Frameworks for *u* Health Care System Using Fuzzy Functions

Haeng-Kon Kim[1]

***Abstract*— In this paper, we show an intelligent disease diagnosis system for public. Our system deals with 30 diseases and their typical symptoms selected based on the report from Ministry of Health and Welfare, Korea. Technically, our system uses a modified FCM for clustering diseases and the input vector consists of the result of user-selected questionnaires. Our modified FCM improves the quality of clusters by applying symmetry measure based on the fuzzy theory so that the clusters are relatively insensitive to the shape of the pattern distribution. Furthermore, we extract the highest 5 diseases only related to the user-selected questionnaires based on the fuzzy membership function between questionnaires and diseases in order to avoid diagnosing unrelated disease.**

***Index Terms*—*u* Health Care , Fuzzy Function, Target Diseases, Symptoms**

I. Introduction

The medical resources accessibility for patients in rural area and low-income people becomes more and more important. Not only they need timely medical treatments, but also, and more importantly, they need early diagnosis of possible medical malfunctions.

Web-based medical information software is a convenient and useful tool to reduce the constraints of people with regional or economic deficit [1]. Among many possible web-based medical information systems, an intelligent health diagnosis system is one of the most needed software. It can greatly reduce the communication procedure between patient and doctor so that the health care system is able to respond fast predictable health evaluation to most needed people [2]. Early diagnosis of serious diseases with such systems can improve the effectiveness and accessibility of health evaluation and avoid long-term medical treatment [3].

In this paper, we propose an intelligent disease diagnosis system for public based on the report from Ministry of Health and Welfare, Korea [4, 5]. These reports summarize information about frequently attacked diseases and burdensome diseases for Koreans. Our system tries to give a simple user interface and usage for the public who may not have enough knowledge related to their health deficit. Technically, we use modified FCM and fuzzy membership function in order to give only symptom related qualitative diagnosis.

The organization of this paper is as follows. In section 2, we explain the target diseases of our health diagnosis system and main technical improvements. Then, we show the snapshots of our experiment in section 3. Section 4 summarizes the concluding remarks.

II. The Proposed System

A. Target Diseases and Symptoms

We select 30 diseases and corresponding symptoms based on the national report "Burdensome Diseases of Koreans 2005". According to another related report, the most burdensome disease for Korean is diabetes mellitus and followed by stomach ulcer, duodenal ulcer, asthma, cerebrovascular accident, rheumatoid arthritis, and depression [4]. There exist clear differences among age groups for the most frequent dangerous diseases. For twenties, the most frequent dangerous disease is depression but it is stomach ulcer or liver cirrhosis for the thirties and forties and cancer for fifties and sixties and myocardial infarction or cerebrovascular accident for elderly people [5]. It has clear relationship with the most frequent fatal (causing death) disease report [4] that picks up cancer, cerebrovascular accident, cardiovascular disease, respiratory disease, diabetes mellitus, and liver disease in order.

In this paper, we collect symptoms of target diseases from 20 medical contents on the web/text and design questionnaires for users. We classify representative symptoms for each target disease and designate abnormal part display considering relationship between symptoms and body part. Figure 1 shows the structure of the system related to the body part display and symptom questionnaires.

B. Overall Structure

In our system, we provide 13 possible abnormal body parts in body-figure and the user select a part that needs diagnosis. Then, the system gives questionnaires for symptoms related to diseases of that part.

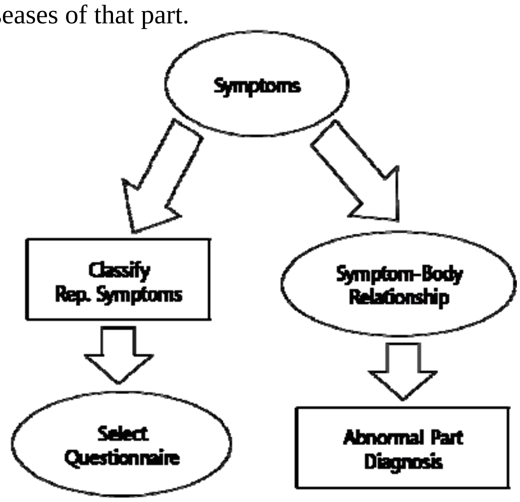


Figure 1: Questionnaire and Diagnosis

Department of Computer Engineering, Catholic University of Daegu, Daegu, Korea 712702
hangkon@cu.ac.kr

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

The results of questionnaires are represented as a vector. Then the system computes the membership rate for each possible disease according to questionnaire results. Then our modified FCM algorithm computes the similarities to patterns only related to those questionnaires selected by the user in order to make the final diagnosis.

Figure 2 shows an example of disease diagnosis process. The system consists of four layers – abnormal part, symptom questionnaire, related disease, and output. In related disease layer, the darker the shadow is, the more this cluster includes related questionnaires. However, the original FCM is not accurate enough in this phase. Our modified FCM applies fuzzy theory in measuring symmetry to the result of the original FCM and re-cluster them. This approach works well in our experiment.

C. Modified FCM Algorithm

Most conventional FCM algorithm optimizes the goal function based on the similarity measured by the distance between input vectors and the center of each clusters [9]. These methods share the disadvantage that the accuracy may depend on the distribution of patterns in cluster space [10, 11]. For example, if the shape of pattern distribution is an ellipse or when patterns are distributed to boundaries of clusters and the basic forms are crossed, the Euclidian distance measure is not an appropriate measure of similarities between patterns and the center of each cluster.

Thus, we apply fuzzy theory to the symmetry measure so that the FCM algorithm should not be sensitive to the shape of the cluster, positions of cluster centers, or the change of cluster distances. The symmetry measure of our modified FCM algorithm is shown in equation (1)

$$Symmetric(x_i, c) = \max_{j=\forall pattern,\ i \neq j} \begin{pmatrix} (1-\alpha)(1-\dfrac{\deg(x_i, x_j, c)}{180}) \\ -(\alpha * ratio_d(x_i, x_j, c)) \end{pmatrix} \quad (1)$$

where $\deg(x_i, x_j, c)$ is the degree between x_i and x_j subject to point c and $ratio_d(x_i, x_j, c)$ is computed as equation (2).

$$ratio_d(x_i, x_j, c) = \begin{cases} \dfrac{d(x_j, c)}{d(x_i, c)} & if\,(d_i > d_j) \\ \dfrac{d(x_i, c)}{d(x_j, c)} & if\,(d_i < d_j) \end{cases} \quad (2)$$

where $d(x_j, c)$ is the Euclidean distance between x_j and c.

The weight factor α is computed by equation (3) based on the fuzzy theory.

$$\alpha = \frac{d(c_i, c_j)}{\sqrt{D_m}} \quad (3)$$

Let $\mu(x)$ be the result of equation (1). Then the center of the cluster is computed as equation (4).

$$V^{(p)} = \frac{\sum \mu(x)x}{\sum \mu(x)} \quad (4)$$

Then, the final similarity U between input pattern and the center of the cluster is computed as equation (5).

$$U = \sqrt{\sum_{i=0}^{k-1} (x_i - c_i)^2} \quad (5)$$

Figure 3 shows the learning process of our modified FCM algorithm.

D. Disease Membership Based in Symtoms

However, our system has a disadvantage that it can diagnose unrelated diseases to the questionnaires that user selected as shown in Figure 4.

Thus, we apply fuzzy membership function in order to remove unrelated disease diagnosis.

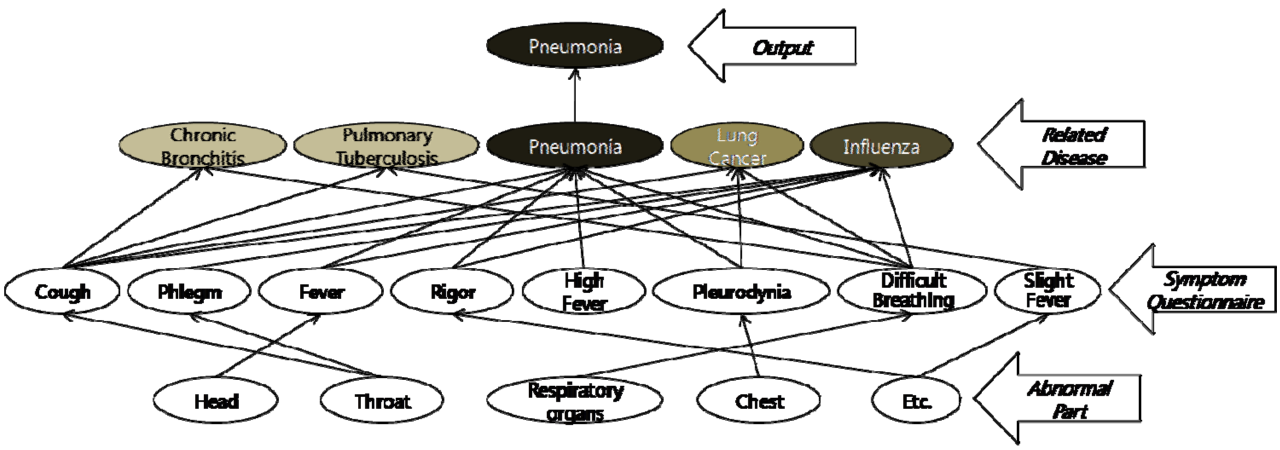


Figure 2. An example of diagnosis process

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

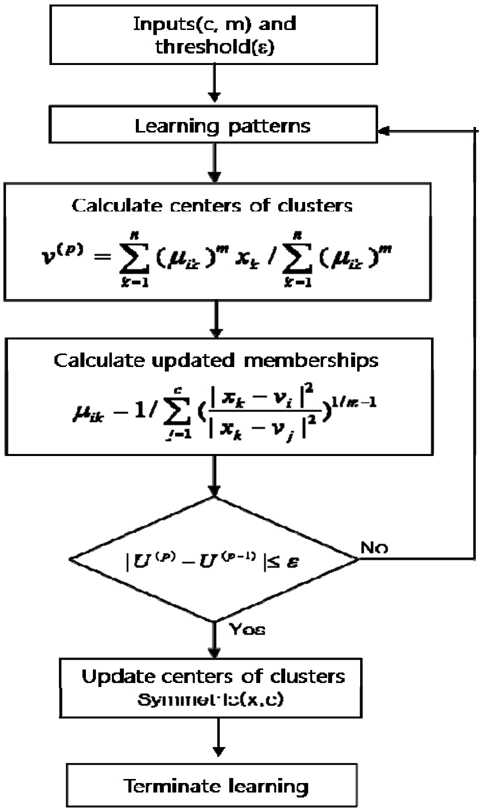


Figure 3: The Learning Process of Modified FCM

The system computes the number of symptoms that are common to the set of possible symptoms of target disease and user-selected questionnaires. Then according to the maximum number of symptoms for each target disease, we compute the membership rate to the related disease based on the proposed membership function. This fuzzy membership function is shown as Figure 5.

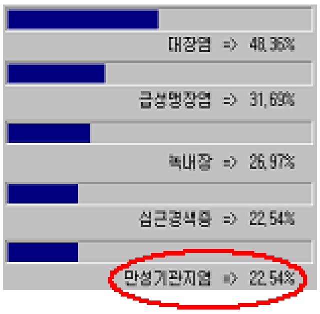


Figure 4: Example Screen of disease diagnosis – red circled disease is not related to user-selected questionnaires.

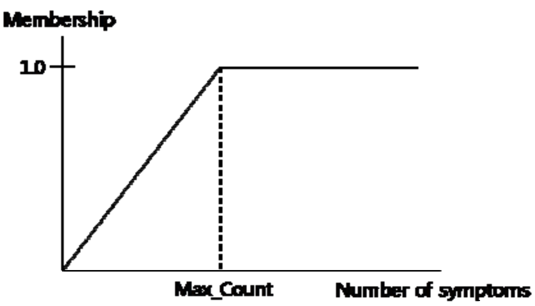


Figure 5: Membership function for each disease

In Figure 5, Max_Count represents the maximum number of symptoms of the target disease. The input of membership function is the number of common symptoms between user-selected questionnaires and the target disease. The fuzzy membership rate is computed as rule (6).

If (input< 0) then

$\mu = 0$

Else If (input > 0 AND input < Max_Count) then

$$\mu = \frac{Max\_Count - Input}{Max\_Count}$$

Else If (input > Max_Count) then

$\mu = 1$ (6)

Our system outputs the five highest similar diseases based on the similarity measure U defined as equation (5) by modified FCM explained in section 2.3 with non-zero membership rate by rule (6).

III. EXPERIMENT AND RESULT ANALYSIS

We used IBM-compatible PC with Intel Pentium_IV 2GHz CPU and 256MB RAM and the software is written in VC++ 6.0. The number of yarget diseases is 30 and the number of questionnaires is 125.

If a user selects a part from the user interface, the system provides symptom questionnaires related to the part the user selects. Figure 8 shows the screen when the user selects respiratory organs and throat.

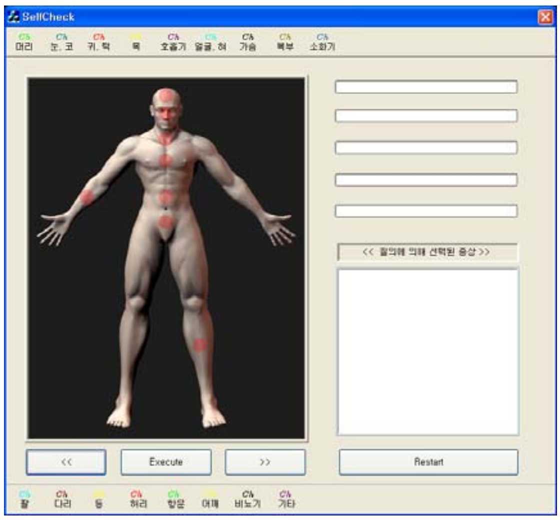


Figure 7: User interface

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

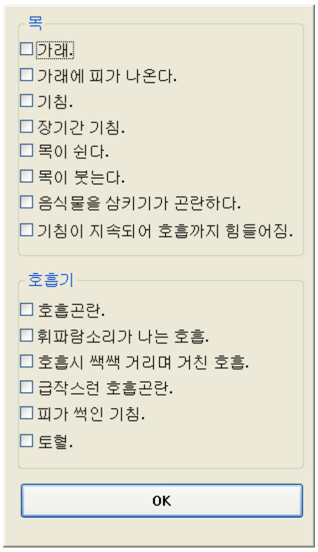
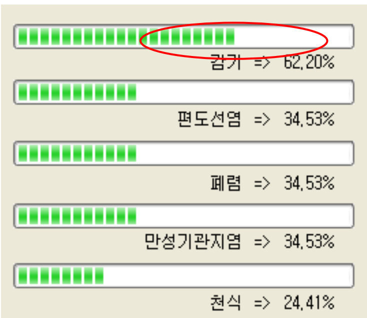


Figure 8: Symptom questionnaire screen

After selecting symptoms, the system outputs 5 highest possible diseases with membership rates.

As explained in section 2.3, we applied modified FCM in order to improve the quality of disease-symptom clustering. When we input the same symptoms as sputum, and dyspnea (difficulty in breathing), two algorithms give different diagnosis shown as figure 9.

(a) FCM output



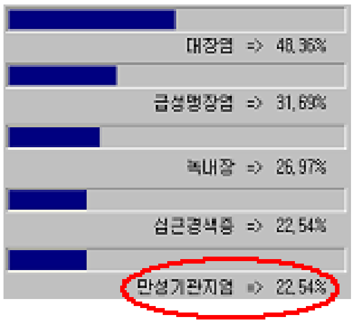
(b) Modified FCM output

Figure 10: FCM modification effect

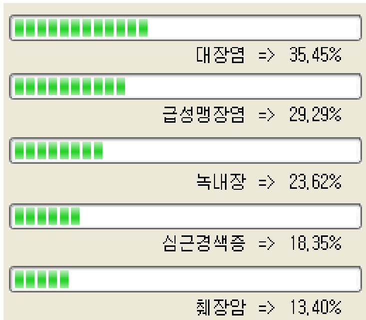
In figure 9 (a), we can find that FCM diagnoses 'herniated inter-vertebral disk' as the third highest possible disease whereas our modified FCM only diagnoses highly related diseases.

However, as explained in section 2.4, without fuzzy membership function, our system can diagnose a disease not related to user-selected symptoms as figure 10 shows. In this example, we input chronic diarrhea, dehydration, vomiting, and stomachache.

The system diagnoses chronic bronchitis that is not related to user-selected questionnaires. This problem is gone when we apply fuzzy membership function and the result with the same input diagnoses pancreatic cancer.



(a) Without fuzzy membership



(b) With fuzzy membership

Figure 10: Fuzzy membership effect

Table 1 summarizes the set of parameters in our experiment. The parameter *m* in the table is the weight index. The pattern consists of symptoms of the target disease whereas the input node is the pattern made from user-selected symptoms. The number of clusters in this experiment is the same as the number of target diseases.

Table 1. Parameters of the Experiment

	# of patterns	*# of input node*	*m*	*# of Clusters*
Modified FCM	**50**	**125**	1000	30

IV. CONCLUSIONS

In this paper, we propose an intelligent health diagnosis system for public use. This kind of system can diminish the accessibility problem of medical resources for people who have regional or economic deficiency. The target disease and symptoms are collected from various sources but based on the national report on death-cause diseases and most burdensome, frequently appearing disease.

We apply modified FCM algorithm in clustering in order to have relatively insensitive clusters to the characteristics of pattern distribution. Also, we apply fuzzy membership function in order to discard diagnosing unrelated diseases.

The experiment shows that our proposed system can respond many complex queries with satisfactory accuracy. The system is successfully implemented and ready to use for web-based medical information system.

ACKNOWLEDGMENT

This work was supported by the Korea National Research Foundation(NRF) granted funded by the Korea Government(Scientist of Regional University (No. R-2009-0067365).

REFERENCES

[1] Mullen K. H. Berry D. L., and Zierler B. K., “Computerized symptom and quality-of-life assess ment for patients with cancer part Ⅱ; acceptability and suability,” Oncology nursing forum, Vol. 31, No. 5, pp.84-89, 2004.

[2] Berry D. L., Trigg L. J., Lober W. B., Karras B. T., Galligan M. L., Austin-Seymour M., and Martin S., " Computerized symptom and quality-of-life assessment for patients with cancer part I; development and pilot testing," Oncology nursing forum, Vol. 31, No. 5, pp.75-83, 2004.

[3] Ford D.F., Pincus H.A., Untzer J, Bauer M.S., Gonzalez J.J., and Wells K.B., "Practice-based interventions," Mental Health Services Research, Vol. 4, No. 4, pp.199-204, 2002.

[4] National Statistical Office, Disease according to Death Cause, Retrieved Oct 1, 2006 from http://kosis.nso.go.kr/cgi-bin/sws_999.cgi

[5] http://www.mohw.go.kr

[6] WHO.ICD-10 Interntional Statistical Classification of Diseases and Related Health Problems, tenth revision, Vol. 2, Geneva, WHO 1993.

[7] [7] Simans, David G., Travell, Janet G. and Simans, Lois S., Myofascial Pain and Dysfunction, The Trigger Point Manual, Vol.1, Vol.2, Williams & Wilkins, 1983.

[8] [8] Arun D. K., Computer Vision and Fuzzy- Neural Systems, Prentice Hall PTR, 2001.

[9] [9] J. Bezdek, "A convergence theorem for the fuzzy ISODATA clustering algorithm," IEEE Trans. Pattern Anal. Machine Intell., Vol.PAMI-2, No.1, pp.1-8, 1980.

[10] [10] Kim K. B., Lee D. U. and Sim K. B., "Performance Improvement of Fuzzy RBF Networks," Lecture Notes in Computer Science, LNCS 3610, Springer, pp.237-244, 2005.

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)