A Study on Multiband FTN Method for improving Transmission Efficiency

Ji-Won Jung
Dept. of Radio Communication Engineering
Korea Maritime and Ocean University
Busan, Republic of Korea
jwjung@kmou.ac.kr

Chang-Uk Baek
Dept. of Radio Communication Engineering
Korea Maritime and Ocean University
Busan, Republic of Korea
cubaek@kmou.ac.kr

***Abstract*— FTN(Faster-than Nyquist) signaling is a technique of transmitting information at a rate higher than the allowed Nyquist limit. In order to improve the transmission efficiency by applying the FTN method according to increasing the interference rate. However, performance degrade arising from ISI(Inter-Symbol Interference). To overcome this problem, we propose the multiband FTN method which is uniformly dividing and allocating the encoded bits into multiband, and transmits the encoded bits using the frequency allocated to each band by applying the FTN transmission method to each band. In addition, it can improve performance through the effect of noise averaging by increasing the number of samples per bit in each band to reduce interference between adjacent channels. We compared the performance by setting the interference ratio to 25%. Through simulation results, we show that the performance of the proposed multiband FTN method is improved about 0.3 dB ~ 0.5 dB than that of single band FTN method.**

Keywords-Faster-than Nyquist; inter-symbol interference; multiband; multiband FTN;

I. INTRODUCTION

Next-generation wireless and/or satellite communications require high transmission efficiency and high reliability to provide various services with subscribers. Recently, many methods for increase of throughput is being researched, as the next satellite broadcast / communication and the 5G based mobile communication demand for throughput is increasing, whilst the bandwidth is limited.

Research on applying the multidimensional modulation method in the modulator/demodulator and speeding up the high-speed channel encoder/decoder to meet the high data rate is already saturated. Thus, research interest has been focused more on transmission methods. However, it is very difficult to improve both throughput and performance, because the two are in a trade-off relationship. Therefore, it is the most important to develop methods which can maintain the performance to the maximum, whilst increasing the throughput.

Recently FTN (Faster-than-Nyquist) method[1,2], which transmits faster than the throughput of Nyquist, is emerging as the standard for the next generation DVB-S3(Digital Video Broadcasting-Satellite Third Generation) and FOBTV(Future of Broadcast Television)[3]. FTN method to increase throughput for limited bandwidth in next generation wireless communication. It is possible to transmit signals faster than the Nyquist rate and maximize throughput with the same channel bandwidth. However, FTN method has a limitation in maximizing transmission efficiency due to performance degradation arising from ISI(Inter-Symbol Interference).

To overcome this problem, we propose the multiband FTN method which is uniformly dividing and allocating the encoded bits into multiband, and transmits the encoded bits using the frequency allocated to each band by applying the FTN transmission method to each band[4-6]. In addition, it can improve performance through the effect of noise averaging by increasing the number of samples per bit in each band to reduce interference between adjacent channels.

II. MULTIBAND FTN TRANSMISSION METHOD

FTN signaling is a technique of transmitting information at a rate higher than the allowed Nyquist limit. In order to improve the transmission efficiency by applying the FTN method according to increasing the interference rate. However, performance degrade arising from ISI.

A. Single band FTN Transmission method

Fig. 1 shows the single band FTN system structure based on turbo equalization.

The source bits to be transmitted bit-stream D is given by

$$\mathbf{D} = \{\mathbf{d_1}, \mathbf{d_2}, \cdots, \mathbf{d_K}\} \quad (1)$$

where k is the size of D, First, D is encoded by the (N, K) LDPC(Low Density Parity Check) encoder[7]. Coded bit stream C is given by

$$\mathbf{C} = \{\mathbf{c_1}, \mathbf{c_2}, \cdots, \mathbf{c_N}\} \quad (2)$$

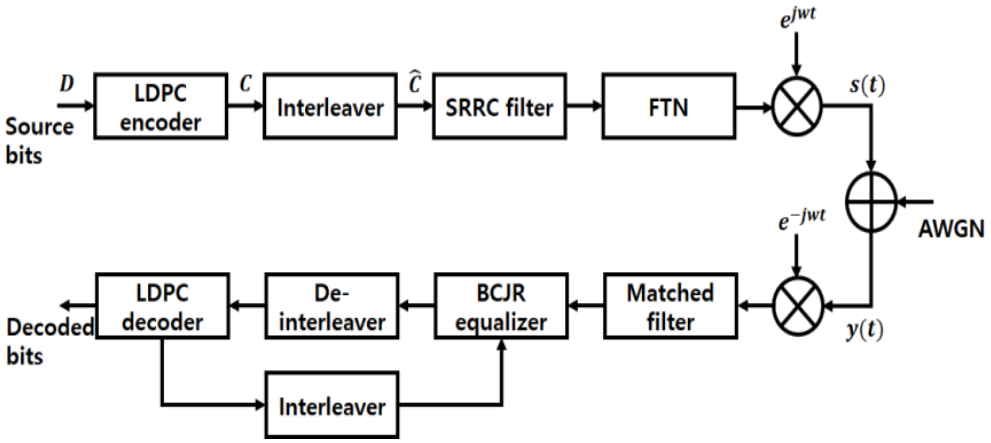


Figure 1. The structure of single band FTN model

The encoded bits passed through the encoder transform burst error into random error via the interleaver. The interleaved output is canceled a posteriori from the proceeding received signal. The interleaving function helps the receiver convergence.

$\hat{\mathbf{C}}$ means the encoded bits has passed through the interleaver.

$$\hat{\mathbf{C}} = \{\hat{c}_1, \hat{c}_2, \cdots, \hat{c}_N\} \tag{3}$$

FTN signaling is a technique of transmitting information at a rate higher than the allowed Nyquist limit. Consequently, ISI necessarily occurred.

Interference transmission signal $\mathbf{s(t)}$ is given

$$\mathbf{s(t)} = \sum_{n=1}^{N} \hat{c}_n h(t - n\tau T) e^{jwt}, \quad \tau < 1 \tag{4}$$

where $\hat{c}_n$ are encoded bit stream after interleaving, $\mathbf{h(t-n\tau T)}$ is a unit-energy baseband pulse, τ is interference time, $\mathbf{T}$ is symbol duration.

The received signal is given by

$$\mathbf{y(t)} = \sum_{n=1}^{N} \hat{c}_n p((n-t)\tau T) + \eta(t) \tag{5}$$

B. Multiband FTN Method

In order to improve the transmission efficiency by applying the FTN method according to increasing the interference rate. However, performance degrade arising from ISI. To overcome this problem, we propose the multiband FTN method which is uniformly dividing and allocating the encoded bits into multiband, and transmits the encoded bits using the frequency allocated to each band by applying the FTN transmission method to each band. In addition, it can improve performance through the effect of noise averaging by increasing the number of samples per bit in each band to reduce interference between adjacent channels.

Fig 2. shows the structure of transceiver for multiband FTN system.

The data that passes through the interleaver of Equation (3) is divided and distributed uniformly into $\mathbf{N_B}$ bands through S/P as shown in Equation (6).

$$\hat{\mathbf{C}}_b = \{\hat{c}_1, \hat{c}_2, \cdots, \hat{c}_{N/N_B}\}, \quad (b = 1, 2, \cdots, N_B) \tag{6}$$

The uniformly divided data is transmitted by passing through a SRRC(Square Root Raised Cosine) filter and applying the FTN method.

Transmission signals given by

$$\mathbf{s(t)} = \sum_{b=1}^{N_B} \sum_{n=1}^{N/N_B} \hat{c}_b(n) h(t - n\tau T_{N_B}) e^{jw_b t}, \quad \tau < 1 \tag{7}$$

Eq.(7) means the transmit data of multiband FTN method.

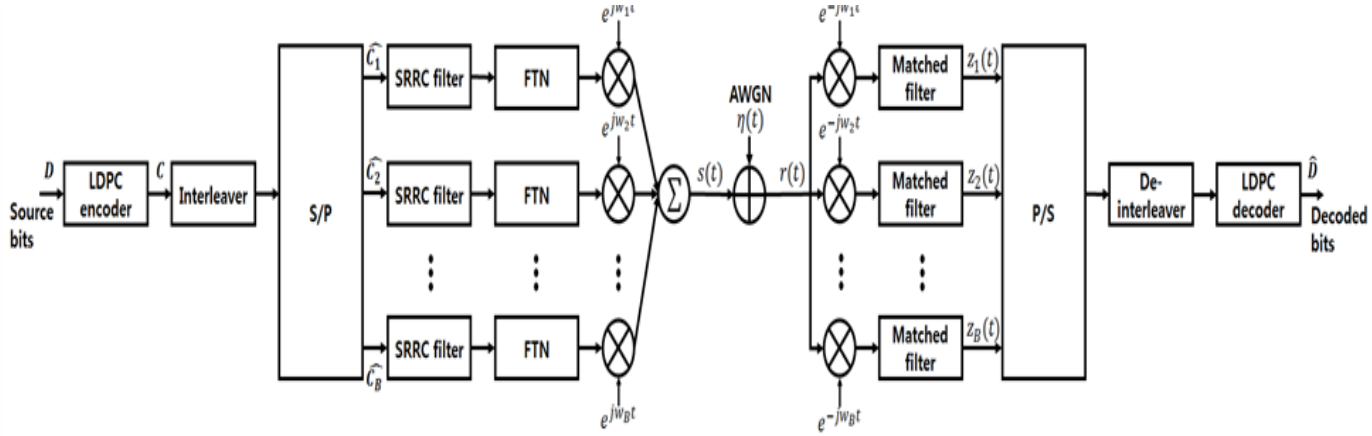


Figure 2. The structure of single band FTN model

In order to reduce ISI, the multiband FTN transmission method minimizes interference by increasing the number of samples per bit depending on the number of band, $\mathbf{N_B}$.

Received signal given by

$$\mathbf{y(t) = s(t) + \eta(t)} \\ \mathbf{= \sum_{b=1}^{N_B} \sum_{n=1}^{N/N_B} \hat{c}_b(n) p_b((n-t)\tau T_{N_B}) + \eta(t)} \quad (8)$$

where, $\boldsymbol{\eta(t)}$ denotes the AWGN(Additive White Gaussian Noise).

Let the output of each band passing through the matched filter be $\mathbf{z_b(t)}$, then $\mathbf{z_b(t)}$ is expressed as in Equation (9).

$$\mathbf{z_b(t) = ((y(t)e^{-jw_b t}) * p_b(t), \quad (b = 1, 2, \cdots, N_B)} \quad (9)$$

By separating the signal of each band after the demodulation process and the matched filter operation, received signals are input to the LDPC decoder after passing through P/S and the de-interleaver.

III. Simulation Results

We analyzed the performance of the multiband FTN method presented in this paper through computer simulations. In order to compare the performance of the FTN transmission method in the existing SISO channel, we comparatively analyzed the performance depending on the increase of $\mathbf{N_B}$ for the same interference ratio.

Simulation parameters are listed in Table 1.

Table **1.** Simulation parameters

Total number of data	10^6
Channel coding	LDPC (K=32400, N=64800)
Coding rate	1/2
Modulation	BPSK
Number of band ($\mathbf{N_B}$)	1 - 6
Roll off factor	0.35
Samples per bit ($\mathbf{N_S}$)	Single band : 8 Multi band : $\mathbf{8 \times N_B}$
Interference ratio ($\boldsymbol{\tau}'$)	25%
LDPC Innter iteration	60
Whole iteration	5

In Table 1. interference rate $\boldsymbol{\tau}'$ is given by [7]

$$\boldsymbol{\tau}'(\%) = 100 \times (1 - \tau) \quad (10)$$

We considered an environment where only the general AWGN channel is present for the simulation of the performance analysis of the multiband FTN. In order to compare the performance of the FTN transmission method in the single band, we compared the performance by setting the interference ratio to 25%. The performance depending on the increase of $\mathbf{N_B}$ is shown in Table 2.

Identify applicable sponsor/s here. If no sponsors, delete this text box. (sponsors)

Table 2. Performance of multiband FTN according to $\mathbf{N_B}$

	N_B					
SNR_R	1	2	3	4	5	6
−0.8	$10^{-1.33}$	$10^{-1.88}$	$10^{-2.24}$	$10^{-2.70}$	$10^{-3.00}$	$10^{-3.40}$
−0.7	$10^{-1.38}$	$10^{-1.99}$	$10^{-2.41}$	$10^{-3.84}$	0	0
−0.6	$10^{-1.42}$	$10^{-2.21}$	$10^{-3.92}$	0	–	–
−0.5	$10^{-1.48}$	0	0	–	–	–
−0.4	$10^{-1.56}$	–	–	–	–	–
−0.3	$10^{-2.11}$	–	–	–	–	–
−0.2	0	–	–	–	–	–

Table 2 shows the performance of the multiband FTN transmission method in AWGN channel environment. Here, $\mathbf{SNR_R}$ is expressed by the received signal strength $\boldsymbol{\sigma}_{\mathbf{s+n}}^{\mathbf{2}}$ and the noise strength $\boldsymbol{\sigma}_{\mathbf{n}}^{\mathbf{2}}$ as in Equation (11),

$$\mathbf{SNR_R} = \frac{\boldsymbol{\sigma}_{\mathbf{s+n}}^{\mathbf{2}} - \boldsymbol{\sigma}_{\mathbf{n}}^{\mathbf{2}}}{\boldsymbol{\sigma}_{\mathbf{n}}^{\mathbf{2}}} \qquad (11)$$

In Table 1, $\mathbf{N_B = 1}$ means the single band FTN transmission method and $\mathbf{N_B > 1}$ means the multiband FTN transmission method. Although the number of samples per bit $\mathbf{N_S}$ increases in proportional to the number of bands $\mathbf{N_B}$, the bandwidth of the transmitted signal is the same. For the single band, we increased proportionally whenever $\mathbf{N_B}$ increased based on $\mathbf{N_S = 8}$. We could confirm that the performance improved as the number of $\mathbf{N_B}$ increased compared to $\mathbf{N_B = 1}$. The performance was improved by approximately 0.5dB when $\mathbf{N_B = 5,6}$ compared with the single band FTN transmission method with $\mathbf{N_B = 1}$, and the performance kept improving as the value of $\mathbf{N_B}$ was continuously increased. However, it was confirmed that the performance improvement became reduced when $\mathbf{N_B = 5}$ or more.

IV. CONCLUSION

Next-generation wireless and/or satellite communications require high transmission efficiency and high reliability to provide various services with subscribers. Recently, many methods for increase of throughput is being researched, as the next satellite broadcast / communication and the 5G based mobile communication demand for throughput is increasing, whilst the bandwidth is limited.

In this paper, we proposed the multiband FTN method which is uniformly dividing and allocating the encoded bits into multiband, and transmits the encoded bits using the frequency allocated to each band by applying the FTN transmission method to each band. In addition, it can improve performance through the effect of noise averaging by increasing the number of samples per bit in each band to reduce interference between adjacent channels.

We compared the performance by setting the interference ratio to 25%. Through simulation results, we show that the performance of the proposed multiband FTN method is improved about 0.3 dB ~ 0.5 dB than that of single band FTN method.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2017R1D1A1A09082161).

REFERENCES

[1] J. E. Mazo, "Faster-than-Nyquist signaling," The Bell System Technical Journal, vol. 54, no. 8, pp. 1451-1462, Oct. 1975.

[2] A. D. Liveris and C. N. Georghiades, "Exploiting fasterthan-Nyquist signaling." IEEE Trans. Commun., vol. 51, no. 9, pp. 1502–1511, Sep. 2003.

[3] M. El Hefnawy, and H. Taoka, "Overview of Faster-Than-Nyquist for Future Mobile Communication Systems," In Proceeding 77th IEEE Vehicular Technology Conference, pp. 1-5, Dresden, Germany, Jun. 2013.

[4] H. Esmaiel and D. Jiang, "Review article: Multicarrier communication for underwater acoustic channel, " International Journal of Communications, Network and System Sciences, vol.6, pp.361–376, Aug. 2013.

[5] P. A. van Walree, E. Sangfelt, and G. Leus, "Multicarrier spread spectrum for Covert Acoustic Communications," In Proceeding of MTS/IEEE OCEANS, Quebec City, QC, Canada, pp. 1-8, Sep. 2008.

[6] F. Rusek and J. B. Anderson, "Multistream Faster than Nyquist Signaling," IEEE Transactions of Communications, vol. 57, no. 5, pp. 1329-1340, May 2009.

[7] R. G. Gallager, "Low-density parity - check codes," IRE Transactions on information theory, vol. 8, no. 1, pp. 21-28, Jan. 1962.

Image retrieval using normalization labeling of multiobjects and rearrangement

Guk-Jeong Kim
BonC Innovators Institute
Gwangju, Korea
sihyun06@nate.com

Tae-Yeun Kim
SW Convergence Education Institute
Chosun University
Gwangju, Korea
tykim@chosun.ac.kr

Young-Eun An
College of General Education Korea
Chosun University
Gwangju, Korea
yean@chosun.ac.kr

***Abstract*—In CBIR with multi-objects, it shows an interest in the image retrieval method based on object features. In this paper, we have proposed an image retrieval method based on normalized labeling features of multi-objects. First, it extracts, normalizes and rearranges the object-labels which present image shape information. Secondly, it measures image similarity using moment features invariant to transition, rotation and scaling of the objects. The simulation results using various coin database showed that our proposed method is superior to the edge histogram method.**

Keywords-image retrieval, invariant moment, labeling, feature, rearrangement, multi-object

I. INTRODUCTION

Recently, the growing Internet and mobile devices made people easily create and save digital images through web sites, blogs and social network services, etc. Along with this trend, applications of images for various uses are getting much interest and this led to the research on search methods based on visual contents of images and multi-media[1].

Most users for media search intend to find similar images based on particular objects included in the questioned images. The search methods to use the formal information of images include the use of varied value as feature information after extracting the outline of the objects[2] or that of pixels' distribution value of the whole image as feature information by utilizing internal information of the objects[3].

The paper intends to present image searching algorithm based on forms by extracting significant and interested objects automatically that could express core image information from images containing various objects. To extract interest objects, a process to choose interest objects should be proceeded.

Among multiobjects exist one prominent object in the normalization labeling and invariant moment feature extraction, which turns out to be very useful expressing relevant information on the images of the object. As such, it has a feature contrary to the surrounding backgrounds in the comparatively bigger sizes and the objects with obvious boundaries are predicted to be well reacted with interest object which efficiently expresses core information included in the images.

II. EXTRACTION OF ATTRIBUTES OF NORMALIZATION LABELING INVARIANT MOMENT

The paper has presented similarity search methods of feature extractions and sequence approximation using invariant moment with image search algorithms of feature vectors of locations and forms of particular objects. The image 3.1 shows the sequence of presented algorithm.

A. Preconditioning

As a preconditioning, noise is removed and hall noise is resolved by contraction and expansion. Using Otsu, the optimal global threshold treatment is preceded with watershed in the end to divisions. Fig. 1 shows preconditioning, Fig. 2 of query image and Fig. 3 of gray scale image.

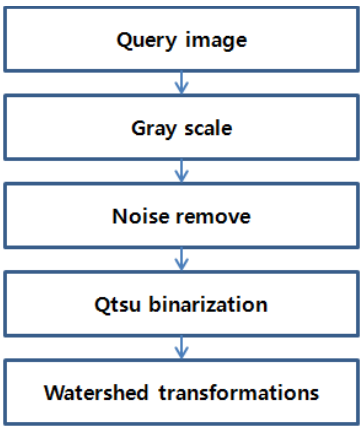


Figure 1. Preconditioning

The method of using Otsu is to create histogram of given images and divide and extract objects with similar values of brightness using this. That is, the Otsu method is to find threshold value to maximize divisions among divided classes.

The watershed algorithm is to divide images using mathematical morphology. Gradient value on the brightness of image pixels is transformed into altitude information to separate spheres. The algorithm is to find out combined sphere in each catchment basin by capturing watershed that tells from catchment basin.

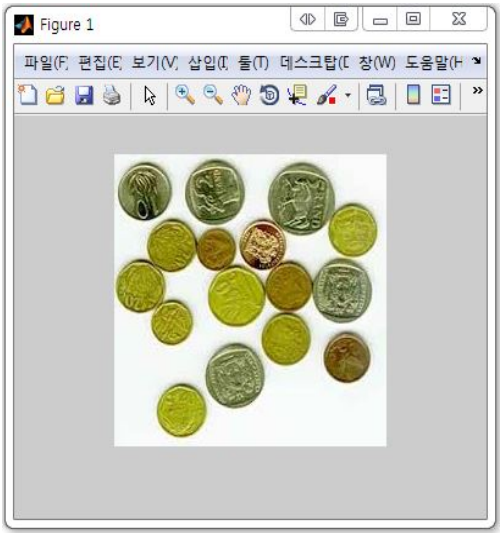


Figure 2. Query image

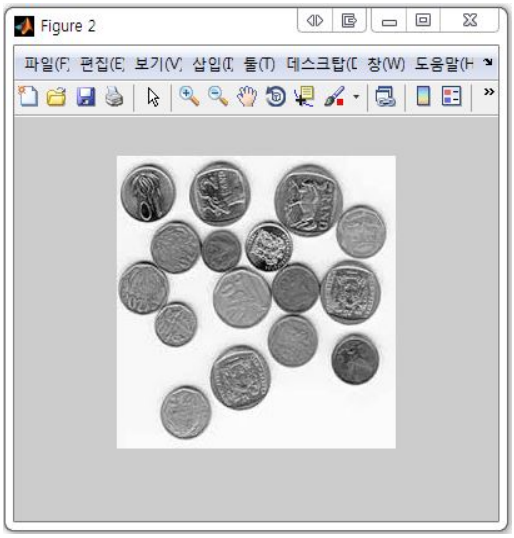
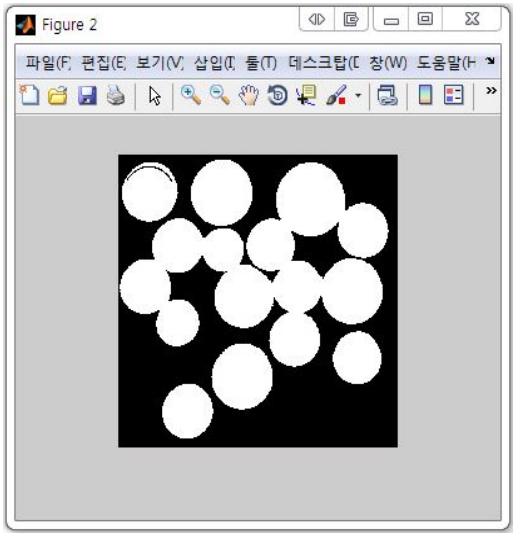
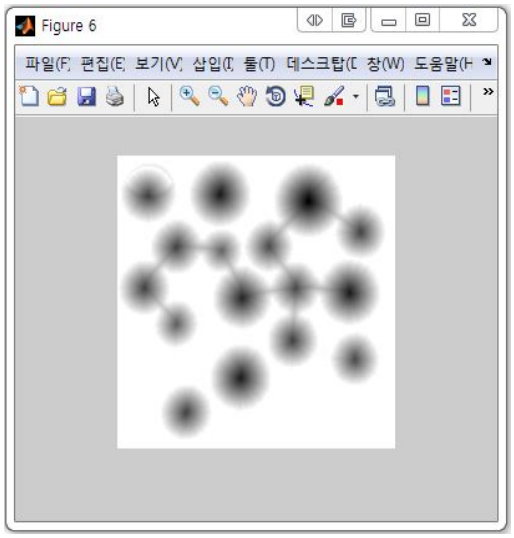


Figure 3. Gray scale image

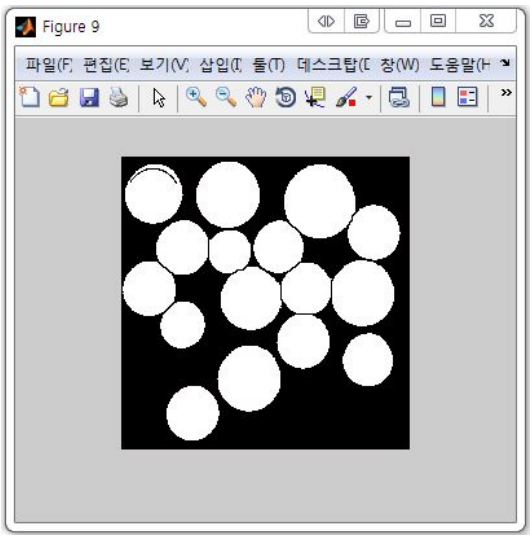
Frequently used tools along with watershed transformations for divisions are distance transformation. Binary distance transformations of the images are a relatively simple concept. That is the distance to the pixel not the nearest zero from each pixel. Fig. 4 (a) shows small binary image matrix with image (b) of distance variation in connection with it. The pixel of the all values of 1 indicates the nearest pixel itself, which means the distance transformation value as 0.



(a) Binary image



(b) Distance transform



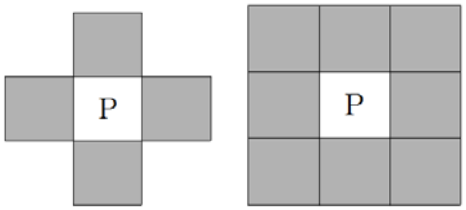
(c) Distance variation

Figure 4. Divisions using the watershed transformations

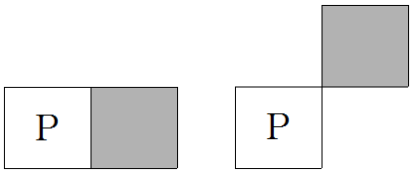
B. Labeling

Labeling is to divide each object using numbering method by using 4-connected or 8-connected in search of all binary pixels after Garber filtering.

In labeling, 4-connected and 8-connected can show Fig. 5 The pixel p on the (x, y) of the image portrays two horizontal neighbor and two vertical of (x + 1, y), (x − 1, y), (x, y+1), (x, y-1) with Fig. 5 (a) of N4(p) as a 4 neighboring groups of p. Four diagonal neighbors of pixel p have (x+1, y+1), (x+1, y-1), (x-1, y+1), (x-1, y-1), and p′s 8 neighboring group along with vertical and horizontal component and note them as N8(p) as shown in Fig. 5.



(a) 4-connected (b) 8-connected



(c) 4 neighboring and 8 neighboring (d) 4 neighboring

Figure 5. 4-connected and 8-connected

C. Extraction of Invariant Moment Feature Vecto

1) Extraction of form Information in the Labeling Sphere

If images are divided in the labeling spheres, each feature information of labeling is extracted. One image is made of numerous labeling and extractions of all feature information of all labeling are likely to have over head. Therefore, the paper extracts feature information only on the labeling within the rank 10th in sizes of labeling spheres. The feature information on the forms use invariant moment strong at movement, rotation, size changes.

D. Normalization Feature Vector

1) Normalization strong at Image Transformation

The paper needs the work of normalization as the features of images are dependent on noises due to distortion, inclination, size changes, etc. In the feature vector extracted ahead extract normalization vector strong at image transformation.

2) Saving Feature Information

Images compared should be indexed in the database after extraction of feature information by method described earlier. To save the feature information of each image, database table should be designed as Table 1.

TABLE I. DATABASE TABLE

Field 1	Field 2	Field 3	Field 4	Field 5
Feature vector	Normalization vector	Number of labels	Image address	Image name
MomentV	NormV	LabelNum	FileWithPath	FileName

E. Similarity Calculations

Among the various images saved in the image database, the realization of image search systems the most similar to the questioned image by users needs a method of efficient use and definitions of similarity measure among images Similarity scale among image can vary according to the attributes of images used as feature vectors. Similarity measurement represents similarity between two image data as figures. Data analysis, index and search are closely related and is a core technology that can influence the search performance based on the content.

This paper has calculated visual similarity of content based image search between questioned images and database images. Therefore, search results are search list distributed in the sequence of similarity not by single image lined up by the similarity of image along with questioned image. Much similarity measurement was devised to image search depending on the experience estimation of recent feature distributions, and each similarity/distance measurement would remarkably influence search performance of image search systems.

This paper uses the following simple distance scale functions to measure the form feature similarity among questioned images and database images.

$$D(Q,I) = \sum_{i=1}^{k} |F_i - F_i'| \quad (1)$$

In the Eq. (1), Q is questioned image, I with database images, F_i and F_i' as respective feature vectors of the two images. The similarity of the images is easily gained by calculating the absolute distance among two feature values and images similar to questioned images from the content based search systems are searched from search database according to the calculated similarity.

F. Performance Appraisal Scale

To analyze the efficiency of content based image search, generally Recall and Precision are used as performance appraisal scales. Recall is the ratio of search video among the questioned images in the image database and Precision represents ratio of images related to questions. That is, let us presume that A is the image groups within the image database and B is that of searched images. Then, Recall and Precision can be defined as conditional probability as the equation below.

$$Recall = P(\frac{B}{A}), \quad Precision = P(\frac{A}{B}) \quad (2)$$

In the real experiment, calculation is used as Eq. (3).

$$R_e = \frac{R_r}{T}, \quad P_r = \frac{R_r}{T_r} \quad (3)$$

T represents the total number of items related to questions in the database, $R_r$is the number of items related to questions in the searched items, $T_r$as that of searched items. That Recall is big in size means having high matching probability and that Precision is high means wrongly searched images are not much.

III. EXPERIMENT AND ANALYSIS

A. Experiment Environment

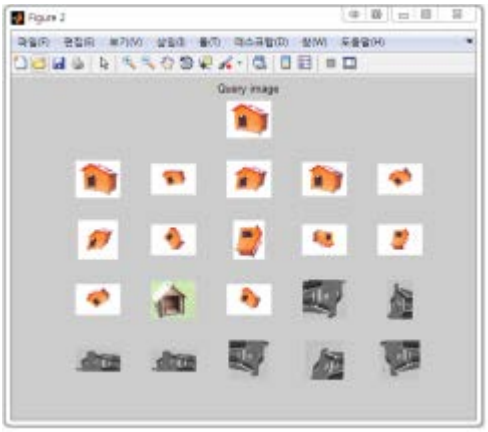
The presented algorithm in this paper was used in the experiment using MATLAB 2014a software on various images. This experiment was analyzed in the same conditions of questioned images. To explain the search method of images, 3,900 database images were laid to find out 20 according to the compared images. The similar images to the questioned images in the experiment are made as suitable number of images. Saved data is not calculated in the input of other questioned images, and only questioned data is calculated. Also, the use of feature values of data led to the comparing similarity measurement with conclusion images.

B. Simulations

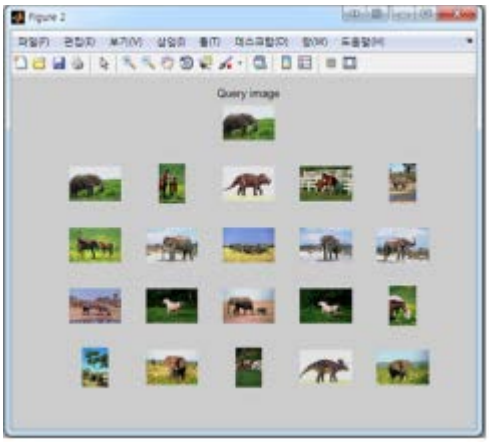
To measure the performance of presented algorithm, the existing algorithms in the questioned image Q1~Q4 in the image 4.1 and presented algorithm were used in the experiment to compare the performance measurement. 4 questioned images were used with each questioned image Q1~Q4 have similar images from questioned images in the 3000 database. After the comparisons of questioned images in the experiment, similar 20 images were defined in the sequence of higher similarity. The numbers of similar images of the questioned images were defined to find 20 with definitions of less than 20 images.

For presented content based image search system appraisal, the presented system using algorithm of the existing corner patch histogram and their Precision and Recall were evaluated.

In each similar images of questioned images, the next image presents search result and performance of the presented algorithm and the existing algorithm in the similar images. Fig. 6 shows the search result of algorithm of corner patch histogram and Fig. 7 shows the presented algorithm.

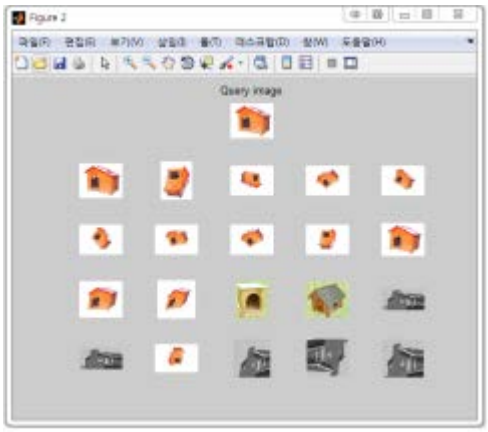


(a) Result of Q1

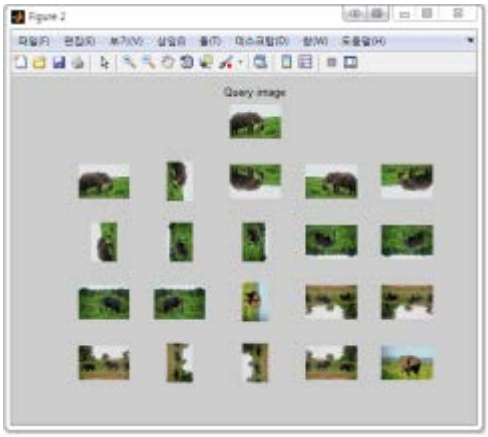


(b) Result of Q2

Figure 6. Result of algorithm of corner patch histogram



(a) Result of Q1



(b) Result of Q2

Figure 7. Result of the presented algorithm

Presented algorithm would show its superior search performance in the images of rotation, expansion and minimization and the existing corner patch histogram algorithm.

Algorithm using corner patch histogram had each performance value, Recall value as (a-0.85), (b-0.60) with Precision value as (a-0.60), (b-0.30).

Compared to the existing algorithm, the labeling moment based form search algorithm in the paper was improved as (a-0.92), (b-0.85) as Recall value, (a-0.65), (b-0.60) as Precision value.

TABLE II. PERFORMANCE ANALYSIS

Simulation \ Method		Corner patch histogram algorithm	Proposed algorithm
Image Q1	Recall	0.85	0.92
	Precision	0.60	0.65
Image Q2	Recall	0.60	0.85
	Precision	0.30	0.60

IV. CONCLUSION

The paper has presented simple and efficient search method for from extraction of multiobjects images based on the normalization labeling moment as form information attributes. The motion sequence rearrange from its maximum area by labeling the object in the image given in the questioned image and extract feature vector using invariant moment. From extracted feature vector, normalization vector is created strong at image transformation. Similarity is measured using measures of feature vectors.

Presented algorithm verifies the search efficiency along with the suppleness of algorithm through simulations. Presented search methods can better search expanded and contracted images and rotated ones than the corner patch histogram.

ACKNOWLEDGMENT

This study was supported by research fund from Chosun University, 2017

REFERENCES

[1] Y. Rui and T. S. Huang, "Image retrieval: current techniques, promising directions, and open issues," J. of Visual Communication and Image Representation, vol. 10, pp. 39-62, 1999.

[2] A.W.M Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349-1380, Dec. 2000.

[3] S. Michael, "Next-generation web searches for visual content," IEEE Computer, pp. 46-52, Nov. 2000.

[4] Xiang Sean Zhou and Thomas S. Huang, "CBIR: From low-level features to high-level semantics," Proc. SPIE Image and Video Communication and Processing, vol. 3974, Jan. 2000.

[5] J. Huanf, S. R. Kumar, and M. Mitra, "Combining supervised learning with color correlograms for content-based image retrieval," Proc. ACM Multimedia, pp. 325-334, Nov. 1997.

[6] J. R. Smith and S. F. Chang, Searching for images and videos on the world-wide web, tech. report TR-459-96-25, Columbia University, Aug. 1996.

[7] C. Carson, M. Thomas, S. Belongie, J. M. Hellerstein, and J. Malik, "Blobworld: a system for region-based image indexing and retrieval, "Proc. 3rd Int′l Conf. on Visual Information Systems, vol. 1614, pp. 509-516, Jun. 1999.

[8] A.H. Kam, T.T. Ng, N.G. Kingsbury, and W.J. Fitzgerald, "Content based image retrieval through object extraction and querying," IEEE Workshop on Content-based Access of Image and Video Libraries, pp. 91-95, Jun. 2000.

[9] Wei Wang, Yuqing Song, and Aidong Zhang, "Semantics retrieval by content and context of image regions," Proc. 15th Int′l Conf. on Vision Interface, pp. 17-22, May 2002.

Big Data Analysis Using the Aggregation Framework in a Shard Cluster Environment

Young-Woon Kim*, Hyeopgeon Lee

Department of Data Analysis, Seoul Gangseo Campus of Korea Polytechnic University, Korea, Republic of

luckkim@kopo.ac.kr, hglee67@kopo.ac.kr

***Abstract*— Performing analysis by extracting and aggregating data in real time by joining several tables in a relational database environment where data are accumulated continuously entails many difficulties. To resolve these issues, the advantages and disadvantages of three big data extraction and aggregation methods—Aggregation Framework, MapReduce, and Group—that operate in a shard cluster environment are compared and an optimal method is selected. The shard cluster is the distributed processing technology of MongoDB. The Aggregation Framework, designed for the improvement of performance and development productivity, was selected to perform big data analysis on similar estimates from an automobile repair estimate system where data have been accumulated continuously over several years.**

***Keywords*— Big Data Analsis, Big Data Analysis System, MongoDB, Shard Cluster, Aggregation Framework**

I. INTRODUCTION

Performing analysis by extracting and aggregating data in real time by joining several tables in a relational database environment where data are accumulated continuously entails many difficulties. To resolve these issues, MongoDB, a NoSQL database system, is used to extract and aggregate tables where a vast amount of data has been accumulated. MongoDB employs the shard cluster technology, which is a distributed processing technology that utilizes a cluster of several computers. MongoDB also utilizes the Aggregation Framework for the extraction and aggregation of large data. By utilizing these, MongoDB performs big data analysis on similar estimates from an automobile repair estimate system where data have been accumulated continuously over several years.

II. LITERATURE REVIEW

MongoDB is a document-oriented database based on reliability and scalability. MongoDB, which aims for low management cost and convenient usability with big data, was developed by 10gen with open source and hence it can be commercially supported[1].

The minimum storage unit in MongoDB is a document. Each document is collected at “Collection” and each collection is managed at the database to support the scope query, secondary index, alignment operation, and set operation of MapReduce.

MongoDB collects the document by collection and does not need a schema. A MongoDB query is created in JavaScript and document based query is conducted. This is a shell with real-time access and supports multiple program languages[2].

MongoDB is able to expand distribution by using auto-sharding. Sharding is a process in which data are divided and stored on different servers separately. By storing data to multiple servers, more data can be managed and processed[3][4][5][6].

MongoDB provides three methods for the extraction and aggregation of large data.

- Aggregation Framework: Designed for the improvement of performance and development productivity, the Aggregation Framework can operate in both non-shard and shard cluster environments.
- MapReduce: The MapReduce provides functions for fast processing of large data sets and operates in both non-shard and shard cluster environments.
- Group: Simple syntax and functions are provided for grouping by a designated key, but the Group does not operate in a shard cluster environment.

III. PROPOSED WORK

To select a method of extraction and aggregation of similar estimates from a large automobile repair estimate system, the advantages and disadvantages of Aggregation Framework, MapReduce, and Group were compared and analyzed as shown in Table 1.

The advantages and disadvantages of the three methods that enable extraction and aggregation of large data and operate in a shard cluster environment were compared. The shard cluster utilizes a distributed processing technology that uses a cluster of several computers. Among the three methods, Aggregation Framework was selected because it was designed for the improvement of performance and development productivity.

IV. CONCLUSION

The Aggregation Framework was designed for the improvement of performance and development productivity, and it utilizes MongoDB’s distributed processing technology

known as shard cluster to extract and aggregate large data. As shown in Figure 1, the Aggregation Framework is used to perform big data analysis on similar estimates from an automobile repair estimate system where data have been accumulated continuously over several years.

ACKNOWLEDGEMENT

This work was supported by the Technology development Program(S2600526) funded by the Ministry of SMEs and Startups(MSS, Korea)

REFERENCES

[1] Doo-sun Park, Yang-se Moon, Young-ho Park, Chan-hyun Yoon, Young-sik Jeong, Hyung-seok Chang , ' big data computing technology', hanbitacademy, 2014
[2] Ju-Jongmyeon, 'NoSQL & mongoDB,, DB, 2014
[3] MongoDB, http://docs.mongodb.com/manual/
[4] Arunkumar Thangavelu, N. Manoharan, 'Design and Analysis of an Effective Channel Distribution Approach for Agricultural Commodities using MongoDB,' Indian Journal of Science & Technology, 2016, 9(47), pp. 1-10.
[5] P. Parthiban, S. Selvakuma, 'Big Data Architecture for Capturing, Storing, Analyzing and Visualizing of Web Server Logs,' Indian Journal of Science & Technology, 2016, 9(4), pp. 1-9.
[6] Young-Woon Kim, Hyeopgeon Lee, 'Implementation of Big Data Analysis System to Prevent Illegal Sales in the Cable TV Industry,' Journal of Engineering and Applied Sciences, 2017, 12(23), pp.6542-6545.

Table 1. Comparison of advantages and disadvantages of big data extraction and aggregation methods in MongoDB

Category	Aggregation Framework	MapReduce	Group
Advantages	• Designed for the improvement of performance and development productivity • Operates in both non-shard and shard cluster environments	• Provides functions for fast processing of large data sets • Operates in both non-shard and shard cluster environments	• Provides simple syntax and functions for grouping by a designated key • Returns the results in-line arranged in an array of grouped items
Disadvantages	• Limits the size of the results document to 16 MB • Allows only the supported operators and expressions, and does not support user-defined functions	• MapReduce functions are difficult to debug. • Not intuitive to programmers who have experience with relational aggregation queries	• Processing takes a long time; hence, this method is used only when necessary. • Does not operate in a shard cluster environment

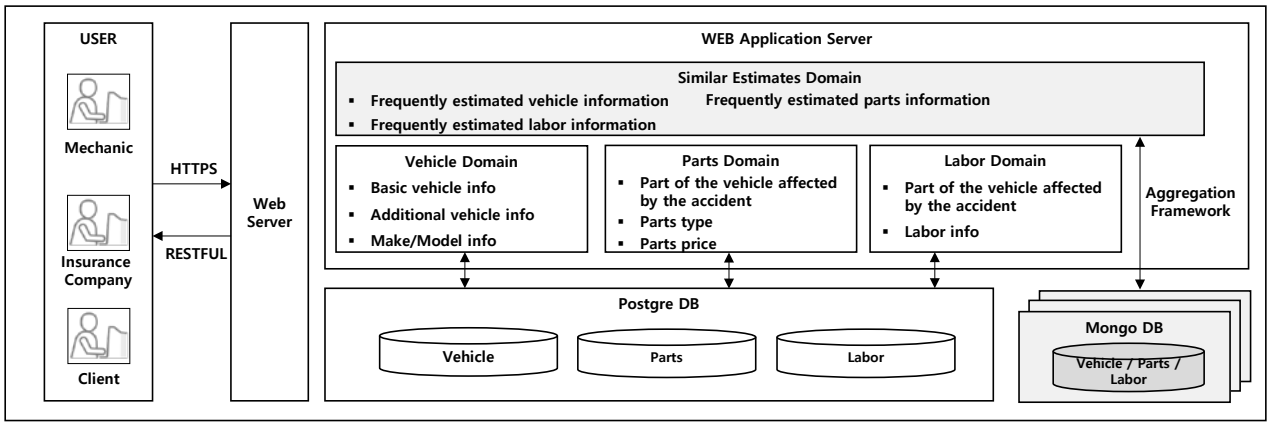


Figure 1. Big data analysis using the Aggregation Framework in a shard cluster environment

Implement of MapReduce Based Big Data File Compression Scheme Using SnappyCodec

Hyeopgeon Lee*, Young-Woon Kim, Ki-Young Kim

Department of Data Analysis, *Seoul Gangseo Campus of Korea Polytechnic, Korea, Republic of*

Dept. of Computer Software, Seoil University, Korea, Republic of

hglee67@kopo.ac.kr, luckkim@kopo.ac.kr, ganet89@seoil.ac.kr

***Abstract*—Large-scale data processing using thousands of nodes with built-in fault tolerance has become widespread due to the availability of open source frameworks, with Hadoop being a popular choice. Hadoop MapReduce is a software framework for easily writing applications which process vast amounts of data in-parallel on large clusters of commodity hardware in a reliable, fault-tolerant manner. However, the data analysis, big data developer and big data engineers are more increases that the performance of MapReduce frameworks for throughput. Thus, we propose the MapReduce based big data file compression scheme.**

Keywords-component; Big Data, Compression Scheme, Hadoop, MapReduce Framworks, SnappyCodec,

I. INTRODUCTION

Hadoop MapReduce is a software framework for easily writing applications which process vast amounts of data (multi-terabyte data-sets) in-parallel on large clusters (thousands of nodes) of commodity hardware in a reliable, fault-tolerant manner [1].

A MapReduce job usually splits the input data-set into independent chunks which are processed by the map tasks in a completely parallel manner. The framework sorts the outputs of the maps, which are then input to the reduce tasks. Typically, both the input and the output of the job are stored in a file-system. The framework takes care of scheduling tasks, monitoring them and re-executes the failed tasks. Typically, the compute nodes and the storage nodes are the same, that is, the MapReduce framework and the Hadoop Distributed File System are running on the same set of nodes. This configuration allows the framework to effectively schedule tasks on the nodes where data is already present, resulting in very high aggregate bandwidth across the cluster [1].

However, the data analysis, big data developer and big data engineers are more increases that the performance of MapReduce frameworks for throughput. And the big data engineer wants more increase the system resource for disk storage.

In this paper, we propose the MapReduce based big data file compression scheme. The proposed the MapReduce based big data file compression scheme uses SnappyCodec. And then the proposed scheme consists of the create file compression algorithm and read file compression algorithm.

The structure of the paper is organized as follows: Section 2 describes MapReduce frameworks; Section 3 describes the proposal big data file compression scheme; Section 4 describes the implement of proposal big data file compression scheme and Section 5 concludes this paper.

II. MAPREDUCE FRAMEWORKS

Hadoop MapReduce is a software framework for easily writing applications which process vast amounts of data (multi-terabyte data-sets) in-parallel on large clusters (thousands of nodes) of commodity hardware in a reliable, fault-tolerant manner [2].

A MapReduce job usually splits the input data-set into independent chunks which are processed by the map tasks in a completely parallel manner. The framework sorts the outputs of the maps, which are then input to the reduce tasks. Typically, both the input and the output of the job are stored in a file-system. The framework takes care of scheduling tasks, monitoring them and re-executes the failed tasks [2].

Hadoop MapReduce is a data processing framework that can be utilized to process massive amounts of data stored in HDFS. As we mentioned earlier, distributed processing of a massive amount of data in a reliable and efficient manner is not an easy task. Hadoop MapReduce aims to make it easy for users by providing a clean abstraction for programmers by providing automatic parallelization of the programs and by providing framework managed fault tolerance support [2].

MapReduce programming model consists of Map and Reduce functions. The Map function receives each record of the input data (lines of a file, rows of a database, and soon) as key-value pairs and outputs key-value pairs as the result. By design, each Map function invocation is independent of each other allowing the framework to use divide and conquer to execute the computation in parallel. This also allows duplicate executions or re-executions of the Map tasks in case of failures or load imbalances without affecting the results of the computation. Typically, Hadoop creates a single Map task instance for each HDFS data block of the input data. The number of Map function invocations inside a Map task instance is equal to the number of data records in the input data block of the particular Map task instance [2].

Hadoop MapReduce groups the output key-value records of all the Map tasks of a computation by the key and distributes

them to the Reduce tasks. This distribution and transmission of data to the Reduce tasks is called the Shuffle phase of the MapReduce computation. Input data to each Reduce task would also be sorted and grouped by the key [2].

The Reduce function gets invoked for each key and the group of values of that key (reduce<key, list_of_values>) in the sorted order of the keys. In a typical MapReduce program, users only have to implement the Map and Reduce functions and Hadoop takes care of scheduling and executing them in parallel. Hadoop will rerun any failed tasks and also provide measures to mitigate any unbalanced computations [2].

Figure 1. Shows that the MapReduce flow.

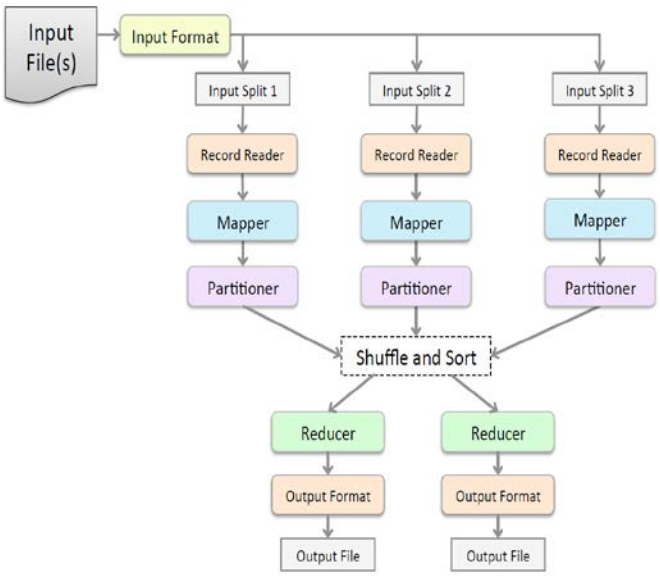


Figure 1. The MapReduce flow

III. Proposal Big Data File Compression Scheme

A. *Create File Compression Algorithm*

The Create File Compression Algorithm uses SnappyCodec that developed at Google for Hadoop Platform and execution time of SnappyCodec is very fast than present compression codec such as gzip and etc. SnappyCodec is a compression, decompression library. It does not aim for maximum compression, or compatibility with any other compression library instead it aims for very high speeds and reasonable compression. For instance, compared to fastest mode of zlib, Snappy is an order of magnitude faster for most inputs, but the resulting compressed files are anywhere from 20% to 100% bigger. On a single core of a Corei7 processor in 64bit mode, Snappy compress at about 250 MB/second more and decompress at about 500 MB/sec or more. The SnappyCodec consist of Binary Key and Value. Block compression is recommended for the SnappyCodec.

MapReduce uses the SnappyCodec for creating sequence file. Hadoop sequence file, containing a succession of key-value pairs, and periodic synch markers to facilitate MapReduce access. Figure 2. Shows that the driver code applied the create file compression algorithm.

```
public class CreateFileCompressionAlgorithm extends Configured implements Tool {

    @Override
    public int run(String[] args) throws Exception {

        Job job = new Job(getConf());
        job.setJarByClass(CreateFileCompressionAlgorithm.class);
        job.setJobName("Create File Compression Algorithm");

        FileInputFormat.setInputPaths(job, new Path(args[0]));
        FileOutputFormat.setOutputPath(job, new Path(args[1]));

        job.setOutputFormatClass(SequenceFileOutputFormat.class);

        FileOutputFormat.setCompressOutput(job, true);

        FileOutputFormat.setOutputCompressorClass(job, SnappyCodec.class);
        SequenceFileOutputFormat.setOutputCompressionType(job,
                CompressionType.BLOCK);

        job.setNumReduceTasks(0);

        boolean success = job.waitForCompletion(true);
        return success ? 0 : 1;
    }

    public static void main(String[] args) throws Exception {
        int exitCode = ToolRunner.run(new Configuration(),
                new CreateFileCompressionAlgorithm(), args);
        System.exit(exitCode);
    }
}
```

Figure 2. the driver code applied the create file compression algorithm

B. *Read File Compression Algorithm*

The Read File Compression Algorithm reads a compression sequence file. For file reading, The Read File Compression Algorithm uses SequenceFileInputFormat object. And then The Read File Compression Algorithm not uses the reducer. So The Read File Compression Algorithm sets property that job.setNumReduceTasks(0).

Figure 3. Shows that the driver code applied the read file compression algorithm.

```
public class ReadFileCompressionAlgorithm extends Configured implements Tool {

    @Override
    public int run(String[] args) throws Exception {

        Job job = new Job(getConf());
        job.setJarByClass(ReadFileCompressionAlgorithm.class);
        job.setJobName("Read Compressed Sequence File");

        FileInputFormat.setInputPaths(job, new Path(args[0]));
        FileOutputFormat.setOutputPath(job, new Path(args[1]));

        job.setInputFormatClass(SequenceFileInputFormat.class);

        job.setNumReduceTasks(0);

        boolean success = job.waitForCompletion(true);
        return success ? 0 : 1;
    }

    public static void main(String[] args) throws Exception {
        int exitCode = ToolRunner.run(new Configuration(),
                new ReadFileCompressionAlgorithm(), args);
        System.exit(exitCode);
    }
}
```

Figure 3. the driver code applied the read file compression algorithm

IV. Implemnt of File Conversion Scheme

We proposes the MapReduce based big data file compression scheme. The proposed the MapReduce based big

data file compression scheme uses SnappyCodec. And then the proposed scheme is consist of the create file compression algorithm and read file compression algorithm. For executing of the MapReduce based big data file compression scheme, we make Hadoop job (MapReduce Job). Hadoop Job consist of compiled CreateFileCompression Algorithm and compiled ReadFileCompression Algorithm.

Figure 4. An executed the create file compression algorithm.

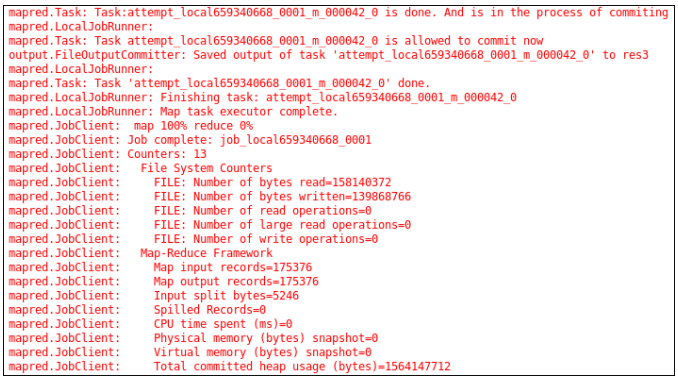


Figure 4. An executed the create file compression algorithm

Figure 5 shows that the result of create file compression algorithm.

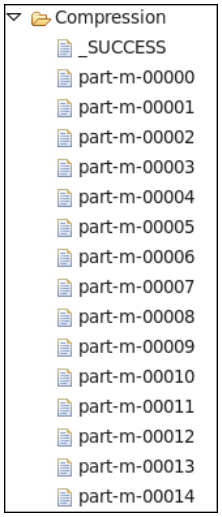


Figure 5. the result of create file compression algorithm.

V. CONCLUSIONS

Hadoop MapReduce is a software framework for easily writing applications which process vast amounts of data (multi-terabyte data-sets) in-parallel on large clusters (thousands of nodes) of commodity hardware in a reliable, fault-tolerant manner.

However, the data analysis, big data developer and big data engineers are more increases that the performance of MapReduce frameworks for throughput. And the big data engineer wants more increase the system resource for disk storage.

In this paper, we propose the MapReduce based big data file compression scheme. The proposed the MapReduce based big data file compression scheme uses SnappyCodec. And then the proposed scheme consists of the create file compression algorithm and read file compression algorithm. The create file compression algorithm is uses SnappyCodec and the read file compression algorithm is uses SequenceFileInputFormat.

ACKNOWLEDGEMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2017R1D1A1B03034729)

REFERENCES

[1] http://hadoop.apache.org

[2] Thilina Gunarathne, 'Hadoop MapReduce v2 Cookbook Second Edition', Packt Publishing, 2015

[3] Hema Jadhav, 'SNAPPY COMPRESSION DECOMPRESSION USING MAPREDUCE TECHNIQUE,' International Journal of Emerging Technology and Innovative Engineering', Volume I, Issue 5, 2015

[4] Nick Pentreath, 'BIRMINGHAM -Machine Learning with Spark', Packt Publishing, 2015

A Study on the BCI-based Intelligent Upper Limb Rehabilitation Robot System

Tae-Yeun Kim
SW Convergence Education Institute
Chosun University
Gwangju, Korea
tykim@chosun.ac.kr

Young-Eun An
College of General Education Korea
Chosun University
Gwangju, Korea
yean@chosun.ac.kr

Sang-Hyun Bae
dept. Computer Science and Statistics
Chosun University
Gwangju, Korea
shbae@chosun.ac.kr

***Abstract*—This study measured and analyzed EEG of users for intelligent upper limb rehabilitation of the system this study proposed, and intended to optimize the system using a genetic algorithm. Then, this study realized a BCI-based Intelligence Upper Limb Rehabilitation Robot and proposed an intelligent fuzzy neural network model. This study evaluated the performances of the system this study proposed through results of EEG electrode positions by making four actions of arms bending and unfolding, and hands closing and opening as actions of intention recognition of specific users. It also found optimal EEG electrode positions and analyzed differences in membership functions, number of clusters, number of learning generations, learning algorithms and the wavelet settings.**

Keywords-component; Rehabilitation robot system; BCI; Fuzzy; Genetic algorithm

I. INTRODUCTION

As previous studies on rehabilitation of the upper limbs focused on system design and peripheral sensors, studies on intelligent rehabilitation robots which should be applied according to users' intentions are still in the beginning stages of development [1][2][3]. In case of existing rehabilitation robots, as it is hard to judge patients' participation and intervention in rehabilitation programs and the programs were manually managed, the effects of rehabilitation may decrease.

Human communication technology that can identify users' intention is developing into the communication technology between humans and machines beyond the technology only for communication between humans. In particular, the BCI technology, which is developing by users' intentions, is defined as all the technology that collects EEG and detects, interprets and classifies it into meaningful information, and as EEG contains diverse information including thoughts and emotions, the system collects and acquires the information to provide unprecedented quality service [4]. The technology to detect and classify meaningful information collected from EEG is important to deliver users' intentions precisely to computers [5][6]. The robots can be effectively used for rehabilitation when they judge patients' intervention through identification of users' intention with the use of EEG and induce them to actively participate in programs [7].

Therefore, this study proposed a new intelligence upper-limb rehabilitation robot system with which rehabilitation programs are available according to users' intentions and which can prevent degeneration of minimal physical functions in musculoskeletal function. That is, this study realized the BCI-based intelligence upper-limb rehabilitation robot system, which can support a regular and qualitative rehabilitation program as it measures and analyses users' EEG and, tested the exactness of the system with the use of an intelligence fuzzy neural network system model.

II. SYSTEM CONFIGURATION AND DESIGN

A. System Configuration Diagram

The BCI-based intelligence upper limbs rehabilitation robot system was composed of a EEG measuring system, a users' intention recognition EEG classification and analysis system, and a upper limbs rehabilitation robot system as seen in Fig. 1.

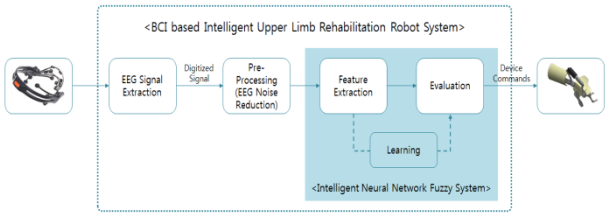


Figure 1. BCI (Brain-Computer Interface) system configuration diagram

This study used EPOC Neuroheadset produced by Emotiv which is accompanied by non-evasive EEG electrodes [8][9]. EEG were measured according to EEG 10-20 electrode standard indices which indicates each signal at each spot where electrode is attached through analyses of EEG features and data measured was made into a database [10]. To realize users' intentions recognition EEG classification and analysis system, this study developed a fuzzy neural network model which can identify users' intentions using EEG. The fuzzy neural network model extracted features of each order users recognized and classified them into intrinsic orders, and to remove noise in EEG, it processed data using a Wavelet system. For optimal rules of the fuzzy system, this study used a cluster classification using cluster analysis and cluster assessment, and for exactness of the system, it tuned membership functions, parameters and weights of the fuzzy system using a genetic algorithm. Finally, this study designed the upper limbs rehabilitation robot system for rehabilitation programs where users' intention is reflected.

B. Intelligent Neural Network Fuzzy System

To control the EEG classification and analysis system that recognizes users' intentions, this study designed a neural network fuzzy system, developed modeling rules and defined membership functions of each rule.

Fig. 2 presents the composition of the intelligence neural network fuzzy system for data detection and learning. In data processing, standards to measure EEG were set up and EEG were measured for a certain period. The EEG data was processed through the wavelet transform, and the data was saved in a database along with data before processing.

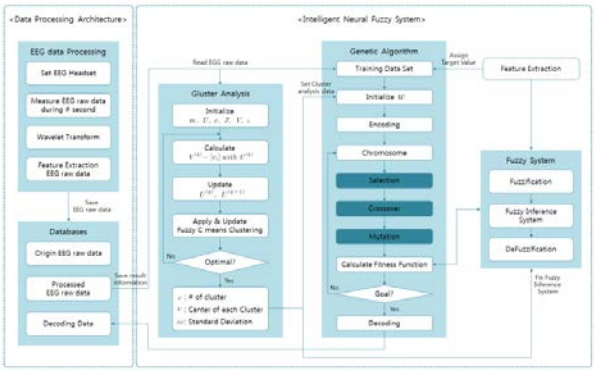


Figure 2. BCI (Brain-Computer Interface) system configuration diagram

The fuzzy neural network classified clusters using FCM (Fuzzy c-Means) and set up the optimal number of clusters through optimal cluster analysis and evaluation. When the number of clusters was decided, it obtained the center of gaussian membership functions under each cluster and standard deviations. Then, arbitrary weights were set up for learning. This study used a genetic algorithm for learning and learning was made based on EEG data which was measured and processed in advance and the values obtained through cluster analysis. Initial weights were arbitrarily set up and chromosomes were expressed with standard deviations and weights in each cluster. As fitness functions were set up to be asymmetrical gaussian values, the right and left values of standard deviations were separated to set up chromosomes and the right and left values of initial standard deviations were the same. If chromosome set-up was completed, learning was made through choice, crossing and variation and fitness functions were evaluated to check if conditions were satisfactory. The fitness functions were set up through the neural network fuzzy system and can be checked through non-fuzzy results. The error ratios were expressed as results of the fitness functions and if conditions were satisfied, learning was completed and if not, learning continued.

The wavelet system, which was used to process the EEG signals collected, changed frequencies into random waveform. When a signal was large, the wavelet extension constant decreased in size, could make more exact local description and separated features of the signal. As there was not only one wavelet, it has to be designed according to each application [11].

The EEG were divided into five stages of slow and fast rhythms according to frequencies and ranged from 0 to 50Hz. This study analyzed the EEG spectrum of the 5-staged frequency band and performed the 4-stage wavelet decomposition presented in Table 1.

TABLE I. WAVELET DE COMPOSITION ACCORDING TO EEG FREQUENCY

Frequency	Band	Analysis Level	Wavelet Constant
Alpha (α) wave	8 ~ 12 Hz	3	H_3
Beta (β) wave	12 ~ 30 Hz	2	H_2
Thetk (θ) wave	4 ~ 7 Hz	4	H_1
Gamma (γ) wave	30 ~ 50 Hz	1	H_1
Delta (δ) wave	0.1 ~ 3 Hz	4	L_1

As a way to extract EEG features, source data of EEG was analyzed by the wavelet transform and the 4-stage wavelet constants were used as multi-dimensional features, and the 4-level wavelet transform whose mother function, , was used. The wavelet coefficients extracted mean EEG data distribution according to time and were divided into the mean (μ), standard deviation (σ) and entropy (ε) of the range of EEG data for a certain period as means to express EEG as a feature. The features are presented as in Eq. (1).

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{N}(x_1-\mu)^2}\ ,\ \mu = \frac{1}{N}\sum_{i=1}^{n}|x_i|\ ,\ \varepsilon = -\sum_{i=1}^{N} p(x_i log_2(x_i)) \quad (1)$$

This study used the means (μ) as a means to express EEG and expressed the mean at each EEG electrode. The membership function was a means to linguistically express ambiguity of a given fuzzy set and deliver the expression. The Gaussian function has right and left symmetry, but this study used the asymmetrical Gaussian membership function as seen in Eq. (2) which shows different right and left standard deviations through the Gaussian transform.

$$A_{ji}(x_i) = \begin{cases} e^{\left(\frac{-(x_i-\upsilon_{ji})^2}{2(s_l^j)^2}\right)} & x_i \le \upsilon_{ji} \\ e^{\left(\frac{-(x_i-\upsilon_{ji})^2}{2(s_l^j)^2}\right)} & x_i > \upsilon_{ji} \end{cases} \quad (2)$$

A(x) indicates the Gaussian membership function, υ indicates the means under each cluster in a rule and s indicates a standard deviation of the given membership function. The right and left standard deviations of the membership functions were managed separately as unsymmetrical membership functions were used. s_l is a standard deviation which takes the left side of the Gaussian function and s_r is a standard deviation which takes the right side of the Gaussian function. For defuzzification after the membership functions were calculated, the fuzzy inference was first defined with the use of knowledge base. The membership value of the j th rule of the given input vector was calculated as in Eq. (3) [12].

$$U_j = A_{j1}(x_1)A_{j2}(x_2) \dots A_{jn}(x_n) \quad (3)$$

For the final value, defuzzification was made with the use of Eq. (4) based on the Center of Gravity. The final value was calculated according to the input value.

$$y^k = \frac{\sum_{p=1}^{n} U_p^k w_p^k}{\sum_{p=1}^{n} U_p^k}, \quad U_p^k = \prod_{i=1}^{m} A_{ki} x_i \quad (4)$$

To determine the structure of the fuzzy neural network system, this study conducted a FCM (Fuzzy c-Means) clustering and a cluster efficacy test. The FCM clustering system is a classification algorithm which clusters data under a cluster according to the degree of belonging and recounts the degree of belonging. X set with n vectors was decomposed into c fuzzy groups and the center of cluster was found from a group where the cost function of dis-similarity measurement became the minimum. The sum of belonging in data set always has to be 1 [13]. To obtain the differences between data, this study calculated weights and output information needed using Eq. (5).

$$J\left(u_i^k, v_i\right) = \sum_{i=1}^{c} \sum_{k=1}^{n} (u_i^k)^m (d_i^k)^2 \quad (5)$$

Objective function is presented as in Eq. (6) when N learning data set was defined as $\{(X_k, y_k^t),\ k = 1\ to\ N\}$.

$$(d_{ik})^2 = \|x_k - v_i\|^2 + (y_k^t - w_i)^2 \quad (6)$$

After the optimal cluster number was identified to express the next chromosome, the mean(u), standard deviation(υ) and weight(w) were calculated as in Eq. (7) [12].

$$u_i^k = \frac{1}{\sum_{u=1}^{c} \left(\frac{d_i^k}{d_j^k}\right)^{\left(\frac{2}{m-1}\right)}}, \ v_i = \frac{\sum_{k=1}^{n} (u_i^k)^m X^k}{\sum_{k=1}^{n} (u_i^k)^m}, \ w_i = \frac{\sum_{k=1}^{n} (u_i^k)^m T^k}{\sum_{k=1}^{n} (u_i^k)^m} \quad (7)$$

v_i was a vector with the center value of the j th rule and w_i output the j th rule. u_i^k had values of [0, 1] according to the degree the k th input data X^k belonged to i th cluster. m used 2 value as a weight of index that indicates the effect on ambiguity degree of membership function. T was a target value of y.

For expression of fitness function and chromosome, genotype definition had to be identified. Genes were arranged based on right and left values of standard deviations and weights of membership functions in fuzzy rules and as the mean value is representative of each cluster, it was fixed during learning. Standard deviation indicated a standard deviation of membership function to be fine-tuned during learning and the weights indicated values to be fine-tuned during learning. To optimize parameters and weights of Gaussian membership functions, this study used a genetic algorithm. The learning results were checked according to the means, standard deviations and weights through the neural network fuzzy system, and y value was extracted through defuzzification. This study compared the extracted y value with the target(t) value as original results to measure error ratio(E) as in Eq. (8).

$$\mathrm{E} = \frac{1}{n} \sum_{k=1}^{n} \sqrt{(t_k - y_k)^2} \quad (8)$$

The measured error ratio was a standard to judge fitness of the gene-learning algorithm and initial genes of learning algorithm that composed gene groups had to be set up based on chromosomes. The initial center value and dispersion value of each rule were defined as in Eq. (9) and Eq. (10).

$$v_{ij} = \frac{\left(\sum_{k=1}^{N} u_{ik} x_{ki}\right)}{\mathrm{N}} \quad \mathrm{i} = 1, 2, \dots, \mathrm{c} \ \ \mathrm{j} = 1, 2, \dots, \mathrm{n} \quad (9)$$

$$(\sigma_{jl}^i)^2 = (\sigma_{jr}^i)^2 = \sum_{k=1}^{N} \frac{u_{ik} \left(x_{kj} - v_{ij}\right)^2}{N} \quad i = 1, 2, \dots, c \quad j = 1, 2, \dots, n \quad (10)$$

The initial weight value and standard deviation value were copied on the final genotype and the rest of the genotype had arbitrary values to create the initial gene group. Fitness of the initial gene group was calculated with the use of the fuzzy inference and Eq. (8) and the next generation was created from operators of the gene algorithm and entities selected. The next generation created performed learning process of selection, crossbreeding and variation. If the error ratio value satisfied the conditions of the fitness function, learning was completed and if not, learning kept going.

C. Intelligent Upper Limb Rrehabilitation Rrobot System

The design of the intelligence upper limbs rehabilitation robot proposed in this study is presented in Fig. 3.

The system proposed was designed to allow wrists and elbows to turn. The wrist turning was designed to gain grasping force of fingers and to replace the elbow turning. As the elbow part only turned, a decelerator was used to make turning force bigger and reduce turning speed with the use of a general DC motors, and a linear motor was used for turning of the wrist part. Also, diverse tuning functions and a physical turning preventive system were installed to allow only a certain part of the elbow turning part to be turned.

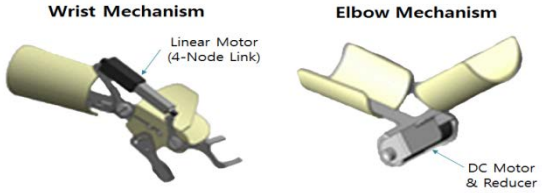


Figure 3. Intelligent upper limb rehabilitation system configuration

III. System Implementation and Performance Evaluation

A. *BCI-based intelligent upper limb rehabilitation robot system*

The intelligence upper limbs rehabilitation robot proposed in this study was composed of wrist turning part, elbow turning part and control part for motor drive and control.

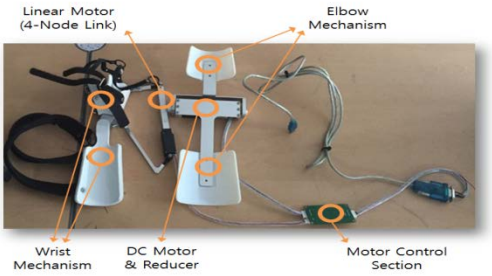


Figure 4. Intelligent upper limb rehabilitation system

The wrist turning part included a 4-joint link turning part and a linear motor for turning, and the elbow turning part included a DC motor for turning force and a decelerator to reduce speed and gain torque. It could control the wrist and elbow joint activation and angles could be controlled through CAN or RS232, and fast calculation and control was available through 32bit ARM Cortex M3. It had a BLDC motor (elbow motor) PID position controller and a Linear motor (wrist motor) for wrist activation and control. In particular, as accident may occur when users use it without recognizing environmental change, a protector from over voltage, low voltage and over current and an angle limit were added. Fig. 4 shows the realized BCI-based intelligence upper limbs rehabilitation robot system.

B. *Performance evaluation*

This study used EPOC Neuroheadset manufactured by Emotiv with EEG non-invasive electrode attachments to test the performance of the realized system, and measured EEG according to the EEG 10-20 electrode standard index, which showed each signal at every spot an electrode was attached through analyses of EEG features. For the data signals to be measured in the test, as CMS and DRL may belong to the parietal or the temporal lobes according to size of human head, this study measured 16 parts except EEG electrode data. For collection of EEG data, four movements of right arm bending, arm unfolding, hand closing and opening were made and EEG data were collected for 4 seconds with identical thinking and body movement.

Approximately 4000 ~ 5000 bits of data were collected for one session and when the intensity of EEG was the most precise, it was designed to use data only with signals. After the EEG extracted for one session was processed with the wavelet transform, each signal was expressed as the mean to make one data where values were saved. To identify users' intention, test data were composed of arms bending and unfolding process with hands open and hand closing and opening process with arms unfolded. For learning, data was classified into learning data and test data. Classification standard was 7:3 and when the learning data and test data were classified, classification standards were set up.

TABLE II. Wavelet de composition according to EEG frequency

Data Information	t_1	t_2	t_3
Arm Bending	0	0	1
Arm Unfolding	0	1	0
Hand Closing	1	0	1
Hand Opening		1	0

Table 2 presents expression of the results according to users' intention. As arms and hands are closely connected with each other, they were expressed as three outputs. t indicates output value and the number of output node in the fuzzy neural network system was three. When t_1 was 0, hands moved and when it was 1, arms moved. When t_2 was 0, it meant 'closed', and when it was 1, it meant 'open'. t_3 is vice versa.

This study conducted fuzzy optimal clustering process with 10 clusters as maximum permitted limits to evaluate what the optimal clusters could be. The whole EEG data measured, according to positions of the electrodes, was used and when measurement results were the highest, the number of clusters was deemed to be optimal.

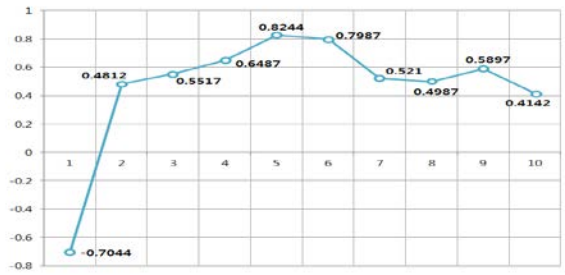


Figure 5. Optimal cluster evaluation

As in Fig. 5, when the number of clusters was 5, efficiency was the best and the number of hidden layers in the neural network and initial parameters were decided based on these results. This study intended to gain optimal solutions using EEG data measured according to positions of electrodes. First of all, right and left EEG signals were classified as in Fig. 6 and learning was made with each signal.

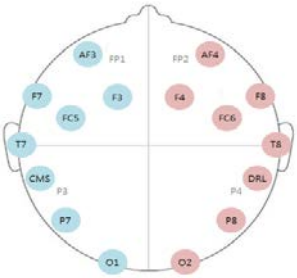


Figure 6. EEG signal left and right classification

To obtain the mean error ratio of learning according to right and left symmetrical EEG data, learning was completed at 100 generations and learning resumed with the use of the same EEG data. When every learning session was made, learning data and test data were randomly set up into five decomposed data sets and learning was made with the use of the decomposed data sets. At every data set, 50 sessions of learning was made and the means and standard deviation of the error ratio were obtained.

Fig. 7 shows the mean learning error ratio according to right and left symmetrical EEG data. As a result, when conditions were no. 13 F7, FC5, F8, FC6 and no. 18 F3, F7, FC5, F4, F8, FC6, error ratio was below 10%, and as difference in standard deviations was not big, it was discovered that they were equal.

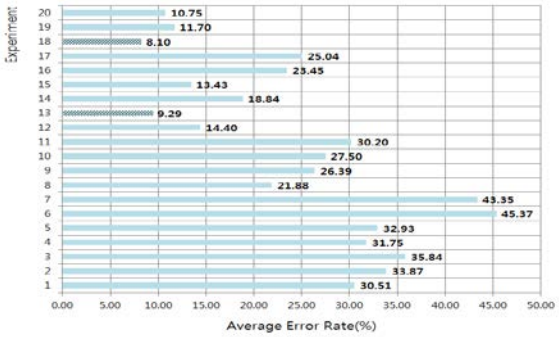


Figure 7. Learning average error rate according to left and right symmetry data

To compare the mean learning error ratio according to right and left decomposed data, this study conducted learning using only EEG under left brain and right brain and learning process was the same as that presented in Fig. 7. Fig. 8 shows comparison of the mean learning error ratios according to right and left decomposed data. As a result of learning with right and left decomposed data, left brain showed a little bit better results and incase of Condition 2 AF3, F7, F3, FC5, the mean error ratio was the lowest.

Figure 8. Learning average error rate according to left and right segmentation data

With the condition that the error ratio of the two conditions was below 10%, this study compared learning results when the number of generations increased to 1,000 as a completion condition. Table 3, As in Table 4, when the number of generations increased to 1,000 learning data mean error ratio was below 8%. In case that F7, F8, FC5, FC6 were input data, the mean error ratio was 5.91%, the lowest. For the test data, the error ratio was 6.31%, the lowest.

TABLE III. AVERAGE ERROR RATE AFTER 1000 LEARNING DATA

	EEG Electrode	**1**	**2**	**3**	**4**	**5**	**Average Error Rate(%)**
1	F7, F8, FC5, FC6	5.98	5.74	5.86	5.92	6.06	5.91
2	F3, F4, F7, F8, FC5, FC6	6.21	6.33	6.21	6.17	6.13	6.21
3	AF3, F7, F3, FC5	7.32	7.17	7.24	7.42	7.25	7.28

TABLE IV. AVERAGE ERROR RATE OF TEST DATA AFTER 1000 GENERATION LEARNING

	EEG Electrode	**1**	**2**	**3**	**4**	**5**	**Average Error Rate(%)**
1	F7, F8, FC5, FC6	6.24	6.73	6.02	6.24	6.34	6.31
2	F3, F4, F7, F8, FC5, FC6	7.02	6.87	7.24	7.19	6.95	7.05
3	AF3, F7, F3, FC5	8.11	7.93	8.01	8.12	8.32	8.10

To compare learning results according to the wavelet transform, this study checked differences according to data processing. With the use of optimal EEG electrode positions obtained through learning results according to electrode positions, learning was made. 1,000 generations were set up as a completion condition and 10 sessions of leaning was made to check the mean learning error ratio. This study decomposed the wavelet system into 4 stages and for comparison, the wavelet system was composed of 3, 4 and 5 stages. A test was also made without wavelet. As a result, there was no significant difference in Stages 3, 4 and 5 of the wavelet. However, there was a difference of approximately over 10% between the test with wavelet and the test without wavelet. Table 5 shows the mean learning error ratios according to presence or absence of the wavelet.

TABLE V. AVERAGE ERROR RATE OF TEST DATA AFTER 1000 GENERATION LEARNING

	Condition	**EEG Electrode**	**Basis**	**Average Error Rate(%)**
1	Wavelet setup	F7, F8, FC5, FC6	D3	6.23
2			D4	5.91
3			D5	6.89
4		F3, F4, F7, F8, FC5, FC6	D3	7.29
5			D4	6.21
6			D5	7.43

	Condition	EEG Electrode	Basis	Average Error Rate(%)
7			D3	7.72
8		AF3, F7, F3, FC5	D4	7.28
9			D5	8.16
10		F7, F8, F3, FC5		21.93
11	Wavelet not set	F3, F4, F7, F8, FC5, FC6		22.78
12		AF3, F7, F3, FC5		24.10

This study checked the mean error ratios with fixed number of clusters to analyze differences in optimal clustering evaluation. With 1,000 generations as a completion condition, ten sessions of learning was made and the mean learning error ratio was identified.

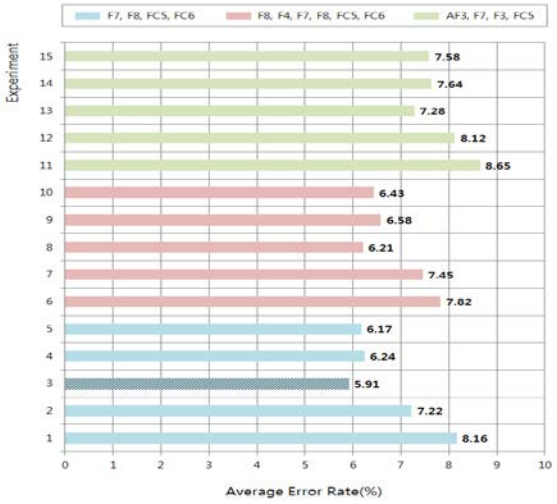


Figure 9. Learning average error rate according to cluster number

As in Fig. 9, in respect to the mean learning error ratios according to the number of clusters, when the number of clusters was 5 or 6, there was no significant difference between the mean error ratios. The lowest error ratio was found when the number of clusters was five and input data was F7, F8, FC5, FC6. As the number of clusters increased, differences in the mean error ratios according to the positions of EEG electrodes were not significant.

This study needed research to see whether the asymmetrical Gaussian function was the most suitable as a membership function. Therefore, this study compared error ratios with the symmetrical Gaussian function and the asymmetrical Gaussian function as membership functions. As there was a difference only in settings of membership functions and the symmetrical function had the right and left identical standard deviations, chromosomes were decided to have the identical settings. With 1,000 generations as a completion condition, ten sessions of learning were made to see the mean learning error ratios. As a result, the difference between the symmetrical Gaussian membership function and the asymmetrical Gaussian membership function was below 2%, which was insignificant. Table 6 shows the mean learning error ratio according to settings of the Gaussian membership function.

TABLE VI. LEARNING AVERAGE ERROR RATE ACCORDING TO GAUSSIAN BELONGING FUNCTION SETTING

	Condition	EEG Electrode	Average Error Rate(%)
1	Asymmetric Gaussian	F7, F8, FC5, FC6	5,.91
2		F3, F4, F7, F8, FC5, FC6	6.21
3		AF3, F7, F3, FC5	7.28
4	Symmetrical Gaussian	F7, F8, FC5, FC6	7.65
5		F3, F4, F7, F8, FC5, FC6	8.33
6		AF3, F7, F3, FC5	8.67

This study compared a genetic algorithm and a back-propagation algorithm to analyze the learning results according to the learning algorithms. The number of the hidden layer nodes in the back-propagation algorithm were 5, the same as optimal number of clusters. F7, F8, FC5, FC6 that showed the lowest mean error ratios were decided as input data and when the maximum repetition number was 50,000, learning was terminated. When the error ratios with the least value were compared, the error ratio of the genetic algorithm was 5.87% and those of the back-propagation algorithm was 6.21%.

As a result of comparing performances of the algorithms, difference between the two learning algorithms was not significant, but the genetic algorithm converged faster than the back-propagation algorithm. Fig. 10 shows comparisons of error ratios between learning algorithms according to the number of generations. When EEG electrodes were F7, F8, FC5, FC6, the mean error ratio was 5.87%, the lowest.

As seen in Table 7 and Table 8, when exactness in learning data and test data was analyzed according to each part, total exactness values were 94.2% and 92.3%. As for exactness according to each of the results, exactness was equal in both algorithms. Therefore, it means that generalization of the system proposed in this study is possible.

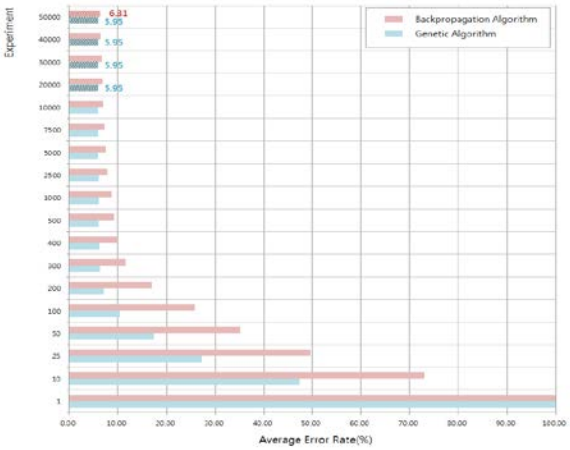


Figure 10. Comparing Learning Average Error Rates by Number of Generations

TABLE VII. ACCURACY ANALYSIS OF LEARING DATA

Data Information	Learning Data Accuracy (%)					
	1	2	3	4	5	Average
Arm Bending	93.8	93.9	93.7	93.6	93.8	93.7
Arm Unfolding	94.6	94.7	94.4	94.9	94.6	94.6
Hand Closing	94.1	94.0	93.9	94.3	94.1	94.1
Hand Opening	94.2	94.8	94.6	94.6	94.5	94.5
Average	94.6	94.3	94.1	94.3	94.2	94.2

TABLE VIII. ACCURACY ANALYSIS OF TEST DATA

Data Information	Test Data Accuracy (%)					
	1	2	3	4	5	Average
Arm Bending	91.4	91.2	91.2	91.3	91.5	91.3
Arm Unfolding	92.7	92.4	92.2	92.0	91.9	92.2
Hand Closing	93.4	93.5	93.1	92.7	92.8	93.1
Hand Opening	92.4	92.7	92.9	92.4	92.2	92.9
Average	92.4	92.4	92.3	92.1	92.1	92.3

IV. CONCLUSION

This study measured and analyzed EEG of users for intelligent upper limb rehabilitation of the system this study proposed, and intended to optimize the system using a genetic algorithm. Then, this study realized a BCI-based Intelligence Upper Limb Rehabilitation Robot and proposed an intelligent fuzzy neural network model.

The intelligent fuzzy neural network model used the wavelet transform to extract features of EEG according to users' intention and remove the noise of signals. It also used the FCM (Fuzzy c-Means) system and optimal clustering evaluation for optimal composition of rules in the system. It used the asymmetrical Gaussian membership function to enhance performance and standard deviations whose right and left was decomposed and weights to express chromosomes. This study enhanced exactness of membership functions, parameters and weights in the intelligence fuzzy system proposed using a genetic algorithm.

This study evaluated the performances of the system this study proposed through results of EEG electrode positions by making four actions of arms bending and unfolding, and hands closing and opening as actions of intention recognition of specific users. It also found optimal EEG electrode positions and analyzed differences in membership functions, number of clusters, number of learning generations, learning algorithms and the wavelet settings. As a result of the performance evaluation, the most optimal EEG electrode positions were F7, F8, FC5, FC6, and exactness rates of the learning data for users' intention recognition and test data were 94.2% and 92.3%. It was suggested that the system proposed can recognize users' intention for specific actions.

Therefore, it is expected that the system proposed in this study can make continuing rehabilitation program possible in everyday life and enhance intention of rehabilitation and quality of living in users.

In the future, we need research on EEG data classification and analysis through five senses such as the color sense and touch, and user recognition in diverse ways through normalization of such data. The BCI-based intelligence upper limbs rehabilitation robot system proposed in this study and studies on rehabilitation contents should be experimentally evaluated.

ACKNOWLEDGMENT

This work (Grants No. C0531442) was supported by Business for Cooperative R&D between Industry, Academy, and Research Institute funded Korea Small and Medium Business Administration in 2017.

REFERENCES

[1] K. Wang, L. Zhang, B. Luan, H. Tung, Q. Liu, J. Wei, M. Sun, and Z. Mao, “Brain-computer interface combining eye saccade two-electrode EEG signals and voice cues to improve the maneuverability of wheelchair,” in Proc. of 2017 International Conference on Rehabilitation Robotics (ICORR), pp. 1073-1078, July 2017.

[2] Z. Su, X. Xu, J. Ding, and W. Lu, “Intelligent wheelchair control system based on BCI and the image display of EEG,” in Proc. of Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 2016 IEEE, pp. 1350-1354, October 2016.

[3] K. Kim, and S. Lee, “Towards an EEG-based intelligent wheelchair driving system with vibro-tactile stimuli,” in Proc. of 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 2382-2385, October 2016.

[4] KISA(Korea Internet & Security Agency), “ Full-scale development of samart convergence security authentication technology using EEG, ECG.” 2015.

[5] M. Teplan, “FUNDAMENTALS OF EEG MEASUREMENT,” Measurement science review, vol. 2, pp. 1-11, 2002

[6] A. Kucukyilmaz, T. M. Sezgin, and C. Basdogan, “Interntion Recognition for Dynamic Role Exchange in Haptic Collaboration,” IEEE Transactions on haptics, vol. 6, no. 1, pp. 58-68, 2013.

[7] T. Y. Kim, and S. H. Bae, “Designing an Intelligent Rehabilitation Wheelchair Vehicle System Using Neural Network-based Torque Control Algorithm,” TIIS(KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEMS), vol. 11, no. 12, pp. 5878–5904, December 2017.

[8] D. S. Benitez, S. Toscano, and A. Silva, “On the use of the Emotiv EPOC neuroheadset as a low cost alternative for EEG signal acquisition,” in Proc. of 2016 IEEE Colombian Conference on, Communications and Computing (COLCOM), pp. 1-6, April 2016.

[9] E. W. Anderson, K. C. Potter, L. E. Matzen, J. F. Shepherd, G. A. Preston, and C. T. Silva, “ A User Study of Visualization Effectiveness Using EEG and Cognitive Load,” Computer Geraphics Forum, vol. 30, pp. 791-800, June 2011.

[10] Trans Cranial Technologies, “ 10/20 System Positoning Manual,” 2012.

[11] S. J. Yoon, G. J. Kim, and C. S. Jang, “Classification of ECG arrhythmia using Discrete Cosine Transform, Discrete Wavelet Transform and Neural Network,” The Journal of the Korea institute of electronic communication sciences, vol. 7, no. 4, pp. 727-732, 2012.

[12] C. H. Lee, “User’s Intention Recognition through an Intelligent EEG Signal Classification in BCI,” Master Thesis, Jeonnam University, 2016.

[13] J. C. Bezdek, R. Ehrlich, and W. Full, “ FCM : The fuzzy c-neans clustering algorithm,” Computers & Geosciences, vol. 10, no. 2-3, pp. 191-203, 1984.

A In-Vehicle Gateway(I-VG) for the Self-diagnosis of an Autonomous Vehicle

ByungKwan Lee
Dept. of software
Catholic Kwandong University
Gangneung-si, Korea
bklee@cku.ac.kr

YiNa Jeong
Dept. of Computer Engineering
Catholic Kwandong University
Gangneung-si, Korea
lupinus77@nate.com

SuRak Son
Dept. of Computer Engineering
Catholic Kwandong University
Gangneung-si, Korea
sonsur@naver.com

KyungDeak Kim
Dept. of Computer Engineering
Catholic Kwandong University
Gangneung-si Korea
chamsol93@naver.com

HoKil Choi
Dept. of Computer Engineering
Catholic Kwandong University
Gangneung-si, Korea
hokilc@gmail.com

EunHee Jeong
Dept. of Regional Economics
Kangwon National University
Samcheok-si, Korea
jeongeh@kangwon.ac.kr

***Abstract*—This paper proposes an In-Vehicle Gateway (I-VG) for the self-diagnosis of an autonomous vehicle which supports the communication between different protocols. The I-VG working in OBD consists of an In-Vehicle Sending and Receiving Layer (InV-SRL), an InV-Management Layer (InV-ML) and an InV-Data Transformation Layer (InV-DTL). The performance analysis of the I-VG shows that the transmission delay time about message transformation and transmission is reduced by average 10.83% and the transmission delay time caused by traffic overhead is improved by average 0.95%. Therefore, the I-VG has higher compatibility and cost effect because it applies a software gateway to the OBD, compared to an existing hardware gateway. In addition, it can reduce the transmission error and overhead caused by message decomposition because of a lightweight message header.**

Keywords-Autonomous vehilce ; CAN ; FlexRay ; In-Vehicle Gateway ; self-diagnosis monitoring system ; MOST

I. INTRODUCTION

In the future, more than 200 billion devices will be connected to the cloud and each other in what is commonly called the Internet of Things (IoT) [1]. Therefore, An IoT gateway that is a physical device or software program as the connection point between the cloud and controllers, sensors and intelligent devices plays important role. IoT gateway provides a place to preprocess that data locally at the edge before sending it on to the cloud and supports communication between heterogeneous devices.

In order to process such a large amount of data, communication protocol inside the vehicle uses various protocols such as FlexRay, MOST, and CAN and Ethernet. CAN, a low-speed (1Mbps) communication method used in conventional in-vehicle communications, is used to process relatively less important vehicle control sensor data. FlexRay, a high-speed (10Mbps) communication method, is mainly used for processing control data of major parts of a vehicle. MOST and Ethernet have speeds of up to 150Mbps. MOST is a ring type network that processes multimedia data such as voice and video inside the vehicle, and Ethernet can be used as a backbone of the in-vehicle network or as a V2X communication protocol WAVE.

In order to support all of the major protocols used for in-vehicle communications and to minimize the transmission delay of gateways, this paper proposes an In-Vehicle Gateway (I-VG) for the self-diagnosis of an autonomous vehicle. The I-VG consists of an In-Vehicle Sending and Receiving Layer (InV-SRL), an InV-Management Layer (InV-ML) and an InV-Data Trans formation Layer (InV-DTL). First, the InV-SRL receives the messages from FlexRay, CAN, and MOST and transfers the received messages to the I-VG. Second, the InV-ML manages the message transmission and reception of FlexRay, CAN, and MOST and an Address Mapping Table. Third, the InV-DTL decomposes the message of FlexRay, CAN, and MOST and recomposes the decomposed messages to the frame suitable for a destination.

II. RELATED WORKS

A. On-board idagnosics

OBD is an automotive term for self-diagnosis and reporting of vehicles. Onboard-diagnostic (OBD) system is developed to detect engine operation conditions for air-pollution monitoring. [2] The basic meaning of OBD is as above and it is used in various ways. OBD related research is as follows.

In [3], exhaust gas sensors for particulate matter (EGS-PM) have been developed based on multilayer ceramic sensor technology. Sensors must be able to detect soot emissions directly after the DPF, be able to withstand harsh exhaust gas environments, and meet future OBD legislation that includes stricter requirements for monitoring the function of the filter.

The [4] reviews a soot sensor currently developed for on-board diagnostics to monitor a diesel particulate filter and detect the failure. Original equipment manufacturers have selected a resistive soot sensor for the on-board diagnostics application because of its commercial feasibility in terms of functionality and cost. The sensor accumulates soot particles

in exhaust gas on the sensing element.

In [5], Every vehicle will be equipped with a central unit which will acquire speed data (and compute the average speed) from the on board sensors of a car using an OBD device, process and deliver the data to the server by using the Zigbee protocol, with a Zigbee transmitter employed in every vehicle and the receivers mounted on light poles. These receivers then route the received traffic information to a localized server shared by several junctions or clusters.

B. *Vehicle protocol*

The CAN (bus arbitration mechanism) [6] protocol is used as standard in most vehicles today. However, it cannot handle a large number of data. Both MOST and FlexRay communication mechanisms have been designed in consideration of the communication speed of these CAN buses and the limitation of data communication load [7] The following is a related study of the vehicle protocol.

The [8] introduces a FlexRay-CAN gateway that uses node mapping. It is possible to solve the problem of the gateway of the vehicle such as the ID change of the message mapping based gateway or the software complexity.

In [9], the transmission reliability and the slot utilization are defined for FlexRay network, and their quantitative calculating formulas are deduced. To ensure reliable network transmission and maximize bandwidth utilization of network, frame packing of FlexRay is formulated as generalized integer linear programming, and an optimized model is presented.

The [10] propose a FlexRay-based software reprogramming optimization, and we analyze the performance of FlexRay and CAN FD software reprogramming. To achieve this, we performed a theoretical analysis on the software reprogramming times of the protocol, and we compared the performance of the software reprogramming optimized FlexRay and CAN FD.

III. A DESIGN OF A IN-VEHICLE GATEWAY(I-VG) FOR THE SELF-DIAGNOSIS OF AN AUTONOMOUS VEHILCE

A. *A Design of a Sending and Receiving Layer (InV-SRL)*

The InV-SRL transmits and receives the messages between the I-VG and a Protocol Bus. A transceiver hardware device is used to transmit and receive the messages of each protocol. A controller used for message processing is embedded in the MCU. The controller lies between an upper Layer and the Transceiver and stores the bus serial bits of messages in the MCU.

If the InV-SRL receives a message, the received message is transferred to the InV-ML. The transferred message is stored in an Input Message Queue. The InV-SRL can also transfer a message from an Output Message Queue to a destination protocol bus according to an output message protocol type and a destination address. Figure 1 shows the block diagram of the Transceiver and the Controller.

B. *A Design of InV-Management Layer(InV-ML)*

The InV-ML has two functions.

First, the InV-ML receives messages and extracts only the necessary data such as source and destination address etc., from the received messages to makes the received messages lightweight. Only the extracted data is stored in an Address Mapping Table. It must rapidly be mapped to an actual data in the Input Message Queue.

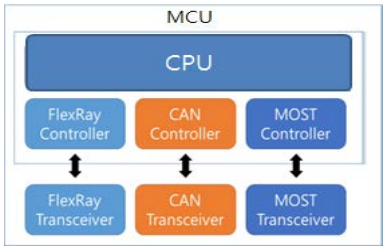


Figure 1. The block diagram of the MCU

Second, if the extracted data is stored in the Address Mapping Table for mapping, the message is stored in the Input Massage Queue according to the priority of a message. The Input Massage Queue consists of a Protocol Priority Multi-Queue and an Event Queue. There are two types in the data transferred from sensors: Real time sensing data processed in the Protocol Priority Multi-Queue and Event sensing data processed in the Event Queue. The Event sensing data is transmitted more rapidly than the Real time sensing data.

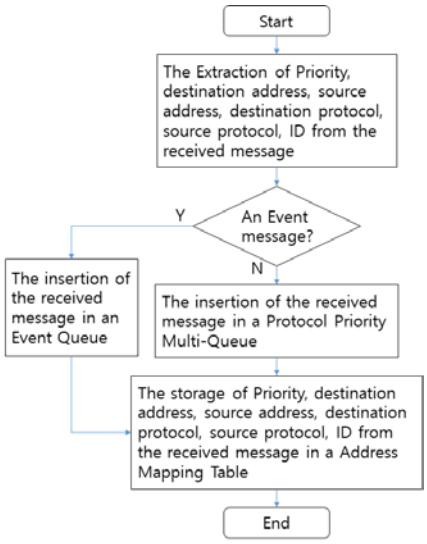


Figure 2. The flowchart to process the received message

The Protocol Priority Multi-Queue is used to decide a priority between protocols and between inner messages of a protocol. The priority of protocol itself prevents the priority collision between protocols. The priority of messages itself provides emergent messages with the highest priority and general messages with higher priority in proportion to the time when a message stays in the Input Message Queue.

The Event Queue is used to process Event sensing messages and the messages of the Event Queue has higher priority than those of the Input Message Queue.

The Output messages of InV-ML is processed in the Output Message Queue. The Output Message Queue is similar to an Input Message Queue. Figure 2 show the flowchart to process the received message.

C. A design of the InV-DTL

The InV-DTL transforms messages by combining the information of an Address Table and the original messages stored in the Input Massage Queue and transfers the transformed messages to the Output Massage Queue of the InV-ML.

The InV-DTL works as follows.

- First, the InV-DTL receives an original message and header information and transforms the received original messages and the header information by using a Mapping Table.
- Second, the mapping Table of the InV-DTL extracts the destination and priority from messages based on the Address Table information received from the InV-ML.
- Third, the Mapping Table of the InV-DTL gives a sequential number to the divided messages in case messages are divided.
- Fourth, the InV-DTL transforms messages based on a Mapping Table.
- Fifth, the transformed messages are transferred the Output Massage Queue of InV-ML.

IV. THE PERFORMANCE ANALYSIS

The model shown in figure 3 uses one gateway connected to two CAN buses, two FlexRay buses, two MOST buses to verify the LI-VEG performance proposed in this paper.

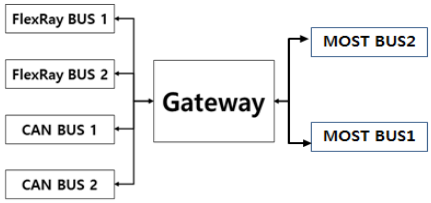


Figure 3. The simulation Model

The Main Processor of the LI-VEG uses ARM Cortex-A7 (900Mhz) and is simulated in the RAM. The performance analysis of the LI-VEG is simulated with these two experiments in PC.

A. The 1st simulation

The transmission delay caused by message transformation is measured while messages are transferred between CAN, MOST, FlexRay. The transmission delay is measured based on the 8 Message transmissions from CAN to FlexRay, the 8 Message transmissions from CAN to MOST during the 1st simulation. The transmission cycle of each message uses 14ms. The transmission delay between an existing message mapping transformation system not using a lightweight header and a message transformation system using a lightweight header is measured. Table 1 shows the base value of the 1st simulation.

TABLE I. THE SIMULATION ENVIRONMENT

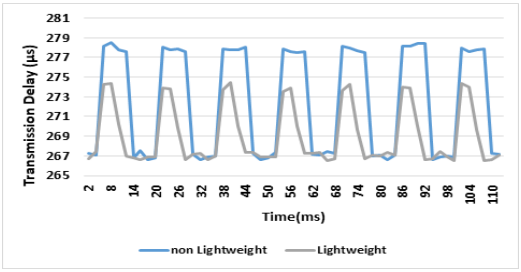
Source protocol	CAN	CAN
Destination protocol	FlexRay	MOST
A sending massage number	8	8
Transmission frequency	14ms	14ms

B. The 2nd simulation

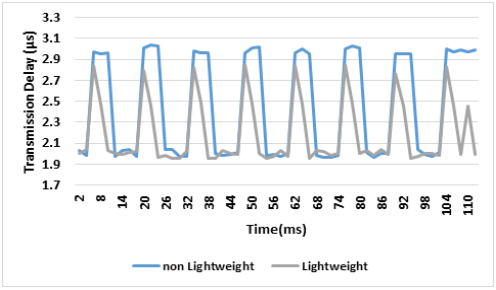
The transmission caused by overhead is measured. The overhead happens in each protocol bus because of the continuous signal from the CAN. The difference between the existing message transformation system and the proposed message transformation system are compared as network traffic increases. In the same way as the 1st experiment, the difference is measured in case of messages transformed by the other 3 protocols in the CAN.

Figure 4 shows maximum transmission delay and its stabilization when CAN messages are transferred to the other protocols. Figure 4(a) shows the transmission delay when messages are transferred from CAN bus to Flexray bus. In the transmission it takes maximum 278μs transmission delay for an existing method to store and transform a message and 8ms to stabilize the transmission delay. And it takes maximum 274μs for the proposed method to store and transform a message and 6ms to stabilize the transmission delay. Therefore, the maximum transmission delay is improved by about 2% and its stabilization time about 25%.

Figure 4(b) shows the transmission delay when messages are transferred from CAN bus to MOST bus. In the transmission it takes maximum 3.03ms transmission delay for an existing method to store and transform a message and 11ms to stabilize the transmission delay. And it takes maximum 2.84ms for the proposed method to store and transform a message and 11ms to stabilize the transmission delay. Therefore, the maximum transmission delay is improved by about 7%.



(a) CAN to Flexray



(b) CAN to MOST

Figure 3. The transmission delay when The CAN messages are transformed

V. CONCLUSION

This paper proposed the I-VG for smooth communication between in-vehicle protocols. The I-VG has some characteristics as follows.

First, the header was much lighter because only the necessary information for transformation was extracted from the header information of a message.

Second, because the Input Queue and Output Queue managing messages are divided into priority and event, the I-VG Transforms emergent messages more rapidly than an existing FIFO.

Third, the I-VG supports communication between all the primary protocols used inside a current vehicle.

According to the experiment based on these characteristics, the error rate when a message is transferred, and the transmission delay when a message is transformed are improved a little highly than a existing method.

In future the standardized gateway connected to OSEK/VDX OS and AUTOSAR must be studied based on the I-VG. And the experiment and estimation is needed in an actual vehicle, not PC.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2018R1A2B6007710).

REFERENCES

[1] G.Leen, D.Hefferman, “Digital Networks in the automotive vehicle”, IEEE Computer and Control Eng. J., vol.10, no.6, pp.257-266, 1999

[2] Man-Ho Kim, Suk Lee, Kyung-Chang Lee, “Performance Evaluation of Node-mapping-based Flexray-CAN Gateway for in-vehicle Networking System”, Intelligent Automation and Soft Computing, vol.21, no.2, pp.251-263, 2014

[3] Jin Ho Kim, Suk-Hyun Seo, Nguyen Tien Hai, Bo Mu Cheon, Young Seo Lee, Jae Wook Jeon, “Gateway Framework for In-Vehicle Networks Based on CAN, FlexRay, and Ethernet”, IEEE Transactions on Vehicular Technology, vol.64, no.10, pp.4472-4486, 2015

[4] C.E. Lin, Ying-Shing Shiao, Chih-Chi Li, Sung-Huan Yang, Shun-Hua Lin, Chun-Yi Lin, “Real-Time Remote Onboard Diagnostics Using Embedded GPRS Surveillance Technology”, IEEE Transactions on Vehicular Technology, pp.1108-1118, 2007

[5] Takeyuki Kamimoto, “A review of soot sensors considered for on-board diagnostics application”, International Journal of Engine Research, 18, pp.5-6, 2017

[6] Yi-Cheng Tsai, Wei-Hsun Lee, Chien-Ming Chou, A safety driving assistance system by integrating in-vehicle dynamics and real-time traffic information″, IEEE, pp.8-10, Nov. 2017

[7] Eunseok Choi, Sekchin Chang, ″A consumer tracking estimator for vehicles in GPS-free environments″, IEEE, pp.450-458, November 2017

[8] Yu-Shang Chou, Yu-Ching Mo, Jian-Ping Su, Wan-Jung Chang, Liang-Bi Chen, Jing-Jou Tang, Chao-Tang Yu, ″i-Car system: A LoRa-based low power wide area networks vehicle diagnostic system for driving safety″, IEEE, pp.13-17, May 2017, July 2017

[9] Junhua Chen, Dong Cao, Xunhong Lv, ″Design of FlexRay-based communication on triplex redundancy flight control computer″, IEEE, pp.27-29, Nov. 2015

[10] Yuefei Wang, Hongjun Liu, Bin Huang, Xuhui Sun, ″Frame Design for Vehicular FlexRay Network Based on Transmission Reliability and Slot Utilization″, IEEE, Vol.2, pp.447 − 452, 2016

A Self-diagnosis System for an Autonomous Vehicle(SSAV)

ByungKwan Lee
Dept. of software
Catholic Kwandong University
Gangneung-si, Korea
bklee@cku.ac.kr

YiNa Jeong
Dept. of Computer Engineering
Catholic Kwandong University
Gangneung-si, Korea
lupinus77@nate.com

SuRak Son
Dept. of Computer Engineering
Catholic Kwandong University
Gangneung-si, Korea
sonsur@naver.com

KyungDeak
Dept. of Computer Engineering
Catholic Kwandong University
Gangneung-si Korea
chamsol93@naver.com

HoKil Choi
Dept. of Computer Engineering
Catholic Kwandong University
Gangneung-si, Korea
hokilc@gmail.com

EunHee Jeong
Dept. of Regional Economics
Kangwon National University
Samcheok-si, Korea
jeongeh@kangwon.ac.kr

***Abstract*—This paper proposes "a Self-diagnosis System for an Autonomous Vehicle (SSAV)" which collects information from the sensors of an autonomous vehicle, diagnoses itself by using Deep Learning and informs drivers of the result. The SSAV consists of 3 modules which are In-Vehicle Gateway Module (In-VGM), Optimized Deep Learning Module (ODLM), and Data Processing Module (DPM). This paper improves simultaneous message transfer efficiency through the In-VGM by 15.25% and diminishes the learning error rate of a Neural Network algorithm through the ODLM by about 3%. Therefore, In addition, by transferring the self-diagnosis information and by managing the time to replace car parts of an autonomous driving vehicle safely it reduces loss of life and the cost caused by it.**

Keywords-Autonomous vehilce ; self-diagnosis monitoring system ; In-vehicle gateway ; optimiased deep learning

I. INTRODUCTION

The self-driving, autonomous vehicle has been getting lots of attention, due to significant development efforts and dramatic progress made by companies such as Google. While general use of autonomous vehicles for widespread use on public roads is likely years away, these vehicles are already being employed in "constrained" applications such as open-pit mines and farming. Among the many technologies which make autonomous vehicles possible is a combination of sensors and actuators, sophisticated algorithms, and powerful processors to execute software [1]

This paper proposes "a Self-diagnosis System for an Autonomous Vehicle (SSAV)" which collects information from the sensors of an autonomous vehicle, diagnoses itself by using Deep Learning and informs drivers of the result.

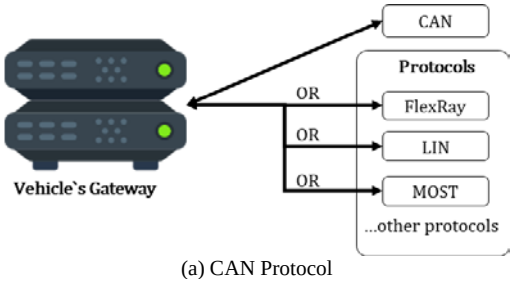
The SSAV consists of 3 modules. The first In-VGM (In-Vehicle Gateway Module) collects the data from the in-vehicle sensors, the media data like black box and radar in driving, and the control message of a vehicle and transfers each data collected through each CAN, FlexRay and MOST protocol to OBD or actuators. The second Optimized Deep Learning Module (ODLM) makes the Training Dataset on the basis of the data collected from in-vehicle sensors and reasons the risk of vehicle parts and consumables and the risk of other parts influenced by a defective part. The third Data Processing Module (DPM) based on Edge Computing has the Edge Computing based Self-diagnosis Service (ECSS) to improve the self-diagnosis speed and reduce the system overhead and the V2X based Accident Notification Service (VANS) to inform adjacent vehicles and infrastructures of the self-diagnosis result which was analyzed in OBD.

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. NRF-2018R1A2B6007710).

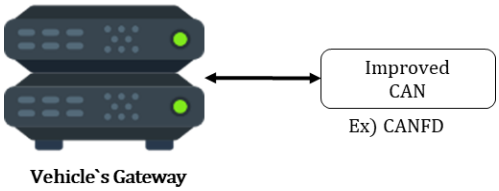
II. RELATED WORKS

A. In-vehicle IoT Gateway

CAN communication, which is the standard for most vehicles today, cannot process a large number of sensor data in real time [1]. Therefore, recent researches use FlexRay [2], which is a high speed protocol, or MOST [3], which is a high speed optical communication network of multimedia devices. Figure 1(a) shows these protocols used with the CAN protocol. And other researches improve CAN protocol. Figure 1 (b) shows that the Gateway uses the improved CAN protocol.



(a) CAN Protocol



(b) Improved CAN Protocol

Figure 1. CAN Protocol

B. Vehicle and Neural Network Model

Currently, Most of self-diagnostics for vehicles use non-OEM or OEM-based OBD methods [4-6]. However, they don't use neural network model. In case of that a neural network is used in a vehicle, studies are underway on electric vehicles and connected vehicles. Besides, there are active researches on various fields.

In [7] provided a new data compression approach and validated it on a method based on neural network (NN) to detect both failures' types and degree in drive system. A brief method was introduced to preprocess training data with a comparison to the standard deviation-based method, via analyzing the linear relationship between features and patterns to be classified. This study`s results indicate the proposed method for data preprocess can significantly improve the efficiency and precision in categorizing all the faults sample especially for fault degree considered in this study.

In [8], they present a novel method to detect vehicles using a far infrared automotive sensor. Firstly, vehicle candidates are generated using a constant threshold from the infrared frame. Contours are then generated by using a local adaptive threshold based on maximum distance, which decreases the number of processing regions for classification and reduces the false positive rate. Finally, vehicle candidates are verified using a deep belief network (DBN) based classifier. This study is approximately a 2.5% improvement on previously reported methods and the false detection rate is also the lowest among them.

In this study, the generated neural network model is simulated using DBN [9], LSTM [10], and BP [11], which are learning models used in real-time processing. And neural network was selected by selecting learning model suitable for vehicle.

III. A Self-diagnosis System for an Autonomous Vehicle

A. A design of an In-Vehicle Gateway Module (In-VGM)

The data types collected from the in-vehicle sensors are diverse. The existing CAN protocol has limits on transmission capacity and speed. But, even if the Flexray and MOST protocol ensuring high communication speed and exact transmission time are used instead of the CAN, they also have another problem like cost and compatibility.

Therefore, this paper proposes the dedicated In-VGM (In-Vehicle Gateway Module) which integrates the CAN, FlexRay, and MOST protocol to solve compatibility and data distribution transmission and supports reliable and real time communication by transforming various types of messages into a message type of a destination protocol.

The In-VGM works as follows.

- First, the in-vehicle data is classified into general sensing data, emergent sensing data, control system and sensor data, media data, and OBD data.
- Second, the data collected and transferred in the sensor and control system is transferred to the In-VGM through the CAN, Flexray, and MOST protocol.
- Third, the transmission data to which ID was allocated is recorded in a mapping table and transmission paths are set on the basis of the ID.
- Forth, because the size of data collected and transferred by CAN communication is very small, compared to that of the Flexray and MOST, the Header information(Start Message Delimiter, Length of message, and Identifier) is to be added to every Fragment so that the Flexray and MOST data can be transferred to the CAN protocol.

B. A Design of an Optimized Deep Learning Module (ODLM)

The ODLM proposed in this paper works in a Cloud Environment and consists of two sub-modules. The ODLM generates Training Data Set with the sensor data of a vehicle, starts a learning and self-diagnoses the vehicle if the learning is completed. The Cloud transfers to a vehicle manager the self-diagnosed result and the difference between current sensor data and normal sensor data.

The first Vehicle Part Diagnosis Sub- module (VPDS) generates a Deep Learning Model (DLM) which diagnoses the condition of parts by using the data collected by the in-vehicle sensors. The DLM uses the Back Propagation algorithm with high accuracy and speed for the safety of a vehicle. The Training Data made out of sensor data is used for the learning of the DLM. Because a Back-propagation algorithm does not decide the number of Hidden layers and Nodes, the number of them must be decided dynamically according to the input sensor data and the number and types of output parts.

The second Total Diagnosis Sub-module (TDS) diagnoses the total condition by using the real-time sensor data of a vehicle and the diagnosed result by the DLN as input data. The TDS is a Lightweight Neural Network Module which diagnoses a vehicle’s total condition by entering each diagnosed result of the autonomous vehicle parts computed by the DLM.

C. A design of a Data Proceesing Module (DPM) based on Edge Computing

The DPM proposed in this paper works as follows.

First, it improves the self-diagnosis speed of a vehicle and prevents system overhead by using ECSS (Edge Computing based Self-diagnosis Service).

The ECSS consists the self-diagnosis system of the OBD and the Main Cloud Server. Because the Main Cloud Server transfers the continuous update of the DLM to vehicles, it improves the exactness of the DLM. A vehicle improves the speed the self-diagnosis by using the DLM transferred from the Main Cloud Server.

Second, it has the V2X based Accident Notification Service (VANS) informing the adjacent vehicles and infrastructure of the self-diagnosis result analyzed in the in-vehicle OBD.

The VANS in this paper informs the adjacent vehicles and infrastructure of the diagnosis result transferred from the Main Cloud Server.

IV. The Performance Analysis

To analyze the ODLM performance, the ODLM is compared with the existing Long Short Term Memory (LSTM)

and Deep Belief Networks (DBN) in the computation speed and reliability of a Neural Network Model. TrainingDataSet and Test Data are used to do them. It learns the TDS by using Back propagation, LSTM and DBN and decides the computation speed and reliability with the computation time and error rate occurring in the computation of TestData.

Table 1 shows a control variable value for simulation. The simulation is done 3 times on the basis of the control variable and compares a Training Dataset learning speed, the response speed of Test Data analysis, and an error rate.

TABLE I. A CONTROL VARIABLE VALUE FOR SIMULATION.

Feature	Size
Number of Hidden Layer	2
Numder of each hidden layer`s Node	20
Learning Rate	0.1
weight(all)	0.2
TrainingDataSet`s Numder of Sensor	20

The 1st experiment compares the means of response time by classifying Test data according to the kinds of sensors and the result is shown in figure 2.

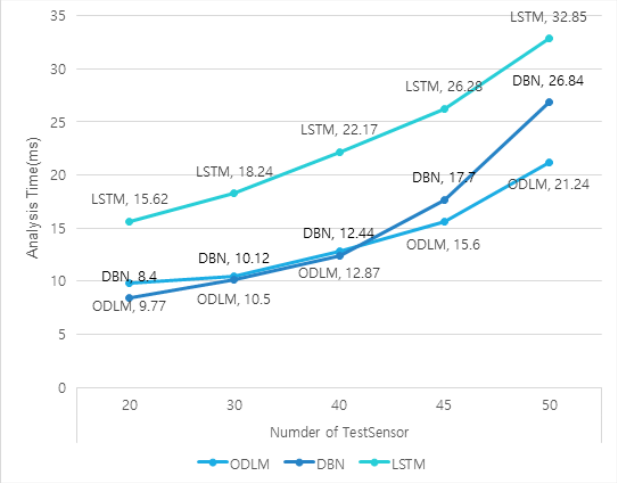


Figure 2. The comparison of the response time according to the kinds of sensors

Therefore, the analysis time of the LSTM is the longest and the more the data, the shorter the analysis time of the ODLM.

V. CONCLUSION

This paper proposes "a Self-diagnosis System for an Autonomous Vehicle (SSAV)" which collects information from the sensors of an autonomous vehicle, diagnoses itself and informs drivers of the result. The system has some advantages as follows.

First, it collects information from the sensors of an autonomous vehicle, diagnoses itself and informs drivers of the result accurately.

Second, it guarantees compatability by transforming different types of data from different communication protocol into the same type of data.

Third, it decides whether vehicles have some faults by informing a vehicle manager of the difference between in-vehicle sensor data and the normal numeric of sensors and the integration result of the self-diagnosis.

Fourth, it improves the response speed processing sensor data and reduces overhead by using Edge Computing.

Fifth, it prevents a chain collision by informing adjacent vehicles and infrastructures of accidents and dangers in advance.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2018R1A2B6007710).

REFERENCES

[1] Guoqi Xie, Gang Zeng, 0Ryo Kurachi, Hiroaki Takada, Zhetao Li, Renfa Li, Keqin Li, "WCRT Analysis of CAN Messages in Gateway-Integrated In-Vehicle Networks". IEEE Transactions on Vehicular Technology, 66, pp.9623-9637, 2017.

[2] William Wong, "design FAQs: FlexRay". Electronic Design, vol.54, no.30, 2006.

[3] Zhipeng Li, Houhua Jing, Zhiyuan Liu, "Design of application layer software and interface circuit for the MOST network in vehicle". Control Conference (CCC), 2013 32nd Chinese, 26-28 July 2013, pp.7720-7725, IEEE, 2013.

[4] Li-ye Wang, Li-fang Wang, Weilong Liu, Yu-wang Zhang. "Research on fault diagnosis system of electric vehicle power battery based on OBD technology," 2017 International Conference on Circuits, Devices and Systems (ICCDS), 5-8 Sept. 2017, pp.95-99, IEEE

[5] Kavian Khorsravinia, Mohd Khair Hassan, Ribhan Zafira Abdul Rahman, Syed Abdul Rahman Al-Haddad, "Integrated OBD-II and mobile application for electric vehicle (EV) monitoring system", Automatic Control and Intelligent Systems (I2CACIS), 2017 IEEE 2nd International Conference on, 21-21 Oct. 2017, pp.202-206, IEEE

[6] Shi-Huang Chen, Wen-Kai Liu, Jui-Yang Tsai, I-Chou Hung, "Vehicle fuel pump service life evaluation using on-board diagnostic (OBD) data". Orange Technologies (ICOT), 2016 International Conference on, 18-20 Dec. 2016, pp.84-87, IEEE

[7] Zheng Zhang, Hongwen He, Nana Zhou, "A neural network-based method with data preprocess for fault diagnosis of drive system in battery electric vehicles", Chinese Automation Congress (CAC), 2017, 20-22 Oct. 2017, pp.4128-4133, IEEE

[8] Hai Wang, Yingfeng Cai, Xiaobo Chen, Long Chen, "Night-Time Vehicle Sensing in Far Infrared Image with Deep Learning", Journal of Sensors, 2016

[9] Shin Kamada, Takumi Ichimura, "Knowledge extracted from recurrent deep belief network for real time deterministic control, Systems, Man, and Cybernetics (SMC)", 2017 IEEE International Conference on, 5-8 Oct. 2017, pp.825-830

[10] Bernhard Lehner, Gerhard Widmer, Sebastian Bock, "A low-latency, real-time-capable singing voice detection method with LSTM recurrent neural networks, Signal Processing Conference (EUSIPCO)", 2015 23rd European, 31 Aug.-4 Sept. 2015, pp.21-25

[11] A. Shyna, Reena Mary George, "Machine vision based real time cashew grading and sorting system using SVM and back propagation neural network". Circuit ,Power and Computing Technologies (ICCPCT), 2017 International Conference on, 20-21 April 2017, pp.1-5.

Data prediction model using R

Qamber Ali Yagob Ali
Dept. of IT Applied System Engineering
Chonbuk National Univ.
Jeonju-si, Republic of Korea
jacobdr4200@naver.com

Jun Lee
Dept. of IT Applied System Engineering
Chonbuk National Univ.
Jeonju-si, Republic of Korea
wartwart@uaram.co.kr

Hyoung Jin Kim*(Corresponding Author)
Dept. of IT Applied System Engineering
Chonbuk National Univ.
Jeonju-si, Republic of Korea
kim@jbnu.ac.kr

***Abstract*— This study is about data analysis and prediction model using R open source language, R is a language and environment for statistical computing and graphics so it's a full function programming language. In this study through using the Titanic datasets, we created a model that predict the survival rates of the test dataset, we used the Train dataset that has the survived variable with levels of "0"(perished) and "1"(survived) data and the test dataset with no survived levels to predict the survival rates on the test. In this study, we used the R functions, packages, and machine learning algorithms that provided in R, to combine, analysis and splitting the data we created utility functions to make features and predictive potential values for our new variables. We Analyzed the Training data set by cross-validation with 82% accuracy, visualization, and decision tree and then leveraged it to test data to predict the survival rate on test data.**

Keywords-Splitting data, utility functions, Cross validation, Analysis, Accuracy, Visualization, Prediction

I. INTRODUCTION

Globally, technology and development are constantly changing, and every year new technologies and method are being developed every year. Among them, data and data analysis techniques have become of interest to companies and schools and become an important issue. Much research has been done on prediction methods and prediction models for analyzing and analyzing big data and general data. This is because mass data itself requires a high probability of accuracy along with fast technology and analysis that finds the correlations and potential values of variables and proves the accuracy. In this paper, we use R statistical programming language to create utility functions to discover the correlation of valuable variable and splitting the data to knows what's going on between survival and perished using RStudio, which has recently been attracting attention as a big data analysis tools, We will analyze the Titanic data to create a survival prediction model. [1] In this study, we present a prediction model using the algorithms and packages of ggoplot2, random forest, rpart, caret, varImpPlot provided by R.

II. BODY

A. *Test data*

With the rapid development of IT technology, data production is so fast that momentarily goes beyond general data to big data. Thus, mass data is being an issue to corporations and public institutions, and various analytical techniques have been developed. In this study we use Titanic data. In this study used the data of Titanic ship that crashed with an iceberg in April 15.1912 and almost 2.200 passengers have perished and 1.514 passengers survived, Dataset configurations are separated by rows and columns, columns are variable names of data, and rows are data and data types in variables. Train data set configuration is shown in [Fig 1].

```
> str(train)
'data.frame':   891 obs. of  11 variables:
 $ Survived: Factor w/ 2 levels "0","1": 1
 $ Pclass  : Factor w/ 3 levels "1","2","3
 $ Name    : Factor w/ 891 levels "Abbing,
 $ Sex     : Factor w/ 2 levels "female","
 $ Age     : num  22 38 26 35 35 NA 54 2 2
 $ SibSp   : int  1 1 0 1 0 0 0 3 0 1 ...
 $ Parch   : int  0 0 0 0 0 0 0 1 2 0 ...
 $ Ticket  : Factor w/ 681 levels "110152"
 $ Fare    : num  7.25 71.28 7.92 53.1 8.0
 $ Cabin   : Factor w/ 148 levels "","A10"
 $ Embarked: Factor w/ 4 levels "","C","Q"
>
```

[Fig 1] The train data configuration

The test data set configuration is the same as the train data set, but there is no survival data variable (Survived) in the test dataset. Test data set configuration is shown in [Fig 2].

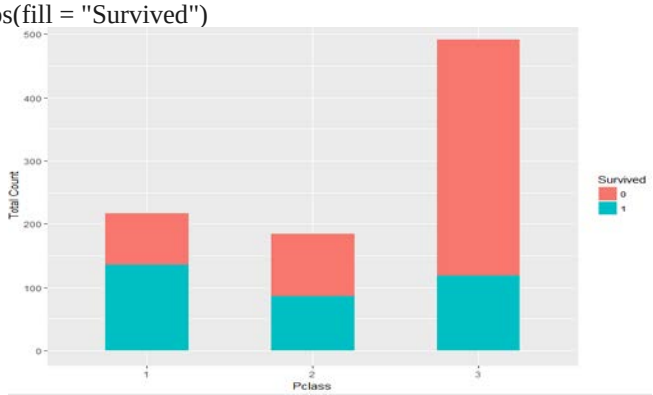
```
'data.frame':   418 obs. of  10 variables:
 $ Pclass  : int  3 3 2 3 3 3 3 2 3 3 ...
 $ Name    : Factor w/ 418 levels "Abbott,
 $ Sex     : Factor w/ 2 levels "female","m
 $ Age     : num  34.5 47 62 27 22 14 30 26
 $ SibSp   : int  0 1 0 0 1 0 0 1 0 2 ...
 $ Parch   : int  0 0 0 0 1 0 0 1 0 0 ...
 $ Ticket  : Factor w/ 363 levels "110469",
 $ Fare    : num  7.83 7 9.69 8.66 12.29 ..
 $ Cabin   : Factor w/ 77 levels "","A11","
 $ Embarked: Factor w/ 3 levels "C","Q","S"
>
```

[Fig 2] The test data configuration

B. Analiysis method

Based on the train data set of Titanic, we used machine learning algorithms and packages applied to R statistical programming language, the train data set have 891 observations 11 variables, 594 peoples perished and 342 survived. We combined the Train and test datasets by using rbind( ) to combine the data by rows and columns to handle the missing values . data.combined <- rbind(train, test.survived) and convert the survived variable as factor. data.combined$Survived <- as.factor(data.combined$Survived) and then use the ggplot2 package to plot survival rates by Pclass variable. [Fig 3] shows the result of survival by Pclass variable, the "0" red color mean perished and "1" blue color means survived

```
ggplot(train, aes(x = Pclass, fill = factor(Survived)))+
 geom_bar(width = 0.5)+
xlab("Pclass")+
 ylab("Total Count")+
labs(fill = "Survived")
```



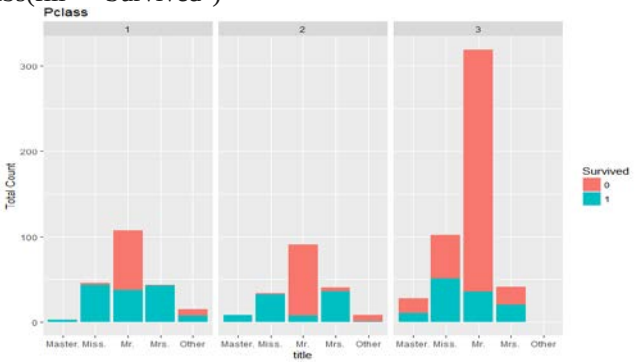
[Fig 3] The Survival rates by Pclass

Used the next code create a utility function to add a new variable named title with predictive potential values to the combined data frame

```
extractTitle <- function(Name) {
  Name <- as.character(Name)
    if (length(grep("Miss.", Name)) > 0) {
    return ("Miss.")
      } else if (length(grep("Master.", Name)) > 0) {
    return ("Master.")
      } else if (length(grep("Mrs.", Name)) > 0) {
    return ("Mrs.")
      } else if (length(grep("Mr.", Name)) > 0) {
    return ("Mr.")
      } else {
    return ("Other")
  }   }
titles <- NULL
for (i in 1:nrow(data.combined)) {
  titles <- c(titles, extractTitle(data.combined[i,"Name"]))
}
```

data.combined$title <- as.factor(titles) and now we add a new variable to data.combined data frame and use the ggplot package to gate more detail of correlation and easy understand of survival rates by Pclass and title by adding to data frame new variable we recognize lots of deference the [Fig 4] shows the survival rates of Master, Mr., Mr., Miss., Mrs., by Pclass and title variables

```
ggplot(data.combined[1:891,],aes(x=title,  fill  =  Survived))
geom_bar() +
facet_wrap(~Pclass) +
 ggtitle("Pclass") +
 xlab("title") +
ylab("Total Count") +
labs(fill = "Survived")
```



[Fig 4] The Survival rates by Pclass & title

Now we are going to handling data to get more accuracy by adding a new variable called family size by using the next code.data.combined$family.size<- as.factor(Temporary.SibSp + Temporary.Parch + 1), adding new variable may make more predictive, use the summary function to knows what going onwith family size, summary(data.combined$family.size)

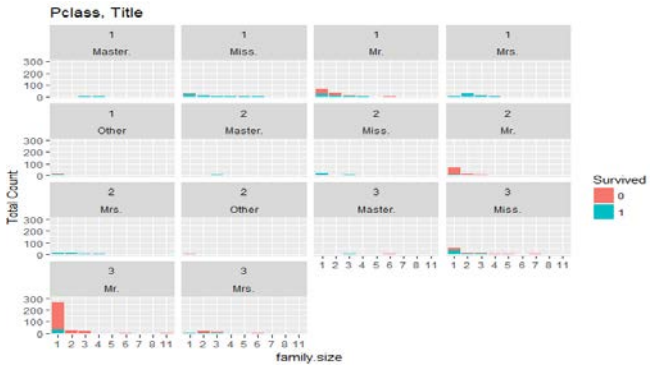
[Fig 5] shows the family relations 1~11

```
> summary(data.combined$family.size)
  1   2   3   4   5   6   7   8  11
790 235 159  43  22  25  16   8  11
>
```

[Fig 5] The sample of family size

By the adding family.size variable to the data frame used ggplot to Visualized the survival rates [Fig 6] shows the survival rate that the big family is not good to survive

```
ggplot(data.combined[1:891,],aes(x=family.size,fill=Survived
)) +
geom_bar() +
facet_wrap(~Pclass + title) +
ggtitle("Pclass, Title") +
xlab("family.size") +
ylab("Total Count") +
ylim(0,300) +
labs(fill = "Survived")
```



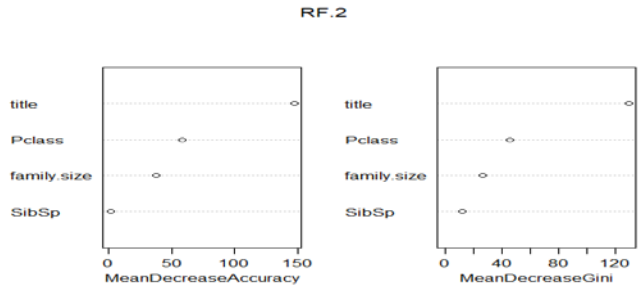
[Fig 6] shows survival rates by Pclass, title & family.size

III. RANDOM FOREST

Random forest algorithm evaluate the importance of variables and select the variable that can be used in the model. Here, varImpPlot prints a result that allows you to precisely view the variable importance you select through the label variables that are dedicated to the random forest algorithm. The Random Forest can handle numeric, categorical, and associative variables. We write the following code to create a temporary data frame with the variables most relevant to the survival rate and also create a private label variable [5]. Wrote the following code to Visualization the importance of the variables on the data frame that we made as a temporary data frame has the most correlation between survived and perished and shows the title as the most important variable [Fig 7] shows the importance of variables top-down

```
RF.train.2<-
data.combined[1:891,c("Pclass","title","SibSp","family.size")]
RF.2 <- randomForest(x = RF.train.2, y = RF.label, importance
= TRUE, ntree = 1000)

varImpPlot(RF.2)
```



[Fig 7] the selection of variables by random Forest

So we got 81% accuracy as the fallowing error rate shows 19.64% error, OOB estimate of error rate: 19.64%.

A. *Tree decision*

During the processing and analyzing data, random forest algorithm given us the most predictive potential values of variables title, Pclass, and family size so used the following code by using 3 most important variable and figure out tree decision for our predict model [Fig 8] shows the survival prediction by a single tree

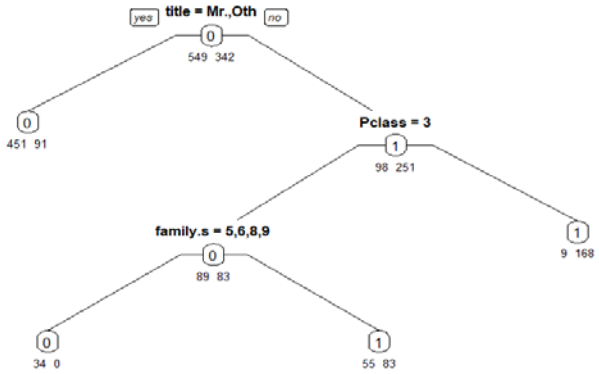
```
features <- c("Pclass", "title", "family.size")

rpart.train.1 <- data.combined[1:891, features]

rpart.1.cv.1 <- rpart.cv(12422, rpart.train.1, RF.labe2, ctrl.3)

prp(rpart.1.cv.1$finalModel, type=1, extra = 1, under = TRUE)
```

Note the right branch of the tree is always no and left branch is yes so if the passenger title is Mr. or other it means perished 2. If Pclass=3 and family size is more than 4 peoples traveling together means perished.



B. *Result*

- Finally, we split the training data using R functions, analysis and found the variables correlations.
- created utility functions make more predictive data and convert the data as needed for analysis
- sued the machine learning algorithm and R packages to make highly accurate for the prediction model
- used the packages to visualize our predictions at the end we got 82.8% accuracy and then used the following code to leverage it to the test data set and we predict 268 perished and 150people survived on test data
- test.submit.df <- data.combined[891:1309, features]

  rpart.3.preds<-predict(rpart.3.cv.1$finalModel, test.submit.df, type="class")
- table(rpart.3.preds)

REFERENCE

[1] Se-Jeong Park1), Seong-Muk Choi2), Hong-Jae Lee3), Jong-Bae Kim4) Spatial analysis using R based Deep Learning, (2016) April

[2] https://www.google.com/search?source=hp&ei=xmvmWtLqO4Ts0gSUwrmoCQ&q=titanic

[3] Se Hoon Jung, Jong Chan Kim, Chun Bo Sim , A Novel Data Prediction Model using Data Weights and Neural Network based on R for Meaning Analysis between Data (2015) April

[4] https://thebook.io/006723/ch10/03/04/03/.

[5] http://blog.heartcount.io/random-forest-ver-10

Design of Safty Monitoring IoT System for Architecture

Kiyoung, Kim
Dept. of Software Engineering
Seoil University
Seoul, South Korea

Hyeop-Geon Lee
Dept. Of Data Analysis
Seoul Gangseo Campus of Korea Polytechnic
Seoul, South Korea

***Abstract*—With the miniaturization of IoT devices and the generalization of low power wireless communication, IoT is widely used throughout society. In this paper, IoT system is designed to check construction and building safety at all times. There is a high demand in the future due to the high interest in application of the latest technology in the construction field. However, IoT communication technology, which is not a legacy communication technology in the field of architecture, belongs to the main technology and suggests a fusion system of IT and construction.**

Keywords-component; IoT, Iot Platformform, LBS

IoT devices are equipped with EEE802.11 and IEEE802.15, and UART is provided for development. In recent years, USB ports are provided for convenience of development, and home appliances equipped with IoT wireless communication module are also being commercialized[1]. The IoT market is becoming visible in 2015. IoT will be formed and spread around platforms. Z-wave is the communication technology used for IoT, and LGU + uses the Z-Wave Alliance member company Z-Wave with lower power than Wi-Fi. In addition, related companies are working on a joint development group. Typically, AllSeen is comprised of Qualcomm and is a member of LG Electronics, MS, Sharp, Haier and Cisco. OIC is led by Intel, and its members are Samsung, Atmel, Broadcom, and Dell. Thread is Google-led, and Samsung Electronics and ARM are participating.

I. RELATEDD WORK

In general, IT technologies are mutually cooperated and converged, and IoT technology is expected to be developed as a fusion technology with other technologies. On the other hand, due to the widespread use of high-speed wireless Internet, various services are expected to appear in the mobile environment, and wearable devices are being actively researched. Wearable technology also includes IoT technology in detail. Therefore, securing IoT technology development can secure a high competitiveness in future technology development and leading the market[2,3]. The following Fig. 1 is a commonly used platform for development.

Figure 1. IoT Platform

The development platforms are as diverse as arduino, raspberry-pi, and edison. For development, you have to decide which platform to use. When choosing a platform, it is important to understand the technology you need[4,5].

IoT consists of the following technologies.

- Technology related to gateway and edge device
- Communication technology that can reliably transfer large amounts of data
- Big data related technology for event processing, data analysis and recommendation
- Platforms and Enabler-related technologies to support convergence ● Device chip level security technology as well as data security through authentication / authorization
- UI / UX technology to interact with users

II. LBS TECHNOLOGIES

The LBS (Location Based Service) can provide various additional services to the existing service using the location information. Utilization of location information started from car navigation is expanding due to miniaturization of mobile terminal and decrease of price of GPS chip. On the other hand, the acceleration of the wireless network makes it possible to utilize the remote storage device. A typical example is cloud computing, and mobile cloud computing using wireless

networks is naturally occurring due to the development of IT technology and network technology. The technology related to the location information of the IoT device is as follows.

A. *Global Positioning System (GPS)*

The GPS calculates the position using a satellite. It is a navigation system using 24 geostationary satellites which controls satellites at the ground base and measures by terminal-based or network-based operation for more accurate position calculation.

B. *Assisted Global Positioning System*

The GPS-supported networks use fixed GPS receivers and the receivers are installed at regular intervals on the network at intervals of 200 km to 400 km. The receiver collects data from the terminal and calculates the position by measuring the time from the satellite to the AGPS receiver without message decoding. The time to first fix (TTFF), which is the time required for the initial positioning, is in the range of 1-8 seconds, which is better than the TTFF of 20-45 seconds supported by the GPS system. The AGPS, which is a hybrid technology, can receive location information from the network, and it is possible to perform positioning in a building or in a shadow area where GPS signals are not received. Enhanced Observed Time Difference

E-OTD is a method of locating location by installing Location Measurements Units (LMU) at multiple points on a wireless network. The LMU provides location information requested by a terminal equipped with E-OTD software. The terminal calculates the position using the difference value of the time stamp` received in the same area. The accuracy of E-OTD is 50-125m and unlike GSP, it is not influenced by weather.

C. *Time of Arrival (TOA)*

Measure accurate arrival time through three or more cell base station signals. The location is calculated using the arrival time and the transmission / reception time of the cell using the characteristics of the frequency having the fixed speed. The error occurring in the multipath section can be improved by using four different cell base stations.

D. *WFi-based location positioning*

The IEEE 802.11 standard WiFi supports 54 Mbps and connects to the wired network through the AP. The positioning of a mobile terminal using WiFi measures the strength of a signal received from two or more APs to determine its position. The signal between the AP and the mobile terminal is referred to as a beacon and the beacon signal may include an information packet.

The downlink signal transmits a beacon including an ID for identifying the AP and information necessary for the communication connection. The terminal receives a beacon signal from adjacent APs, selects an AP having a relatively high signal quality among the APs, determines a location after connection.

III. PROPOSED SYSTEM

The building safety diagnosis system using IoT is shown in the following Fig. 2. The position of sensors and Io T devices can be determined when installing the sensor in the building. The proposed system collects wind velocity, vibration, slope and temperature information for safety monitoring of buildings and sends them to the server. The server analyzes the received information and monitors the state of the building.

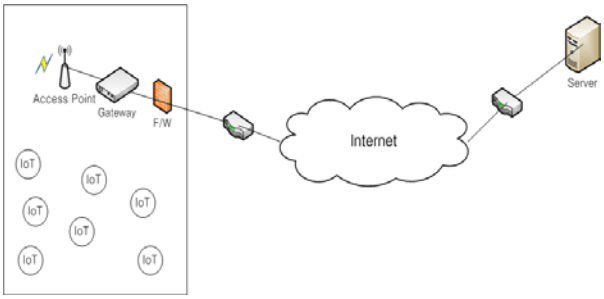


Figure 2. Proposed system

It is easy to take location information of sensors and IoT devices in the building. In case of sensor is not fixed, Proposed system use WiFi LBS technology partially.

IV. CONCLUSION

In this paper, IoT system design for safety inspection of buildings is proposed. The development of IoT, sensor, and wireless communication technology enables various applications. In the field of architecture, interest in IoT technology is high. Convergence research is needed for IoT application development to accommodate the needs of architecture. On the other hand, it is necessary to carry out research to solve the communication distance limitation in the building.

REFERENCES

[1] Jayavardhana Gubbia, Rajkumar Buyyab, Slaven Marusica, Marimuthu PalaniswamiaInternet of Things (IoT): A vision, architectural elements, and future directions, Future Generation Computer Systems 29(2013), pp.1645-1660.

[2] L. Tan, N. Wang, "Future internet: The internet of things", *Proc. 3rd Int. Conf. Adv. Comput. Theory Eng. (ICACTE)*, pp. V5-376-V5-380, 20–22 Aug. 2010.

[3] Z. Shelby, C. Bormann, 6LoWPAN: The Wireless Embedded Internet, UK:John Wiley & Sons Ltd, 2009.

[4] A. Javadpour, H. M. Tehran, F. Saghafi, "A Temperature Monitoring System Incorporating an Array of Precision Wireless Thermometers", *International Conference on Smart Sensors and Application (ICSSA)*, pp. 155-160, May 2015.

[5] https://www.mobodexter.com/wp-content/uploads/2018/05/Whitepaper-on-A-complete-Guide-to-Choosing-the-Right-IOT-Platform.pdf