3—2 A Proposal of Multimedia Intelligent Databases for Medical Diagnosis

Toshio MODEGI Communication & Information Operations,
Dainippon Printing Co., Ltd.
Shun-ichi IISAKU Communications Research Laboratory,
Ministry of Posts & Telecommunications

Abstract: For constructing an intelligent multimedia database system for medical diagnosis, we are focusing on two technological points. One is a retrieval algorithm of databases, and the other is a coding algorithm of multimedia contents. For the first, previously we proposed a front-end database preprocessor called "keyword-network", and in this paper we present its extended model providing an intelligent logical AND searching function especially for medical differential diagnosis. For the second, we present examples of multimedia intellectual coding methods for cardiovascular examination records.

1 Introduction

For constructing a medical diagnosis system, we are focusing on the following basic diagnosis steps[1]:

(1) identifying diseases of a facing case.
(2) evaluating illnesses of the identified disease.

These statements tell us, although we can learn diseases from textbooks, we cannot learn illnesses without studying a lot of cases with patients.

In order to enable these two points, we are considering two kind of retrieval functions of databases as follows:

(1) Searching the most associated past cases with a facing case.
(2) Browsing multimedia examination records related to the past cased, and comparing them to those of the current case.

For the first item, we propose locating a network structure of keywords called "keyword-network"[2] at the front-end of a database. This structure is consist of multiple nodes and links which indicate keywords and the relations between these keywords. Moreover, the parameters of the network would be changed or the network would be restructured by users' retrieval operations. In this paper, we propose also an extended logical AND searching function on this model.

For the second item, we propose multimedia intellectual coding algorithms, which will not only decrease the amount of archived data but also help users' understanding of browsed data. Through a medical application, we show several examples of intellectual coding algorithms for cardiograms.

2 Model of an Intelligent Multimedia Database

Our proposed multimedia database is based on a front-end of database technology and a database coding technology. The former enables a rapid retrieval of the target file from distributed databases and the latter enables a rapid search of the target point in the retrieved file. The following description is concepts of these two technologies.

2.1 Structuring of an Intelligent Multimedia Database

The multimedia source materials are digitized and classified to multiple types of data so that each type of data will be individually encoded by using the most suitable coding method. As this classification is based on the contents of source materials, even though the digitized format of source materials were the same, they could be encoded by different algorithms. For example, the intellectual coding algorithms used in X-ray images and UCG images are different each other.

While the encoded data are archived into a database, several keywords are selected out from the content of the data, and these keywords are edited into a keyword-network.

2.2 Retrieval of an Intelligent Multimedia Database

Figure 1 shows the retrieval process of an intelligent multimedia database. Giving an initial keyword, users can begin surfing the keyword-network. if they find some node which interests them, they can browse the contents linked to the nodes. In this case, the decoder or browser program relating to the contents will be down-loaded at first, then this program sends the encoded data to the specified multimedia data generators for initiating play-back operations. While this browsing operation is being progressed interactively by users' instructions, another keyword relating to this browsed data can be found out. Users can try searching again using this related keyword into the keyword-network.

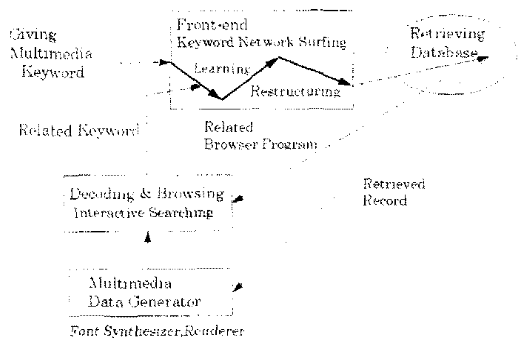

Figure 1: Retrieval of Intelligent Multimedia Database

These each user's accessing logs update the parameters of the keyword-network every time, and in some case restructure it. Therefore by watching the parameters of each user's keyword-network, you can recognize each user's working efforts.

3 Front-end Technology

Medical diagnosis are progressed by the past knowledge that a disease causes some fixed symptoms and signs which can be detected as clinical findings extracted by several examinations. Therefore, a data model that findings and diseases are linked with some weight seems to be preferred for medical diagnosis. This model is similar to what we has proposed generally for a front-end of multimedia databases called "keyword-network"[2] and its extension to distributed environments called "distributed keyword-networks"[3]. By this distributed model, we have also prototyped the database of cardiovascular cases.[3] In this section, we will overview the basic functions of a keyword-network.

3.1 Elements of a Keyword-Network

A keyword-network is expressed as a traditional graph which consists of nodes and links. Each node represents a group of keywords described below, and link defines some relation between groups. In a distributed model, a group of nodes is defined as a class-node and a group of links is defined as a class-link, in this sense, each node is called a instance-node and each link is called a instance-link. Figure 2 shows an example of distributed keyword-networks.

3.2 Definition of Keywords

Each node has at least one representative keyword and several equivalent keywords[3], which are the extended expression of synonyms and indicate the same functional object as the representative does. Certain equivalent keyword has the address data pointing to some record of databases. The term keyword is defined as an abstract expression of source materials, and not necessarily consists of character codes. The binary data such as reduced images, icons, sound clips and reduced video clips can be treated as keywords.[3]

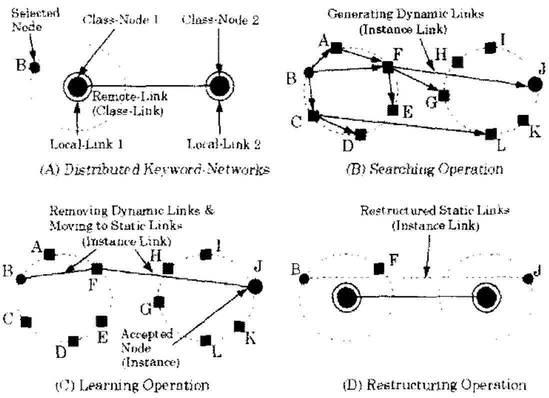

Figure 2: Action of Distributed Keyword-Networks

3.3 Relation of Keywords

Instance-links between instance-nodes called dynamic link[3] are generated based on matching of keywords in both instance-nodes while referring to a thesaurus defined by those class-link. In Fig.2-(B) example, 9 dynamic links are generated from user's selected node B. In medical terms, are categorized by the following three types of keyword[1]:

(1) nominal which is expressed as generic name such as "hypertension".
(2) ordinal which is expressed with grade or stage such as "hypertension is ++".
(3) metric which is expressed with data such as "blood pressure is 180/100 mmHg".

In terms of our proposing thesaurus dictionary called "numeric thesaurus", which has keywords such as "hypertension : blood pressure is greater than 150/90.", the item (1) and (3) keyword above can be linked by mathematical matching.

3.4 Searching Operation

After generation of dynamic links is complete, a unit virtual signal to the initial node is given, and it transfers and fans out the signal up to nodes of specified hops. A calculated value of transferred virtual signal level of each node determines to become a candidate and their priority when a listing of the candidates is presented to the user.
In Fig.2-(B), node F, A, C, D, E, G, J and L are the candidates.

3.5 Learning Operation

Selecting an accepted node from these candidates, weight values regarding the accepted paths between the initial node and the accepted node will be increased, while the other values of the paths regarding the disregarded candidate nodes will be decreased. Then dynamic links whose weight value increased become static links, which will be used for next searching operation, and the other dynamic links are removed. In Fig.2-(C), by the accepted node J, two links between node B and F, and F and J have become static links.

3.6 Restructuring Operation

After similar operations to those described above are done several times, weight values of instance-nodes and remaining static-links can become unbalanced. By predetermined condition, low-weight unused instance-nodes or static-links will be removed, and the network will become more simple. In Fig.2-(D), the intermediate node F between two static links is removed.

4 Front-end Technology for Medical Diagnosis

The problems on diagnosis are that most diseases cannot be identified by some unique finding, moreover in some case, multiple diseases tormenting a patient must be identified simultaneously. For diagnosing these cases called such as differential diagnosis[1], we propose an intelligent logical AND searching operation on the keyword-network, by providing multiple keywords one after another. This function also provide to the keyword-network, more independence of the database designer's intention and capturing more users' interests by automatic addition feature of static links.

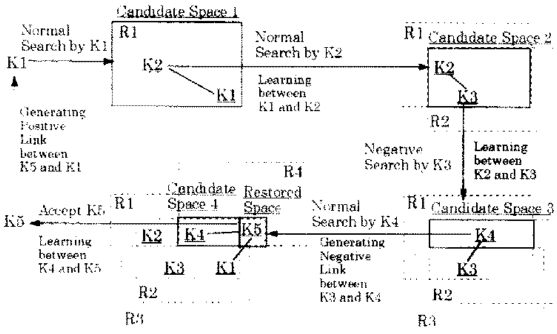

Figure 3: Logical AND Searching Operation

4.1 Basic AND Searching Operation

In Fig.3, each region indicates a set of keywords or nodes, which are such as clinical findings, and diseases, and the candidate space is showed to become narrow by repeating a keyword search. At first giving the keyword K1, we can retrieve the region R1, which is the same as the candidate space 1.

Selecting K2 from that, the link between K1 and K2 is made learn, in other word, the weight of the link is increased. Then using K2, we can retrieve the region R2, and the candidate space 2 is obtained by a logical AND operation between R1 and R2. This operation is basis for differential diagnosis.

4.2 NOT Searching Operation

From the candidate space 2, selecting K3, the link between K2 and K3 is made learn, then using K3, we can retrieve the outside region of R3 by a negative search, which is done by giving to the node of K3 a negative virtual signal. The candidate space 3 is obtained by a logical AND operation between the candidate space 2 and Not-R3. This operation is useful for a case that some finding can be obviously excluded.

4.3 Consolation Match

From the candidate space 3, selecting K4, the negative link between K3 and K4 is created because this link does not exist. Then using K4, we can retrieve region R4, and the candidate space 4 is obtained by a logical AND operation between the candidate space 3 and R4.

The right square to the candidate space 4 is consistently inside of R2, Not-R3 and R4 by recent three searching operations, it might be added to the candidate space 4, although it is outside of R1. In this case, selecting K5 from this restored space, not only the link between K4 and K5 is similarly made learn, but the link between K5 and K1 is newly created. This operation is useful for diagnosing complications or secondary diseases which cause mixture of findings.

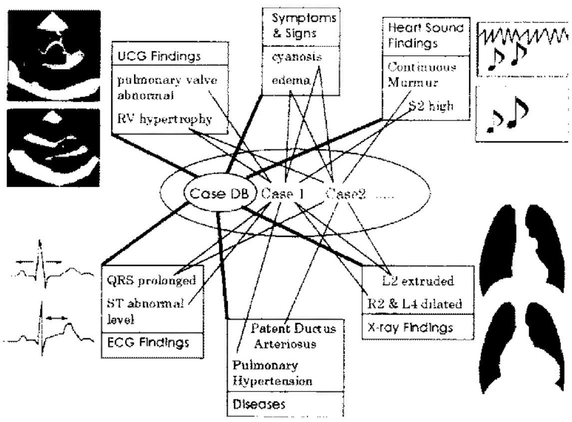

Figure 4: Network Structure of Cardiovascular Diagnosis Application

4.4 Cardiovascular Diagnosis Application

Figure 4 shows a network structure of a cardiovascular-diagnosis application based on reported in [3]. All of the keywords in this application are selected from reference [6]. Six class-nodes, which are UCG findings, X-ray findings, ECG findings, heart sound findings, symptoms and signs, and diseases, are star-shaped connected to a main class-node, cardiovascular cases. Each class-node links with a specified database, especially, the class-node of cardiovascular cases links with examination records archived by variety of database formats described later.

Giving a clue-keyword showing some clinical finding regarding a facing case, a user can retrieve associated cases and other findings related to the retrieved cases, and compare various multimedia data of the retrieved cases to data of the current. Adding the other clinical findings, we can determine the most associated past case and identify its disease.

5 Intellectual Database Coding Technology

In this section, we are focusing on contents of databases, a multimedia intellectual coding method which is extending the previously proposed methods especially for image data [4], to multimedia data.

5.1 Concept of Intellectual Database Coding

The following two points are our definition of multimedia intellectual coding.

(a) Content-Based Coding:
The contents written in data will be coded in the format specified by each data type, whereas the conventional coding methods are based on sampled electronic signals.

(b) Signal-Generating Decoding:
The wave of the similar signal to the original will be created by passing the encoded data to the generator module specified by each type, whereas the conventional decoding methods are just re-forming the wave of the signal.

Currently, there are four types of multimedia data, text, speech sound, musical sound and CG (Computer Graphics), which meet our definitions described above.

Text is coded by characters and its decoding generators are fonts.

Speech sound is also coded by characters and its decoding generator is voice synthesizer.

Musical sound is coded by MIDI (Musical Instruments Digital Interface) and its decoding generator is MIDI synthesizer.

CG Image (2D or 3D, still or motion) is coded by model-based[4] such as PostScript, Open-GL and VRML. its decoding generators are RIP (Raster Image Processor) and graphic renderer.

5.2 Features of Intellectual Database Coding

As no signal data are included in encoded data, we can easily understand its encoded code-size will become very small. We have listed up the excellent features regarding to this method as follows:

(1) Extremely Low Bit Rate.
(2) Resolution Free.

(3) Editable Decoding.

As for the second item, there are no concept of quantization and sampling intervals, or more specifically concept of pixels in case of image coding. Therefore, you can decode images at any types of resolution using the same encoded data, and those frame rates are also variable. In a musical data case, you can play back data at any tempos. As for the third, you can easily modify the contents of encoded data, so it is useful to prepare a variety of materials using the same encoded data.

6 Intellectual Coding Methods for Cardiograms

Compared with the other medical clinical subjects, there are some special features in a cardiovascular diagnosis subject that time-based examination records are often used in order to analyze dynamic functions of hearts. Owing to those features such as using motion pictures and sounds for diagnosing, this application can become a good benchmark of multimedia databases. We are using the following general types of multimedia formats.

(1) Text document: shift-JIS coded plain text

(2) chemical analysis: tabulator tool format

(3) X-ray, pathologic pictures: bitmap, JPEG still image

(4) UCG, CAG: compressed video

(5) PCG: ADPCM, wave-form sound

However, in case of time-based recorded cardiograms, we will face a difficulty to search and browse points of interests, therefore an innovative method that some intellectual attributes are expressed in encoded data is expected. In this paper, we propose a MIDI coding method for PCG (Phonic-Cardiogram), QRS Coding Method for ECG (Electro-Cardiogram), and M-mode Coding Method for UCG (Ultra-sonic echo Cardiogram).

6.1 MIDI Coding Method for PCG

For coding the features of heart-sound rhythms into MIDI format, there are three types of codes for encoding each MIDI note: strength, time position and tone position of sound. We can extract these three codes at each note from the PCM sampled data of heart-sounds as follows:

(1) Time position: Extracting of uttering section. Getting the starting and ending time of uttering section and converting the data to the unit of quantization. This one section is encoded as one note

(2) Strength: Detecting of a local peak in this section. Converting the peak amplitude of the section to 256 levels of strength.

(3) Tone position: Analyzing frequency Getting the basic frequency of this section and converting the data to the name of note, C, D, E, F, G, A and B. As one sequence of these three can generate data of one note, this sequence must be repeated until the end of the sampled signal.

However, if several phrases of encoded MIDI codes are almost same for several periods, encoding sequence can be skipped, instead the number of skips will be recorded. This skipping is not only saving the amount of coded data but also convenient for diagnosing because the uninterested sections for diagnosis could be removed beforehand.

Figure 5, which is a kind of musical score for MIDI, shows an example of three cases which have the significant feature of heart rhythms. The MIDI codes in this chart are encoded by us manually based on the PCM data in reference[5]. Like shown in Fig.5, adding physiological attributes[5] such as S1 and S2 to the encoded data, helps diagnosing.

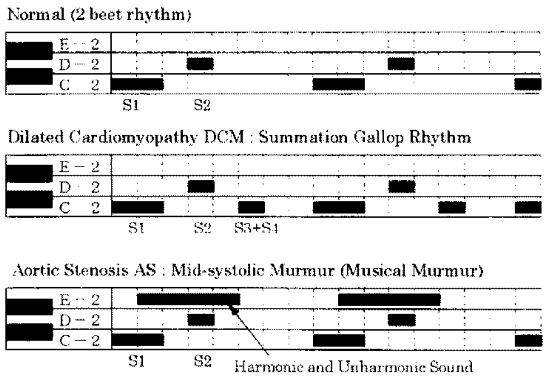

Figure 5: MIDI Encoding Examples of PCG

6.2 QRS Coding Method for ECG

QRS detection is a fundamental analyzing technique for ECG in daily diagnosing activities.[6] As shown in Fig.6, we propose encoding the coordinate data of the standard diagnosing points P, Q, R, S, T and U, and the intermediate points P1, P2, Q1, J, T1, T2, U1 and U2, which are defined by us for decoding smooth wave patterns, into a particular graphic format. Here, the point J is also used for diagnosing. Moreover, if time periods between several waves are almost same, encoding sequence of waves can be skipped similarly to MIDI coding. This encoding method applies to all 12 types of different signals used in a standard ECG examination.

For decoding to a wave signal, a popular Bezier interpolation technique can be applied. Except in the case of high-frequency noises included in the original signal, we can get similar waves as the original. If 12 signals described above were recorded synchronously, we could plot a complete three-dimensional vector-ECG using computer-graphics technique.

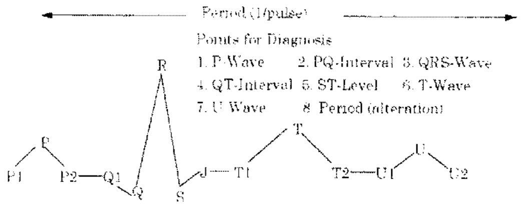

Figure 6: QRS Encoding of ECG Signal

6.3 M-mode Coding Method for UCG

An ultrasonic echo tomograph is viewed as arc-shaped images and seems to be suitable to a polar coordinate system. If we extract scan-lines of the same sector from every frame and locate them in the horizontal direction, we can get a M-mode image like shown in Fig.7. Similarly, changing a extracted position of a sector little by little, we can get M kinds of M-mode images.

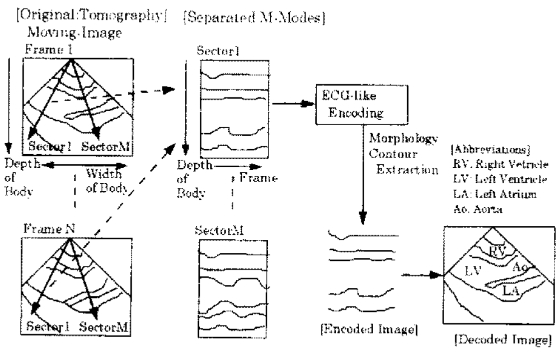

Figure 7: M-Mode Encoding Flow of UCG Images

By processing extraction of contours at each M-mode image, ECG-like wavy patterns will appeared on it. Here, we can encode these patterns by the similar method as described in 6.2. As each wavy pattern indicates the edge of some organ, we can add its anatomical attributes to the wavy pattern, then its decoded image can be presented with the captions like shown in Fig.7. In this method decoded moving images will be composed of only contour lines, and different from the original images, but they are sufficient and friendly to diagnose except in the case of showing hazed echoes such as a thrombosis case or showing color Doppler echoes such as a regurgitated blood stream case.

Moreover, encoded data are based on the M-mode format, then we need not record additionally this mode of images for a minute measurement as the conventional way.

7 Conclusion

In this paper, we have proposed two concepts of constructing multimedia databases for medical diagnosis, and as an example of applying these technologies we have presented a prototype of a database on cardiovascular cases.

Through making a prototype of this application, we have confirmed a technological feasibility of applying our proposed concepts. However, it is not currently clear whether the effect of our proposed concepts will become what we expect, therefore we must evaluate this prototype including the practical databases partnered with some staff in such as a medical school.

Currently we are progressing applying a concept of keyword-network with our proposed models to several other kinds of applications especially for CAI, and we are going to evaluate suitability of each educational application to several concepts of keyword-network.

References

[1] M.H.Delp and R.T.Manning, "Mayor's Physical Diagnosis 9E -An introduction to the clinical process-." Igaku-shoin/Saunders Ltd., (1984)

[2] S.Iisaku and T.Modegi,``An intelligent retrieval algorithm for image archiving system," Proc. IEICE General Conference, .D-62. (Mar.,1996)

[3] T.Modegi and S.Iisaku. ``A proposal of the CAD for intelligent multimedia databases and its application." Proc. IEEJ Electronics, Information. and System Society Convention. B-12-9.(Sep.,1996)

[4] H.Harashima, ed., ``Image Information Compression," ITE, Ohm-sha. (1994)

[5] T.Sawayama, ``Cardiac Auscultutions -- Exercise with Compact Disc," Nankodo, (1994)

[6] ``Clinics Today," CD-ROM edition. vol.5. Igaku-shoin. (1995)