine using n^2 nodes, where n is the number of concavities, and a sigmoidal transfer function of the form

$$g(x) = 1/(1 + \exp(-x/T)). \tag{10}$$

Initial states of the neurons are given by

$$U_i(0) = U_{00} + \delta U_i \tag{11}$$

with $U_{00} = g^{-1}(1/N)$, due to the constraint that only N neurons corresponding to the consistent match of N primitives of the candidate will be on, subject to the condition

$$\sum_i \sum_j V_{ip} = N \tag{12}$$

for the single layered model. $|\delta U_i| < 0.1 U_{00}$ is a perturbation to the initial states of the neurons which prevents the network from getting stuck at the initial configuration. For the above transfer function $U_{00} = -T\log(N-1)$.

The property-value pairs of the individual attributes are the width and distance to the height point for entrance and exit points, height to width ratios for the base point, and height and distances to the entrance and exit points for the height point. Relational constraints are the interconcavity distances measured between height points. Selector functions $P(i,a,b)$, $C(p,a,b)$, $f(i,j,t)$ and $g(p,q,t)$ ensured that only the corresponding concavities between primitive and candidate shapes were taken into account. Several values of T were experimented with and it was found that the best result was obtained when T was chosen to be a tenth of the initial state. For larger T,

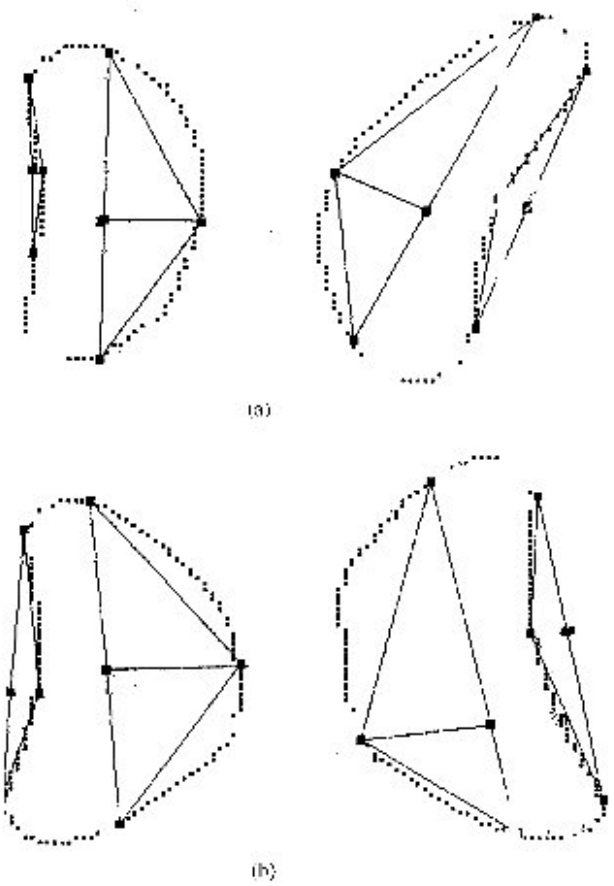

Fig. 6. Concavities and their characteristics extracted from the convex hull of the edge image of (a) CT and (b) MR of Fig. 5.

convergence was slower and the network had a propensity for getting stuck at spurious local minima for smaller T. Neurons in the network were updated synchronously and so the time step taken was reasonable (0.001) although good results were also achieved for 0.01. Weights of the attributes, relations and relational constraints were chosen according to how close the numerical parameters are, based on the results of Table 2. Lagrange multipliers A and B decide the weightage of domain constraints, while C and D are associated with the relative significance of the spatial constraints among primitives. Choosing larger values of A and B as compared with C and D results in mismatch due to violation of spatial constraints, whereas a reverse choice would result in mismatch of domain constraints and a compromise had to be made. In the present simulation, $A = B = 1$ and $C = D = 10$ were chosen. The goodness of match (computed as the average of $r1$, $r2$, $r3$, $r4$ with the ratio $r4/r3$ from Table 2) between corresponding primitives of prototype and candidate are given in

Table 1
Numerical parameters for arcs in CT and MR image

Fig.	Modality	Concavity	$r1$	$r2$	$r3$	$r4$	$r5$	$r4/r3$
3, 4	CT	L (Left)	98	94	164	44	(B) 138	0.27
		R (Right)	135	153	272	67	(B) 138	0.25
		T (Top)	118	139	223	57	(L) 115	0.26
		B (Bottom)	106	88	166	40	(L) 138	0.24
	MR	L	140	165	276	68	(B) 144	0.25
		R	135	162	279	73	(B) 135	0.26
		T	181	172	301	82	(L) 92	0.27
		B	114	91	178	49	(L) 144	0.27
5, 6	CT	I	120	116	192	60	(II) 92	0.313
		II	100	168	260	56	(I) 92	0.215
	MR	I	156	136	208	84	(II) 92	0.404
		II	160	120	216	84	(I) 92	0.389

Table 2
Agreement between same concavities of CT and MR images

Fig.	Concavity	$r1$	$r2$	$r3$	$r4$	$r4/r3$	$r5$
4	L	0.70	0.57	0.59	0.65	0.92	0.96
	R	1.0	0.94	0.97	0.92	0.96	0.98
	B	0.93	0.97	0.93	0.83	0.89	0.96
	T	0.65	0.81	0.74	0.70	0.96	0.80
6	I	0.77	0.85	0.92	0.71	0.77	1.0
	II	0.63	0.71	0.83	0.91	0.55	1.0

Table 3
Goodness of match between corresponding primitives of Fig. 4

Prototype candidate	L	R	T	B
L	0.772	0	0	0
R	0	0.963	0	0
T	0	0	0.817	0
B	0	0	0	0.919

Tables 3 and 4. Results indicate that the right and bottom concavities give the best match for Fig. 4 and concavity I for Fig. 6. The results of these two tables can also be expressed as neural units with 1.0 denoting a perfect match and 0 denoting a total mismatch.

Table 4
Goodness of match between corresponding primitives of Fig. 6

Prototype candidate	I	II
I	0.79	0
II	0	0.66

5. Discussion

We presented a scheme for matching corresponding shapes within complex images of the same object obtained from an atlas and a sensor or those obtained from two different sensors, by analyzing concavities corresponding to the intersection of the bitangent with the concavity giving entrance and exit points, and the point of inflection giving the concavity height point. Other characteristics chosen for the analysis include the width of the concavity which is the intercept of the bitangent with the concavity, and the height of the concavity which is the perpendicular from the point of inflexion to the bitangent. The choice of these point landmarks from the convex hull is computationally simple and obtained from purely geometric considerations. Distances chosen were chamfer distance measures.

Matching of shapes constitutes a constraint optimization problem, based on an appropriate expression for the Hopfield network energy function which caters to the possibilities of partial mismatches between primitives and their spatial constraints. In this scheme a differential criticality factor has also been incorporated between primitives and their spatial relations. Since numeric parameters corresponding to distances have been assigned as property-value pairs of the attributes, as well as spatial constraints, spatial relations are symmetric and the Hopfield network does not have asymmetric connection weights. In future work, attempts will be made to exploit the rotational and translational invariances to reduce the number of possible correspondences, specially when the concavities refer to several objects which can vary in position relative to each other. The network is designed to eliminate the possibility of false matches between primitives.

A parallel implementation of the network is also possible for complex scenes consisting of a number of concavities. With proper indexing, several concavities can be assigned to individual processors of a parallel machine. The network need not be fully interconnected. Thus, this methodology can be applied to biomedical images containing a number of concavities.

Acknowledgements

This work was funded by the Department of Science and Technology, Government of India. The images were provided by JMD Medicare, of Calcutta. The authors wish to thank the anonymous referees for their helpful comments.

References

[1] D.H. Ballard, C.M. Brown, Computer Vision, Prentice-Hall, Englewood Cliffs, NJ, 1982, p. 131.
[2] S. Banerjee, D.P. Mukherjee, D. Dutta Majumdar, Point landmarks for the matching of CT and MR images, Pattern Recognition Lett. 16 (1995) 1033–1042.
[3] J. Basak, S. Chaudhury, S.K. Pal, D. Dutta Majumdar, Matching of structural shape descriptions with Hopfield net, Int. J. Pattern Recognition Artificial Intelligence 7 (1993) 377–404.
[4] G. Borgefors, G. Sanniti di Baja, Analyzing convex 2D and 3D patterns, Computer Vision and Image Understanding 63 (1996) 145–157.
[5] G. Borgefors, Distance transformations in digital images, Comput. Vision, Graphics and Image Process. 34 (1986) 344–371.
[6] G. Borgefors, Hierarchichal chamfer matching: a parametric edge matching algorithm, IEEE Trans. Pattern Anal. Mach. Intelligence 10 (1988) 849–865.
[7] J. Canny, A computational approach to edge detection, IEEE Trans. Pattern Anal. Mach. Intelligence 8 (1986) 679–698.
[8] J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, Proc. Natl. Acad. Sci. USA 79 (1982) 2554–2558.
[9] J. Hopfield, Neurons with graded response have collective computational properties like those of two state neurons, Proc. Natl. Acad. Sci. USA 81 (1984) 3088–3092.
[10] J.J. Hopfield, D.W. Tank, Neural computation of decisions in optimization problems, Biol. Cybernet. 52 (1985) 141–152.
[11] W. Li, M. Nasrabadi, Object recognition based on graph matching implemented by a Hopfield style network, Proceedings of the International Joint Conference on Neural Networks II, Washington, DC, 1989.
[12] R.P. Lippman, An introduction to computing with neural nets, IEEE Accoust. Speech Signal Process. 61 (1987) 4–22.
[13] F.P. Preparata, M.I. Shamos, Computational Geometry: An Introduction, Springer, New York, 1985 (Chapter 3).
[14] L.M. Shapiro, R.M. Haralick, Organization of relational models for scene analysis, IEEE Trans. Pattern Anal. Mach. Intelligence 4 (1982) 575–602.
[15] L.M. Shapiro, P.G. Moriarty, R.M. Haralick, P.G. Mulgaonkar, Matching three dimensional objects using a relational approach, Pattern Recognition 17 (1984) 385–405.

Sreeparna Banerjee obtained her Ph.D. in theoretical Physics from the University of Virginia in 1987. Subsequently she held appointments in theoretical and computational Physics in Virginia. Since 1992, she has been with the Indian Statistical Institute, Calcutta (I.S.I.). Her current interests include neural networks, fuzzy logic, pattern recognition, computer vision and image processing with applications to biomedical imaging, and she has publications in these fields. She was a recipient of the National Science Talent Search scholarship (NCERT)

D. Dutta Majumdar received his Ph.D. in memory technology from Calcutta University. In 1955 he joined the Indian Statistical Institute (ISI). He was head of the Electronics and Communication Sciences unit of ISI, from 1972 to 1992. He has published more than 400 research papers and six books in the fields of memory technology, speech and other types of pattern recognition, image analysis, computer vision, AI, ANN, cybernetics and knowledge based computing. Dr. Dutta Majumdar is Professor Emeritus of ISI, Secretary-Director of the Institute of Cybernetics Systems and Information Technology (ICS + IT), Calcutta, and Emeritus Scientist of CSIR. He is a fellow of IAPR, INSA, INAE, IASc, IETE and a member of the Governing Boards of the International Association of Pattern Recognition, World Organization of Cybernetic Systems and International Fuzzy Systems Association.