Predicting Group Emotion in Kindergarten Classes by Modular Bayesian Networks

Jun-Ho Kim
Dept. of Computer Science
Yonsei University
Seoul, Korea
barabarada@sclab.yonsei.ac.kr

Sung-Bae Cho
Dept. of Computer Science
Yonsei University
Seoul, Korea
sbcho@yonsei.ac.kr

***Abstract*—Conventional methods predict emotion directly by measuring equipment like electrode. However, this approach is not suitable for education, especially for children. In this paper, we propose modular Bayesian networks for predicting the emotion with the environment information from the sensors. The Bayesian network is constructed as modules divided by Markov boundary. To evaluate the proposed method, we use data collected from kindergarten classes. The results show more than 84% accuracy and 20 times faster than the single Bayesian network.**

Keywords-emotion prediction, modular Bayesian networks, Markov boundary, kindergarten classes

I. INTRODUCTION

The emotion can affect human's mind. Recently, the sensory service to promote user's purchases by sensory stimuli has been applied in many fields. The human emotion is very important elements at the education service, as well. Emotion can affect to the effect of education, and some curricula are designed so that the students feel specific emotion.

However, the technology of changing human emotion requires to measure and predict emotion. Especially, predicting future emotion is quite difficult. Prediction needs not only measuring method, but also more information about human and surround environment. Conventional methods use many equipments like electrode attached on subject's skin. Because it disturbs educations and can affect to subject's emotion, we need another method to predict emotion.

The human emotions can be affected by stimulus of environment [2]. Shopping malls use lights to highlight goods, and movies use sound and color to incite viewers' emotion. The stimuli that we can control are listed in Table 1. Child's emotion is subject to change and affected by environment more than adult's emotion. Their emotion changes very quickly, and we must consider when we predict emotion.

Table 1. Stimuli that can control in kindergarten

Types	Elements
Sight	Light, Background movie
Hearing	Volume, Background sound
Olfactory	Scent
Touch	Temperature, Humidity

In this paper, we propose modular Bayesian networks which predict human's emotion with the relations of stimulus and emotion. The network structure is defined by four types and designed with reference to the stimulus-organism-response model by Russell [2]. In addition, Bayesian network is modularized based on Markov boundary theory in order to reduce the time complexity and make it possible to reuse the modules. The relationship of each module maintains using a virtual linking method.

We have conducted two experiments with the data collected from kindergarten classes to confirm the usefulness of the proposed method. The time complexity is calculated using the LS algorithm and execution time to evaluate the performance of the proposed method. In addition, we evaluate the accuracy.

II. RELATED WORK

A. Emotional service

Emotion is a fundamental component of human being. Traditionally, human-computer interaction focused on how to use emotion at services [1]. Recently, various services such as marketing, education, and design apply emotional components to increase effectiveness. Especially, Mehrabian-Russell proposed the stimulus-organism-response model. This model represents the relationship between environmental stimuli and emotional states that affect to decision making [2]. The state of emotion, which consists of negative, positive, arousal, and relax, is sensitively responded by the environment variables surrounding the human, such as temperature, humidity, incense, sound and so on. Donavan and Roinssiter applied the model to retail store and conducted positive analysis about customer's decision making [3]. The analysis verified the relationship between customer's emotion, purchase rate, and state of the environment in retail shops. Therefore, we need to recognize emotion, and make decisions for good performance of the emotional service.

B. Predicting emotion

The prediction of emotion from data such as facial, vocal, and text has been the focus of a large number of studies over the past several decades. Table 2 shows the emotion recognition research which use raw-data.

Black extracted features from facial image and defined rules to classify emotion [4]. Eyharabide recognized emotion

978-1-4673-9360-7/15/$31.00 2015 IEEE

using ontology and applied e-learning service [5]. Yacoub compared with four types of classification method using speech data [6]. SVM shows the best performance and decision tree was the worst. Lee classified speech data for human and robot interaction [7]. The robot recognizes emotion and responds depending on the current states of emotion. Utane applied emotion recognition to natural conversation between human and computer [8]. Hayamizu designed BN to recognize group emotion considering the previous state of emotion [9]. Previous works focused on correct recognition using raw-data. However, in the emotional service, we not only recognize current emotion, but also predict the change of emotion in the current environment in order to maintain the optimal emotion. Ivon proposed Bayesian network and regression models to predict emotion of students [15]. But, it targeted high school and ungraduated students and need special personal devices like chair, webcam, and so on.

Table 2. Research for predicting emotion

Autors	Data	Proposed method	Domain
Black and Yaccob [4]	Image	Rule-based	HCI
Eyharabide, et al. [5]	Text	Ontology	E-learning
Yacoub, et al. [6]	Speech	k-NN, SVM, NN, DT	Prediction
Lee, et al. [7]	Speech	DT	HCI
Utane, and Nalbalwar [8]	Speech	GMM + HMM	HCI
Hayamizu, et al.[9]	Facial Image, Speech	BN	HCI
Ivon, et al.[15]	Facial Image, GSR, Accelerometer	BN	Prediction

C. Modular Bayeisan networks

Bayesian network is a directional non-cyclic graph and used to infer some information using conditional probability tables (CPT). CPT can be calculated by equation (1).

$$\begin{aligned} P(U) &= P(A_1, A_2, \dots, A_n) \\ &= P(A_1)P(A_2|A_1) \dots P(A_n|A_1, A_2, \dots A_{n-1}) \\ &= \prod_{i=1}^{n} P\left(A_i|pa(A_i)\right). \end{aligned} \quad (1)$$

Modular Bayesian network is extension of Bayesian network and consists of multiple BN units (BN modules) which are connected with other units according to their causal relationships.

BN consists of a variable V, edges $E = (V_i, V_j)$, and conditional probability table $P(V)$. A BN module $\Psi_i = (G_i, P_i)$ is a Bayesian network represented by $G_i = (V_i, E_i)$ where V_i's are variables and $E_i = (V_i, V_j)$ are directed edges from V_i to V_j and P_i is the conditional probability. A BN module is a basic unit of problem domain for perceiving contextual information.

MBN consists of a 2-tuple (M, R) where M represents BN modules, and R indicates the causality between BN modules. Let two BN modules be defined as $\Psi_i = ((V_i, E_i), P_i)$, and $\Psi_j = ((V_j, E_j), P_j)$, and have an influence on each other. Then, a link $R = \{< \Psi_i, \Psi_j > | i \neq j,\ V_i \cap V_j \neq \emptyset\}$ is created and able to be affected by defining a sharing node.

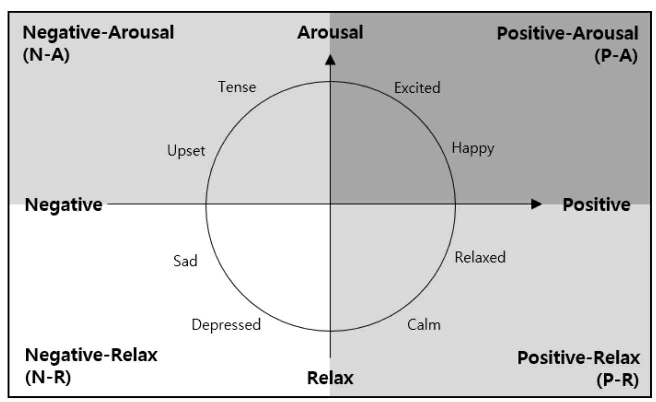
III. The Proposed Method

A. Classifiing emotions

At the kindergarten class, it is impossible to measure or predict students' emotion exactly. Therefore we categorize emotions into four groups. We map emotions to two dimensions. One dimension is a positive degree, and the other dimension is an arousal degree. Each dimension can be represented by two elements. A positive degree is represented by positive and negative, and an arousal degree is represented by arousal and relax. This dividing method is called as Valance-Arousal (V-A) model proposed by Russell [2].

V-A model comes from the employed strategy that simplifies the problem of classifying the six basic emotions to a three-class valence-related classification problem. V-A model reduces the dimensional emotion classification problem to a two-class problem or a four-class problem. For instance, Wagner et al. analyzed four emotions, each belonging to one quadrant of the V-A emotion space: high arousal positive valence (joy), high arousal negative valence (anger), low arousal positive valence (relief), and low arousal negative valence (sadness).

We compose four emotion groups to positive-arousal (P-A), positive-relax (P-R), negative-arousal (N-A), and negative-relax (N-R) using V-A model. It can be represented in coordinate plane. Figure 1 shows the V-A coordinate plane. Each emotion group has own area that includes emotions.



B. I/O elments of Bayesian networks

Figure 1. V-A model coordinate plane

At the proposed Bayesian networks, input is determined by the changeable stimuli in the kindergarten classes. The space for kindergarten class has limited size, and there is many restrictions to control environment.

The input and output of the proposed Bayesian networks are listed in Table 3. Input elements are possible stimuli and a current emotion group. Each stimulus has three or four states that are changeable in the kindergarten, and current emotion is the four emotions categorized by V-A model. The output is an emotion group predicted. The prediction time is set to a minute.

Table 3. I/O elements of the proposed Bayesian networks

Types	Elements	Possible states	Comment
Input	Brightness	200lx, 700lx, 1000lx	Brightness of the light
	Color Temperature	1000K, 3000K, 7000K	Color of the light
	Sounds	S_1, S_2, S_3, S_4	Background sounds
	Volume	20db, 40db, 60db	Volume of sound
	Scent	lemon, jasmin, lavender, rose	
	Temperature	18^0C, 23^0C, 25^0C, 28^0C	
	Humidity	30%, 40%, 50%, 70%	
	Current emotion	P-A, P-R, N-R, N-A	
Output	Predicted emotion	P-A, P-R, N-R, N-A	

C. Tree-structured Bayesian network

To minimize the time complexity of sensitivity prediction using the Bayesian network is designed for a network in a tree structure to reduce the computation time. We separate modules to reduce the processing time and the complexity. The initial structure of the Bayesian network is constructed in the form that every node is connected with each other.

Maximum spanning tree algorithm is used to construct tree-structured Bayesian network. Maximum spanning tree can represent simple tree that has no-cycle and the minimum number of edges. We assume that the weight of edge means influence when the edge is removed. Let the state of two nodes be X and Y, and the state of output nodes be C. Then, we can represent the weight of each node I_p as equation (2):

$$I_p(X,Y) = \sum_{x\in X, y\in Y, c\in C} P(x,y,c) log \frac{p(x,y|c)}{p(x|c)p(y|c)}. \quad (2)$$

After calculating the weight of each edge, we can construct maximum spanning tree from Bayesian network. Edges are sorted by decreasing order, and select every edges that make no cycle. The converting process from initial Bayesian networks to tree-structured Bayesian network (TBN) is described by algorithm in Figure 2.

D. Modular Bayesian networks

We can expedite the computing speed by reducing the complexity of Bayesian networks. One of the best known and most useful features of modular modeling is the ability to reduce the complexity of model by the expert knowledge.

Input: Single Bayesian Network (SBN)
Ouput: Tree-Structured Bayesian Network (TBN)

TBN = SBN
Remove all node in TBN

While (N_i is node in SBN)
  Select node N_i
  Select edge set E_i that includes every edge of N_i

  While (E_i is not empty)
    Select edge E_{ij} that has maximum weight in E_i
    Make connection N_i and N_j TBN

    If (cycleExist(TBN))
      Delete connection between N_i and N_j
      Delete E_{ij} from E_i
    Else
      Escape the loop

Figure 2. Algorithm to construct tree-structured Bayesian networks

After calculating the Markov boundary, we find the middle nodes that are overlapped more than two Markov boundaries. Middle nodes are border of modules. Modules are divided by the middle node, and middle nodes are removed from both modules. Both modules are connected with a pair of virtual nodes inserted in each module. Modules M_i, and M_j can be represented as equation (3):

$$\begin{aligned} M_i &= \left(B_i(TBN) - Md_{ij}\right) \cup V_{ij} \\ M_j &= \left(B_j(TBN) - Md_{ij}\right) \cup V_{ji} \\ V_{ij} &= V_{ji} = \{S_0, S_2, \dots, S_n\} \end{aligned} \quad (3)$$

Md_{ij} is a middle node between M_i and M_j, and V_{ij} means the virtual node that is in M_i and connects M_i and M_j. Virtual nodes have the same states and conditional probability table with the middle node. Virtual nodes in a pair have the same probability in each state. By this process, we can convert tree-structured Bayesian network to modular Bayesian networks (MBNs).

E. Bayesian network to predict gruop emotion

The modular Bayesian networks constructed to predict emotion in the kindergarten class have six modules and fifty-five nodes. The constructed modules are divided to the three parts: "Collecting stimuli data part", "Combining stimulus part", and "Predicting emotion part". The constructed MBNs are represented in Figure 3.

There are four modules for collecting stimuli data, touch, sight, hearing, and olfactory sensory. We name them by the role of each module. They collect stimuli data about each sense like sound type or volume. "Combining stimuli" module gathers their results. It combines data and predict emotion that calculated with stimuli. "Predicting emotion" module predicts emotion from the previous emotion, current emotion, and the result of "Combining stimuli" module.

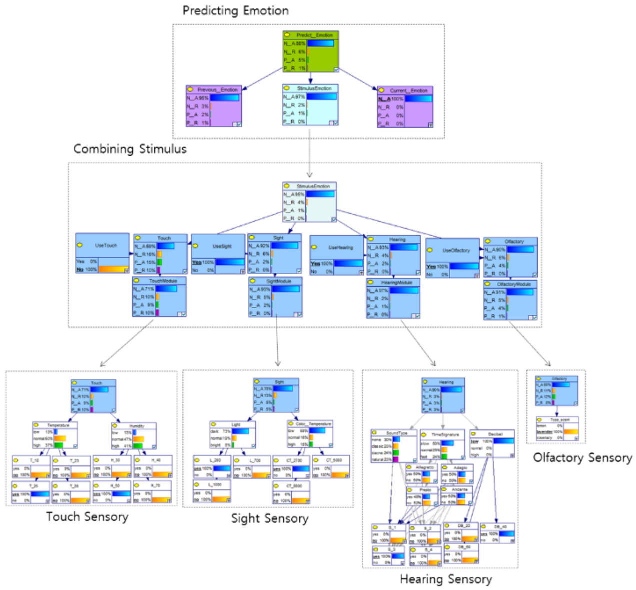


Figure 3. Modular Bayesian networks to predict emotion in the kindergarten classes

IV. EXPERIMENTS

A. The data set

We collected real data for ten children from the kindergarten classes where volume indicator, thermometer, hydroscope, illuminometer, and color sensor were installed. We used VIBRA system [12] to measure group emotion.

Classes for collecting data are processed for five days, and two classes in a day. Data were collected every minute and each class was processed in twenty minutes. We collected two-hundred data from all classes. Data are constructed with brightness, color temperature, humidity, volume, sound, video, scent, current emotion, and next emotion.

Table 4 shows the number of data in each emotion. Positive emotional data is almost three times more than negative data. It is caused by curriculum of kindergarten classes that makes children more positive.

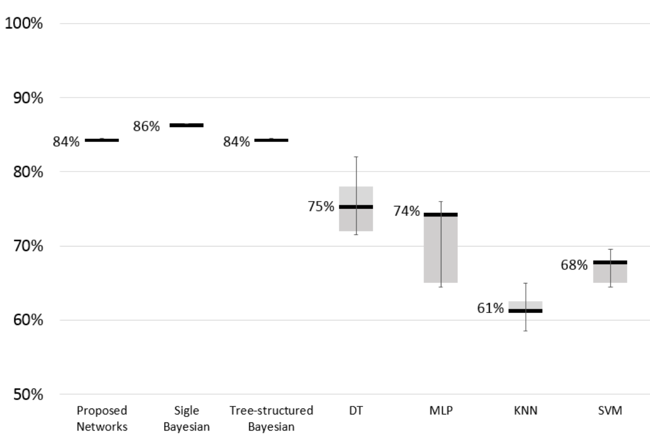


Figure 4. Accuracy comparing result

B. Accuracy test

We compared accuracy with single and tree-structured Bayesian network, and other classification methods: Decision Tree (DT), Multi-Layer Perception (MLP), K-Nearest Neighbor (KNN), and Support Vector Machine (SVM). Classification methods were evaluated by five-fold cross validation.

Figure 4 shows the results of accuracy. The proposed method shows accuracy of 84%, which is little lower than single Bayesian networks 86%. However, the proposed method produces much higher accuracy than other machine learning methods. The most outstanding method is decision tree that shows accuracy of 75%, and the worst method is k-nearest neighbor that shows accuracy of 61%. This means the proposed method has better performance than other prediction methods, and similar performance with single Bayesian network. If the situation needs faster speed than accuracy, using the proposed method can be better choice than single Bayesian network.

Table 4. Number of data in each emotion

	P-A	P-R	N-A	N-R
Number of data	80	73	18	29

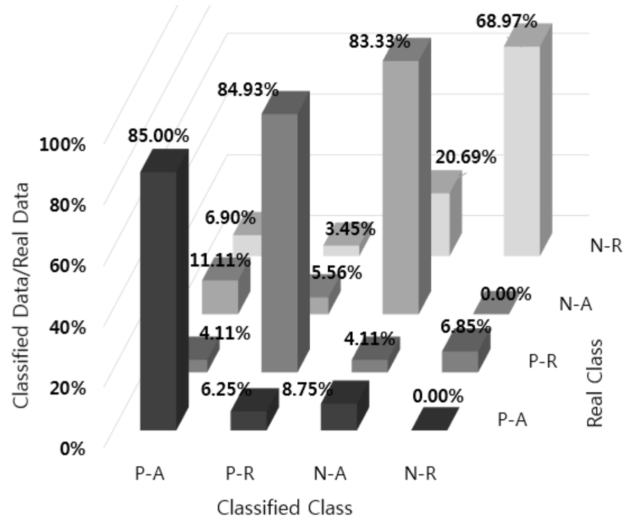


Figure 5. Classified results of the proposed method

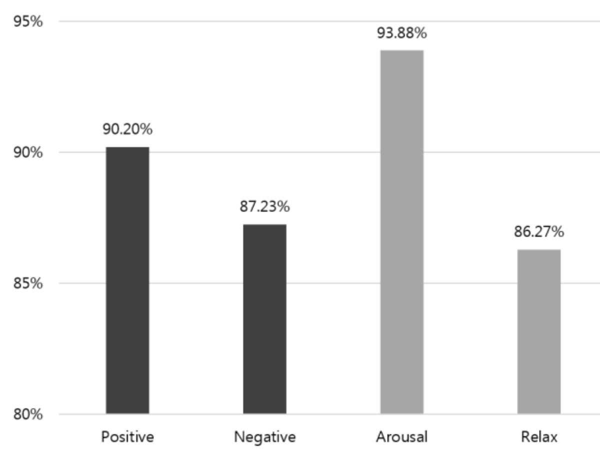


Figure 6. Accuracy of classifying each dimension

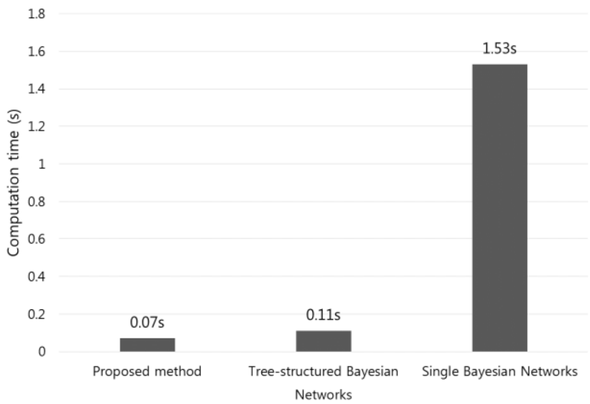


Figure 7. Computation times

Figure 5 shows the classified results of the proposed method. The emotions P-A and P-R show better classification performance than N-A and N-R. It may be affected by number of data. Figure 6 shows the classification result with only a dimension. In this case, accuracy has jumped to maximum 93.88% and the worst accuracy also shows 86.27%. The positive class is shown better performance than negative class in almost 3%, and the arousal class shows much better performance than relax class in almost 8%. It shows the positive and arousal emotion has distinguishing property comparing with negative and relax. This difference of performance may come from the characteristic of kindergarten class. At the kindergarten class, curricula are constructed with joyful contents, and it makes lack of data about negative and relax emotions. This problem could be resolved by gathering more data, until the data can reflect the curriculum`s characteristic enough.

C. Computation time

We compared computation time of the proposed method; tree structured Bayesian networks and single Bayesian network. We calculate the computation time by average of hundred times of tests. In kindergarten classes, most of curricula are constructed in a minute interval. Then, we must predict emotion in 10 seconds to save time for other tasks that improve the quality of classes using predicted emotion.

In this result, the proposed method is 1.5 times faster than tree-structured Bayesian network and twenty times faster than single Bayesian network. In this case, number of nodes are limited and all methods show proper computation time, even single Bayesian networks. However, if the number of nodes are increasing, the computation time of single Bayesian network would increase sharply than the proposed method.

V. CONCLUSION

In this paper, we have proposed modular Bayesian networks to predict emotion in kindergarten classes. To obtain modular Bayesian networks, we construct tree-shaped networks using maximum spanning tree algorithm, and divide it to modules.

The proposed method shows above 80% performance, and the performance is much better than other machine learning methods. It may show the better performance for classifying positive and arousal than negative and relax emotions. The computation time is decreased twenty times more than single Bayesian network.

For the future works, we need to collect more data to verify the proposed method, especially for negative emotions. We have to compare computation time in cases that the number of networks are increased more than current networks. Also, we need to evaluate that the proposed modular Bayesian networks can be used not in the kindergarten, but in the other institutions.

ACKNOWLEDGMENT

This work is supported by the National Strategic R&D Program for Industrial Technology (10044828, Development of augmenting multisensory technology for enhancing significant effect on service industry), funded by the Ministry of Trade, Industry and Energy (MOTIE).

REFERENCES

[1] S. Brave and C. Nass, The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, pp. 81-96, 2003.

[2] A. Mehrabian and J. Russell, An Approach to Environmental Psychology, Cambridge, MA: MIT Press, 1974.

[3] R. Donovan and J. Rossiter, "Store atmosphere: An environmental psychology approach," Journal of Retailing, vol.58, no. 1, pp. 34-57, 1982.

[4] M. Black and Y. Yacoob, "Tracking and recognizing rigid and non-rigid facial motions using local parametric models of image motion," Int. Conf. on Computer Vision, pp. 374-381, 1995.

[5] V. Eyharabide, A. Amandi, M. Courgeon, C. Clavel, C. Zakaria, and J. Martin, "An ontology for predicting students' emotions during a Quiz," IEEE Workshop on Affective Computational Intelligence, 2011.

[6] C. Lee and S. Narayanan, "Emotion recognition using a data-driven fuzzy inference system," Proc. of Eighth European Conf. of Speech Comm. and Technology, pp. 157-160, 2003.

[7] C. Lee, E. Mover, C. Busso, S. Lee, and S. Narayanan, "Emotion recognition using a hierarchical binary decision tree approach," Speech Communication, vol. 53, pp. 1162-1171, 2011.

[8] A. Utane and S. Nalbalwar, "Emotion recognition throught speech using Gaussian mixture model and hidden Markov model," Int. Journal of Advanced Reserch in Computer Science and Software Engineering, vol. 3, pp. 742-746, 2013.

[9] T. Hayamizu, S. Mutsuo, K. Miyawaki, H. Mori, S. Nishiguchi, and N. Yamashita, "Group emotion estimation using Bayesian network based on facial expression and prosodic information," Int. Conf. on Control System, Computing and Engineering, pp. 23-25, 2012.

[10] G. Kleiter, "Propagating imprecise probabilities in Bayesian networks," Artificial Intelligence, vol. 88, no.1, 1996.

[11] Y. Lee and S. Cho, "Recognizing multi-modal sensor signals using evolutionary learning of dynamic Bayesian networks," Pattern Analysis and Applications, vol. 17, no. 4, pp. 1-13, 2012.

[12] J. Park, J. Kim, and S. Hwang, "A study on emotional classification algorithm using Vibraimage technology," ACED Asian Conf. on Ergonomics and Design, 2014.

[13] Handbook of educational data mining, FL: CRC Press, Boca Raton, 2010, pp. 323-338