COMPUTERS IN CARDIOLOGY

September 12-15, 1987
Leuven, Belgium

Computer Society Order Number 886
Library of Congress Number 80-641079
IEEE Catalog Number 87CH2544-5
ISBN 0-8186-0886-2
ISSN 0276-6574

THE COMPUTER SOCIETY OF THE IEEE

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

IEEE COMPUTER SOCIETY PRESS

A RECOGNITION METHOD FOR CORONARY ARTERY STENOSES

Krzysztof Ciesielski+
Krzysztof J. Cios*
Lucy S. Goodenday#

+Bowling Green State Univ., Bowling Green, OH, USA
*The University of Toledo, Toledo, OH, USA
#Medical College of Ohio, Toledo, OH, USA

SUMMARY

A method was developed for recognition of the site of coronary artery stenosis in patients having sentosis of the left anterior descending (LAD), circumflex (CCX) or right coronary (RCA) artery. The input data to the system was derived from the stress T1-201 radioactive counts in ten anatomical regions of the heart measured in three different views for each patient. These original gray level images were then preporcessed resulting in a series of numbers representing perfusion defects for each patient. The normal patient's data and the set of abnormal patients with arteriographically-proven coronary artery stenosis comprised learning data for the system. A diagnosis was given by cardiologists in the form: LAD - k%, CCX - n%, RCA - m%. The developed system is an algorithmic function which associates with every vector M a diagnosis d=S(M) i.e. the function from the set {LAD, CCX, RCA} into the interval [0, 100] representing the percentage of each stenosis. The percentages of the stenoses from a learning data set constitute essential information for the method.

INTRODUCTION

Frequently, in a patient with ischemic heart disease, it is important to know the functional significance of a particular stenosis, or to know which artery is causing myocardial ischemia. Thallium-201 scintigraphy is often used to estimate myocardial perfusion, and several computer analysis programs are available to determine the presence and size of a myocardial perfusion defect. Other clinically important data, however, such as the site of a coronary artery obstruction and its relative severity compared to stenoses in other arteries, has not been as readily derived from scintigrams. We have approached this problem before, trying to determine whether perfusion patterns exist which are diagnostically specific for location of stenoses in the major coronary arteries. It was found [1, 2] that there is a correlation between perfusion defect patterns and stenosis of the three main coronary arteries, but that it is difficult to recognize with satisfactory specificity and sensitivity by standard techniques.

In this paper, we describe an approach to the problem using a recognition method which works by splitting up overlapping information into atomic units. The mathematical basis of the technique comes from the theory of vector spaces. A diagnosis is determined from the similarity of the test case to the known prototype vectors derived from patients having arteriographically-proven stenosis at known locations. Validity of the diagnosis is estimated from the distance of the test vector from the hyperplane defined by typical vectors.

METHOD

Study Group Preprocessed Thallium-201 scintigraphic data from patients who had undergone T1-201 stress scintigraphy and coronary arteriography comprise the group designated as abnormal. Data were selected from patients who had a greater than 70% diameter reduction in at least one major vessel. There were 35 such patients. The other 25 patients used were those who were found to have entirely normal coronary arteriograms and left ventricular function, and their data comprise the normal group. The two groups together constitute the learning data set.

Thirty features were originally measured on each patient and then preprocessed resulting in a composite of 30 regional percentage values for each patient [1]. The ten regions for each view: anterior (ANT), lateral (LAT) and 30 degrees left lateral oblique (LAO) are shown in Fig. 1.

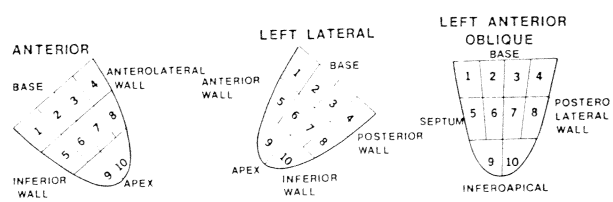

Fig. 1. Scheme of the numbering pattern for each region.

The location and severity (as a percentage of diameter obstruction) of coronary artery obstructions were determined by arteriography, which had been previously read and reported by two cardiologists.

The set of abnormal patients is divided into three groups, corresponding to LAD, RCA and CCX, depending on the coronary artery having the greatest single severity of stenosis.

Algorithm Our goal is to find the percentage severity of the obstructions in LAD, RCA and CCX arteries. The arteries supplying blood to the walls of the heart represent here the "sources"

0276-6574/88/0000/0123$01.00

of the three signals. The "intensities" of these sources depend on severity of stenoses in the arteries.

Let us assume that we have $k > 1$ sources of signals. We also assume there is no opportunity to change the signal intensity. Thus in each measurement the signals overlap. How can we find how each of the sources influences a measurement?

Clearly we have to know the relation between the intensity of each source and the features of the measurement vector. Throughout we assume the following properties of the signals.

(A1) The intensity of signals from separate sources has a multilinear effect on a measurement vector.

Let us consider three sources of signals A, B and C and a measurement $M = X + Y + Z + N$, where the components X, Y, Z, N are induced respectively by sources A, B, C and a noise N. Then, by increasing the intensity of signals from the sources A, B and C, a-, b- and c- times respectively a measurement $a X + b Y + c Z + N$, is obtained [3]. Let us also notice that in almost all physical measurements assumption (A1) is satisfied, in particular in the case of radiation, provided that the interference between the signals is not essential. The second assumption concerns noise.

(A2) If a measurement is repeated in the same configuration of sources then the noise components (the time and intensity of the sources may be different) are approximately the same in each measurement.

To solve the problem at least k independent features in each measurement are needed. This observation concerns the technical part of an experiment (the measurements can be always, at least theoretically, set up better) rather than the assumptions (A1) and (A2) which have more fundamental character.

Finally, let us notice that the problem would not be very difficult if all measurements were precise. However, in our approach we assume that there may be essential errors in measurements because, in reality, assumption (A2) is not entirely satisfied. Thus it contributes to the error. Therefore, a statistical approach to the problem is used.

To be more specific let us assume that we have $k > 1$ sources of signals and in each measurement $n > k$ features are measured. So a measurement can be represented as a vector $M = (m_1, \ldots, m_n)$ of n real numbers.

We also assume there is a fairly big set S of results from similar experiments (i.e. the elements of S are n-dimensional vectors) and for some subset S_o of S we have our problem solved, i.e. we know the intensity of each source (i.e. for each S from S_o we have a vector $V_S = (v_1, \ldots, v_k)$ describing intensity of sources 1, 2, . . . , and k respectively). We also assume the existence of a set Z of results for which the intensity of all sources is equal to zero. This set will be used for noise estimation. The set S_o along with set Z constitute learning data set for the method.

Step 1. Normalizing the measurement vectors.

All measurement vectors from set S are first normalized, in order to take care of the situation where there is a "filter" between the sources of the signals and the measurement vectors. Such a filter acts by scaling measurement vectors which has an undesirable effect on measurements. In our case the action of the filter may be caused by different thickness of a heart wall, different average diameter of arteries, etc.

Step 2. Elimination of noise.

Assumption (A2) is used to estimate a noise vector N. For that, the average vector representing noise N, of the set of measurements Z for which intensities of the sources represent standard level is calculated. We will be looking for deviation from this standard level. Then, N is subtracted from all measurement vectors thus yielding the set of noise-free measurements.

Step 3. Finding typical vectors for the sources.

Using assumption (A1) we can find prototype vectors. To do that we notice that for each vector S from S_o the following equality holds

$$v_1 X_1 + v_2 X_2 + \ldots + v_k X_k = S \qquad (*)$$

Using the assumption that at least k of them are linearly independent, we decompose the set of all equations from the set S_o into the k disjoint families. Each family corresponds to the group of equations for which the intensity of one source is bigger than those of the others. For each family we find an average linear equation of the form (*). To find typical vectors $X_1, \ldots, X_k$ a system of k linear vector equations with k variables is solved.

Step 4. Solution of the general problem.

Now, in order to find the intensity of each source for a new measurement vector $S = (s_1, \ldots, s_n)$ we have to solve the vector equation:

$$v_1 X_1 + \ldots + v_k X_k = S$$

with respect to $v_1, \ldots, v_k$. The numbers $v_1, \ldots, v_k$ describe the intensity of the sources 1, . . . , k respectively. To achieve this we project vector S onto the hyperplane H_o spanned by prototype vectors. In order to give a simpler formula for calculating the projection, let us suppose again that we have only three sources of signals and, for clarity, let us call prototype vectors X, Y and Z instead of $X_1, \ldots, X_k$. Each of the three represents a typical vector for a source with known intensity. These vectors are calculated in Step 3.

What we are really interested in, though, is the coordinates, as spanned by X, Y and Z, of a projected vector. Let us call these coordinates a, b and c respectively, instead of $v_1, \ldots, v_k$. Then we can write a system of the three equations in the following form:

$$a \sum_{i=1}^{n} X_i X_i + b \sum_{i=1}^{n} X_i Y_i + c \sum_{i=1}^{n} X_i Z_i = \sum_{i=1}^{n} X_i S_i$$

$$a \sum_{i=1}^{n} Y_i X_i + b \sum_{i=1}^{n} Y_i Y_i + c \sum_{i=1}^{n} Y_i Z_i = \sum_{i=1}^{n} Y_i S_i \qquad (**)$$

$$a \sum_{i=1}^{n} Z_i X_i + b \sum_{i=1}^{n} Z_i Y_i + c \sum_{i=1}^{n} Z_i Z_i = \sum_{i=1}^{n} Z_i S_i$$

where S_i is the i-th measurement vector to be recognized. The solution of the above system in respect to a, b and c constitutes the solution of our problem.

Step 5. Estimation of an error.

By the assumption (A1) all measurement vectors should be the elements of one k -dimensional hyperplane H_0 of R^n spanned by prototype vectors $v_1, \ldots, v_k$ and vector N defining the origin. Clearly, because of measurement errors, it will not be the case. We take care of that by using projections in Step 4. Therefore, to estimate the credibility of the results from Step 4 we treat the distance of S from H_0 as a measure of an error.

To define such a measure of credibility we have to find a function, such that, for a vector lying in the hyperplane (distance zero) it should take on a value of one and for distance approaching infinity the value of zero. As a model an exponential function is used. To find its form the ratios $||s-p(s)||/||p(s)||$ are first found for all measurement vectors with known intensities of the sources, where p(s) is the projection of vector S into hyperplane H_0. Then, using these values the parameters of the credibility function can be estimated.

The method is first applied to the learning data set. It consists of 35 patients with arteriographically proven stenoses of which the first 8 patients had CCX stenosis, next 18 LAD stenosis and the remaining 9 RCA stenosis. The percentages of stenoses in the three arteries comprise essential information for finding prototype vectors for each coronary artery disease i.e. LAD, RCA and CCX.

In the first step, the whole learning data set of 60 patients is normalized. Then, in order to eliminate the unavoidable noise the average vector of 25 normal patients is calculated and subtracted from abnormal patients data.

To calculate prototype vectors (Step 3) we use the subset of 35 abnormal patients with arteriographically proven percentages of stenoses. This set is divided into three subsets corresponding to CCX, LAD and RCA obstructions. We add percentages of stenosis for each group and the corresponding measurement vectors. As a result, we obtain a system of three equations similar to system (**). The difference being, that the percentages of stenosis a, b and c are known. Solving it results in obtaining prototype vectors.

RESULTS

First, we treated the set of 35 abnormal patients as the test data to check the validity of the method and obtained a diagnosis for each of them. The combined results are given in Fig. 2.

Pt.#	LAD	CCX	RCA		
	Algorithm Diagnosis				
1	-0.16	1.15	-0.48		0.91
2	0.10	1.05	-0.72		0.92
3	-0.32	0.53	-0.23		0.85
4	-0.10	1.00	0.15	CCX	0.92
5	-0.25	0.95	0.12		0.92
6	-0.05	0.83	-0.87		0.92
7	0.39	0.97	-0.13		0.91
8	-0.30	0.70	0.12		0.90
9	0.68	0.24	-0.45		0.93
10	0.55	0.36	-0.32		0.90
11	1.11	0.13	-0.16		0.96
12	0.87	-0.54	0.25		0.94
13	0.87	0.03	-0.23		0.94
14	0.84	0.26	-0.06		0.93
15	0.68	0.02	-0.30		0.92
16	1.12	-0.10	0.30		0.95
17	0.08	0.62	-0.81	LAD	0.92
18	0.84	-0.16	0.13		0.93
19	0.32	-0.60	-0.31		0.92
20	0.93	0.17	-0.48		0.96
21	0.97	-0.09	-0.08		0.94
22	1.01	-0.13	0.10		0.95
23	0.83	0.40	-0.07		0.93
24	0.76	0.16	-0.22		0.93
25	0.87	0.05	0.30		0.93
26	0.69	-0.21	-0.30		0.93
27	-0.12	0.28	0.41		0.89
28	0.08	-0.40	0.96		0.94
29	-0.01	0.10	0.39		0.85
30	0.10	0.16	0.84		0.94
31	0.05	-0.32	0.18	RCA	0.69
32	0.19	0.06	0.69		0.92
33	-0.18	0.69	0.45		0.93
34	0.07	-0.37	1.08		0.95
35	-0.08	0.00	-0.06		0.43

Fig. 2. Results of recognition of learning data.

Fig. 3. Credibility function.

The largest number for each patient represents the most severe stenosis according to the algorithm.

Next, using the credibility function as a measure of credibility of results from Fig. 2, we find values shown in Fig. 3.

Looking at Fig. 2, we notice that only one LAD patient, number 17, was misrecognized as CCX and two RCA patients, numbers 33 and 35, were misrecognized again as CCX ones. The diagnostic accuracy, for the learning data, is 91% with sensitivity and specificity for each group shown below:

Specificity		Sensitivity
100%	LAD	94%
89%	CCX	100%
100%	RCA	78%

Finally, since any technique is only as good as its performance in real-life situations, the method was tested on a test scintigraphic data with unknown stenoses. There were 12 such patients selected by cardiologists for testing the system. Results are shown in Fig. 4 along with credibility values shown in Fig. 5.

	Algorithm Diagnosis			
Pt.#	LAD	CCX	RCA	
1	0.83	-0.14	0.93	0.92
2	1.22	-0.06	0.37	0.97
3	1.05	-0.09	0.61	0.94
4	0.07	0.31	0.72	0.41
5	0.07	0.54	0.71	0.36
6	1.09	0.04	0.28	0.97
7	0.85	-0.51	0.77	0.92
8	0.67	1.05	0.29	0.91
9	0.59	0.15	0.99	0.89
10	0.63	-0.49	1.11	0.88
11	0.23	-0.30	1.21	0.72
12	0.92	0.28	0.01	0.95

Fig. 4. Results of recognition of test data. Fig. 5. Credibility function.

After these results were produced by the recognition algorithm the cardiologists revealed the correct classification, as proved by arteriography, which is shown in Fig. 6.

Pt.#	LAD	CCX	RCA
1	0.60	0.50	1.00
2	0.80	0.25	0.25
3	0.99	0.00	0.00
4	0.00	0.00	0.90
5	0.00	0.00	1.00
6	0.90	0.00	0.00
7	0.00	0.00	0.99
8	0.25	0.50	0.95
9	0.50	0.00	0.95
10	1.00	0.25	0.25
11	0.00	0.00	1.00
12	0.70	0.00	0.00

Fig. 6. Percentages of stenoses for test data.

From Fig. 4 we see that two RCA patients, number 7 and 8, were misrecognized: one as LAD and one as CCX. Only one LAD patient, number 10, was misrecognized as RCA. The specificity and sensitivity for the test data are shown below:

Specificity		Sensitivity
86%	LAD	80%
67%	RCA	71%

DISCUSSION

The method gave reasonable results despite the fact that the learning data set, from which the prototype vectors were derived, constitutes only a small statistical sample. Typical vectors for RCA stenosis were estimated from only 9 learning patients and for CCX from 8, as opposed to 18 for LAD patients. This is due to the fact that LAD stenosis is more common than the other two stenoses. Unfortunately, it is much more difficult clinically to distinguish between RCA and CCX than to distinguish between LAD and either RCA or CCX. This is partly due to the overlapping of the anatomic regions supplied by those two arteries.

A potential problem is created by the reliability of the arteriographic data. First, without a quantitive cineangiographic technique, estimates of the severity of coronary artery stenosis may be rough [4]. Also, if collateral blood flow is supplied from one normal artery to the perfusion bed of a stenosed artery, the resulting scintigraphic perfusion defect may not be typical [5].

Another problem encountered is the scintigraphic data itself, which may not be totally consistent for a specific diagnosis. For example, the scintigram from patient number 10, of the test data, misrecognized by this system was also misrecognized by the cardiologist. If this patient were to be removed from the test data, then both the specificity and sensitivity for test LAD patients would rise to 100%. Since the gold standard we use for comparisons of the results is coronary arteriography and not the cardiologist's diagnosis we have reported this patient as being misrecognized.

The basic ideas of the method are simple and easy for computer implementation. Despite the small size of the learning data set they give better results than any of the fuzzy clustering techniques used on the same test data set or by the expert system [6] at its present state. The method may also be applicable to solving many other diagnostic problems which can be formulated in numeric terms.

REFERENCES

1. Goodenday, L. S. et. al.: Prediction of the Site of Coronary Artery Obstruction from Tl-201 Scintigrams by a Quantitative Computer Technique. Proc. IEEE Conf., Computers in Cardiology, 1981.
2. Cios, K. J. and Goodenday, L. S.: Design and Testing of a Classification System which Recognizes Coronary Stenosis by Site and Relative Severity, Using Myocardial Tl-201 Scintigrams. Proc. IV Int. Conf. Image Analysis and Processing, Cefalu, Italy, Plenum Press, 1987, in press.
3. Hoffman, K. and Ray, K.: Linear Algebra. Prentice-Hall, 1964.
4. Berger, B. C. et. al.: Effect of Coronary Collateral Circulation on Regional Myocardial Perfusion Assessed with Quantitative Tl-201 Scintigraphy. Am. J. Cardiol, 46:365, 1980.
5. Brown, B. G., Bolson, E. L. and Dodge, H. T.: Arteriographic Assessment of Coronary Atherosclerosis: Review of Current Methods, their Limitations, and Clinical Applications. Arteriosclerosis, 2, 2, 1982.
6. Cios, K. J., Fraker, W. A., Freasier, R. E. and Goodenday, L. S.: Hybrid Diagnostic System: Applications in Coronary Arterial Stenosis. Proc. IEEE Conf., Computers in Cardiology, Leuven, Belgium, 1987.