

Research Article

An Improved Non-Face-to-Face, Contactless Preference Survey System

KwangCheol Rim 

College of Basic & General Education, Dongsin University, NaJu 58245, Republic of Korea

Correspondence should be addressed to KwangCheol Rim; rim1201@hanmail.net

Received 26 February 2021; Revised 31 March 2021; Accepted 16 April 2021; Published 24 April 2021

Academic Editor: Hoon Ko

Copyright © 2021 KwangCheol Rim. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Conventional election-related public opinion polls have utilized the automated response system (ARS) method. The ARS public opinion polls are predicated on the convenience of use and require random telephonic responses. However, the actual response rate is less than 5%. As a result, discrepancies between recent public opinion polls and the actual election results have become an issue. In this study, we propose a system that quantifies the preferences by region, age, and gender by quantifying emotions based on the behaviors and facial expressions of the citizens passing by at the campaign site and utilizes them as basic statistics. Furthermore, a previously published facial recognition artificial intelligence (AI) was used to obtain age, gender, and various facial recognition data, along with citizens' emotions. The published facial recognition AI produced stability of over 99% recognition rate. The data structure followed a weighted reverse tree structure, and facial expressions, gender, and age were analyzed using the published facial recognition algorithm. Moreover, the expressions as well as the behaviors showing emotions were merged to gather and analyze data with weights.

1. Introduction

Till date, significant errors existed between the public opinion survey results released after elections in various countries and their actual voting results. A significant portion of the cause of these errors is attributed to the response errors. A variety of causes have been brought up as the cause of the response errors. These include low response rate, polling methods that primarily rely on phone calls to landlines, and failure in the sampling design, which is represented by the exclusion of respondents in their 20s. To reduce the response error in the public opinion polls, we must obtain ways to ensure homogeneity in the population and the sample groups through various polling methods, rather than simply focusing on improving the response rate. However, considering that voters tend to change their support as it gets closer to the election, inevitable errors cannot be avoided under the current system, which prohibits the release of public opinion survey results about one week prior to the election date, even if the public opinion polling methodology is improved. Studies seeking an empirical approach to the causes of public

opinion poll errors have also proposed a variety of causes for the public opinion poll errors. Examples of these causes include the bias introduced by the surveying agency, the validity of the wireless random digit-dialing telephone surveys, which is a part of the polling methods, the bias of the online public opinion polls, and the weighting method. However, even these studies have not adequately answered the question regarding the key problem of the public opinion poll errors [1]. There are various types of election-related public opinion polling methods. Thus, even for polls conducted on the same day, the results may show differences depending on the sampling frame and on the sampling, data collection, and analysis methods. Furthermore, changes in public opinion need to be evaluated [2].

- (1) Was the candidate support rating collectively aggregated by reflecting the strengthening of the voting intentions as the election day approaches?
- (2) Was the overall established trend based on the candidate support rate formed through the election structure?

- (3) Has the candidate support rate changed owing to election campaigns and external events?

The answers to the above three questions may vary and remain uncertain.

Among the polling results reported by Gallup Korea to the Korean National Election Survey Deliberation Commission on April 13, 2017, the response rate for contacts made through wireless phone calls was 25.1%. Although the company labels this figure as the response rate, this was derived by using the formula for the cooperation rate because it describes that 812 from a total of 3,230 people who answered the phone call completed the survey. In fact, according to the same report by Gallup, the number of failed contact cases was 10,069, indicating that 10,069 people did not answer the call. Thus, in this case, Gallup Korea actually received 812 completed surveys by making 13,299 calls instead of 3,230 calls. Therefore, the actual response rate significantly decreases when the number of failed contact cases is added to the denominator. Therefore, to improve the irrationality and uncertainty of the current election polls, in this study, we intend to derive positive and negative emotions by capturing the images of the facial expressions and behavioral responses of the citizens at the election campaign site in real time. A face recognition algorithm was incorporated into the existing non-face-to-face, contactless preference survey system to understand the situation at the campaign site in greater detail [3–5]. Furthermore, an algorithm that can apply weights based on user selection for each data being passed from a lower-level module to a higher-level module was used.

Section 2 looks into the non-face-to-face, contactless preference survey system. In Section 3, the structure and flowchart of the improved non-face-to-face, contactless preference survey system, as well as the system using the face recognition AI application and weighted reverse tree, are designed.

2. Non-Face-to-Face, Contactless Preference Survey System

By utilizing artificial intelligence (AI), the preference survey server recognizes the research subjects in the image frames received from the photographing devices. If there are multiple photographing devices, the preference survey server groups the image frames provided by each photographing device in terms of the time slot. The server recognizes at least one research subject in the grouped image frames and checks the facial expressions and behaviors of the recognized research subjects. Specifically, the facial expressions include the faces of the research subjects, and the behaviors refer to the entire upper body of the research subjects, including their hands.

The server counts the positive or negative emotions of the research subjects depending on their identified facial expressions and behaviors. Specifically, the facial expressions and behaviors of the research subjects in the image frames are checked when they are stimulated, such as when the candidate's business card is handed over to them. If the

identified facial expressions or behaviors are recognized as positive emotions, the positive count is incremented by one. If they are recognized as negative emotions, the negative count is incremented by one. Positive facial expressions include smiling with eyes, raising the corners of the mouth, and making eye contact. Behaviors associated with positive emotions include receiving promotional materials, taking pictures, shaking hands, and changing the direction of movement. On the contrary, facial expressions associated with negative emotions include frowning, lowering the corners of the mouth, and avoiding eye contact. Behaviors associated with negative emotions include not receiving promotional materials, refusing to shake hands, waving hands, and changing the direction of movement.

Figure 1 shows an example of the method for counting the positive and negative behaviors of the research subjects. This method was implemented using a smartphone application. The election campaign proceeded from the upper administrative district of the region to the lower administrative district of the region, following a reverse tree structure. In the screenshot of the smartphone at the bottom, the screen comprises counters that can count both positive and negative behaviors. This algorithm was used to test the implementation in the local mayoral elections in Korea. There were two contending candidates. A week before the election, the ARS poll showed that the gap between the two was within the margin of error, with 23 percent and 24 percent. The algorithm was more than twice as different as 29% and 60%. Since then, the election results have also doubled the gap between the two issues, indicating a stochastic appropriateness.

The third screen in Figure 1 displays a data structure wherein the positive and negative data counted for the campaign sites are gathered together. The data count represents the final total summed up by traversing up the reverse tree structure. The screen on the left in Figure 1 first classifies the campaign sites according to the regional division. The middle screen in Figure 1 shows the experimental forms of the positive and negative behaviors. The application has been implemented such that these data are aggregated and counted in the administrator screen. As a result, the citizens' positive and negative responses for each region can be checked in real time.

3. The Improved Non-Face-to-Face, Contactless Preference Survey System

The main idea of our proposal is that our survey system is designed to quantify and analyze the behavioral patterns and facial expressions of the research subjects without being in direct contact with them. The collected behavior type data can be weighted and aggregated to predict the preference for a particular candidate. Moreover, AI is used to analyze the facial data of people at the research site to organize and analyze age, gender, etc. A free, verified algorithm was used for the face recognition algorithm. Our proposed system does not require asking survey questions to the research subjects (voters) in person. Instead, the system utilizes the facial expressions and behaviors of the

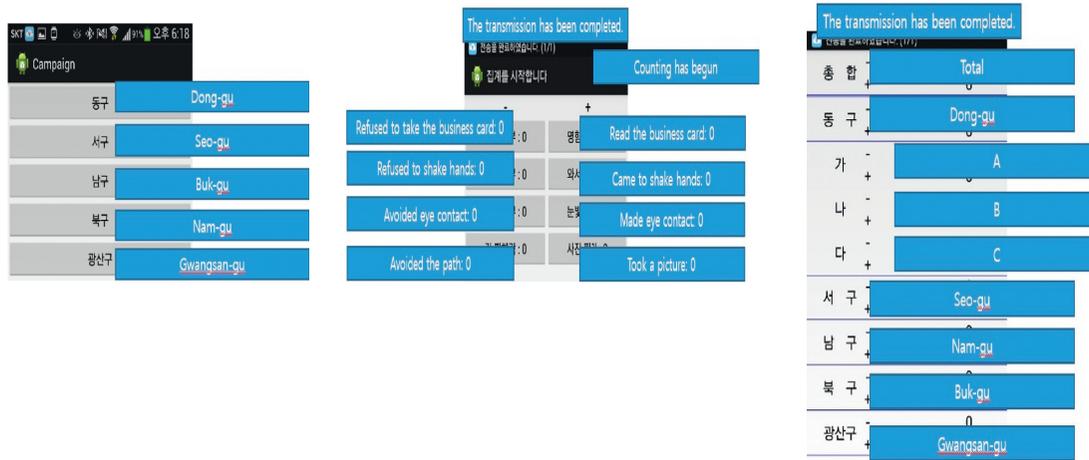


FIGURE 1: Campaign site counts and administrator screen.

research subjects at the counting site. Hence, it is a contactless preference survey system. This non-face-to-face, contactless system can analyze the facial expressions and behaviors of the users at the site of an election campaign or product promotion and survey the preference for a candidate or a product. The improved non-face-to-face, contactless preference survey system comprises photographing devices placed at the promotion site or election campaign site to obtain image frames, and the preference survey server that recognizes the research subjects in the image frames captured by the photographing devices. The server generates the preference survey results that quantify the positive and negative emotions of the recognized research subjects by judging their facial expressions and behaviors. Furthermore, the age and gender of the research subjects at the campaign site are transmitted by utilizing a previously verified face recognition algorithm. The server presets the types of facial expressions and behaviors to determine positive or negative emotions according to the promotion target. The improved system has been designed to apply weights to each data based on the positive or negative emotions determined based on the research subjects' facial expressions and behaviors in the image frames captured by the photographing devices. The system has replaced the experiment with structural design and modality implementations because it uses known facial recognition algorithms.

Figure 2 is a flowchart of the improved non-face-to-face preference survey. The process for the improved non-face-to-face, contactless preference survey system is as follows.

- (1) Input: uses the screen where counting can be performed and creates a work table.
- (2) AI: saves and analyzes various data using a verified face recognition algorithm (Naver Clova etc.). The counting program saves pictures or videos (duplicate pictures and videos are detected). The saved data are sent to the server.

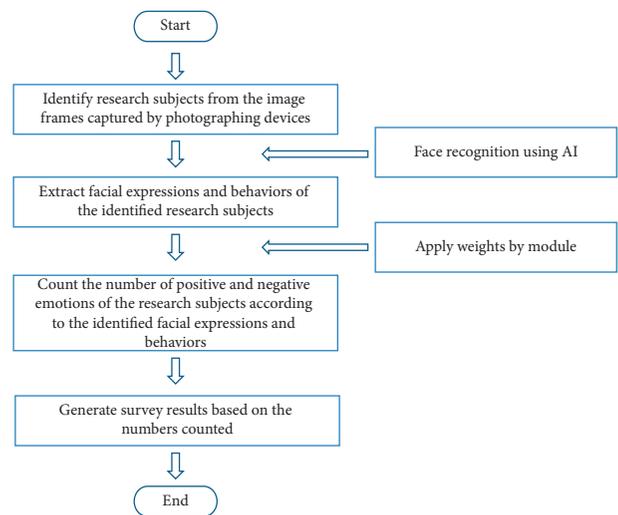


FIGURE 2: Flowchart of the improved non-face-to-face preference survey.

- (3) Weights: the server can adjust the weights. The server adjusts the weight for each behavior and stores the weighted data and unweighted data separately. The database is designed using the weighted reverse tree structure.
- (4) Server: the centrally managed main server can control all data and each site. The server classifies and stores the data from each site. The server then transmits the data to the administrator server. The server can adjust weights for each site and saves data types (by region and time slot).
- (5) Administrator server: automatically moves the data from the server to the administrator server. It organizes the database screen by time slot and region and stores basic statistics data, as well as stores basic statistics per date and region.

- (6) Output: supports printer output. It organizes the output and administration screen.

3.1. Input Process. As shown by the campaign site image in Figure 3, cameras C1 and C2 capture the reactions of groups of people P1 and P2, while the stage is used for the election campaign. Once the images are transferred to the server, AI-based facial recognition and behavior recognition are applied to generate data on the positive and negative behaviors of citizens. Subsequently, the generated data are quantified by applying weights according to the given behavior and subsequently analyzed as basic statistics to transmit and display the results to the administrator screen in various modes. In Figure 3, *B* denotes people who turned around after seeing the campaign site and *A* denotes people showing a reaction toward the campaign, such as asking the candidate for a handshake. Furthermore, *R* represents the prevention of duplicate image capturing by the AI system when the areas captured by the two cameras overlap. As the data collected through this process are transmitted to the server, they are stored as *A* and *B* data, with *A* denoting a positive emotion and *B* denoting a negative emotion.

3.2. Facial Recognition AI. Facial recognition technology could be utilized in real life because the error rate has decreased, and its accuracy has improved due to advances in AI technology. In 2007, Google conducted studies on facial recognition technology and developed a tool that was tested using the Labeled Face in the Wild (LFW) database. The LFW database was published to contribute to advancements in technology and performance improvement. In 2015, it produced a face recognition rate of 99.96%. However, unlike the LFW DB, which was built using still image pictures, the accuracy dropped to 95.12% when the test was performed on a face database built using YouTube videos with various lightings and face angles [2]. Apple announced Face ID technology that performs recognition at a high speed of less than one-thousandth of a second and has an error rate of one-millionth. The Face ID technology was introduced in the iPhone X smartphone. However, recognition error and failure cases, such as errors with facial recognition of twins and recognition failure when the lighting is dark, have been reported occasionally. Alibaba adopted facial recognition technology as a secure authentication method for Alipay, which is an online payment method. In China, besides Alibaba, face recognition companies founded in the early to mid-2010s are growing rapidly, aided by the government's active support. These companies achieved an accuracy of over 99.9% based on the recognition of 10 million people in a Face Recognition Vendor Test competition held by the National Institute of Standards and Technology in 2018. Currently, the facial recognition field is considered to be in a stagnant period overall. However, the face classification field, a subtechnology of the facial recognition field, is growing. And, the emotion recognition field is believed to be in the early stages of technology. The face classification field may be appropriate from the perspective of developing

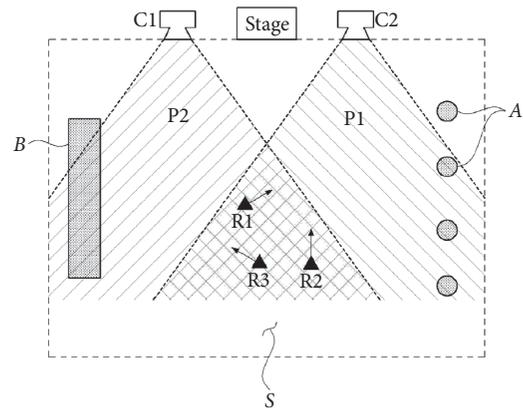


FIGURE 3: Input site image.

applied technology. However, the trend shows that there are more interests in the emotion recognition field to develop a foundational technology. The Clova platform service by the Naver Developer Center provides eyes, nose, mouth, and facial expressions of faces and processes 1,000 cases per day for free. Kakao Vision API can distinguish between face detection and product detection. MS Azure provides face detection, emotion recognition, and searching similar faces [6–10]. In this study, facial recognition by Clova was used. The Clova Face Recognition API (CFR API) receives the image data and then returns the facial recognition results in a JSON format. The CFR API provides a face detection API that recognizes the face in the image and provides the analysis information. It also provides a celebrity face recognition API, which informs which celebrity resembles the face in the image. The CFR API is an HTTP-based REST API, and it is a nonlogin open API that does not require user authentication. The CFR API does not have any limitation on the image format sent via the client's request. It merely performs facial recognition based on the first frame image of the image format, similar to a GIF. The size of the image sent to the server is limited to 2 MB.

When the facial recognition request is sent to the facial recognition server, the facial recognition server returns the analysis result data in the JSON format as a response message (Algorithm 1).

The basic processing detects faces in the input image and returns how many faces have been detected and where each face is located in what size and shape. The output information of the CFR API is as follows.

- (i) The number of faces detected
- (ii) The analysis information of each face detected
- (iii) The coordinates and size of each face detected
- (iv) The coordinates of the eyes, nose, and mouth for each face detected
- (v) The estimated gender and estimates of the face detected
- (vi) The estimated age and estimates of the face detected
- (vii) The emotions analyzed from the face detected

```

// CFR API Python code
import os
import sys
import requests
client_id = "YOUR_CLIENT_ID"
client_secret = "YOUR_CLIENT_SECRET"
url = "https://openapi.naver.com/v1/vision/face" // 얼굴감지
files = {"image": open("YOUR_FILE_NAME", "rb")}
headers = {"X-Naver-Client-Id": client_id, "X-Naver-Client-Secret": client_secret }
response = requests.post(url, files = files, headers = headers)
rescode = response.status_code
if(rescode == 200):
    print (response.text)
else:
    print ("Error Code:" +rescode)

```

ALGORITHM 1

(viii) The direction of the face detected

3.3. The Structure of the Weighted Reverse Tree Data. In a reverse n -ary tree structure, a structure wherein a specific weight is assigned to each component of the deepest module is referred to as an n -ary tree, and a structure wherein a specific weight is assigned to a module rather than a component is called a module-weighted 10-ary tree [11–15]. A weighted 10-ary tree is derived by assigning a weight of m_i ($i = 0-n-1$) to each of the n components corresponding to the deepest module and varying the size of each component. As shown in Figure 4, the state transition method for applying the classification according to the cluster analysis in the parent module and multiplying the weight, i.e., m_i ($i = 0-9$), by the sum of the components of each module according to the significance of each group is called a module-weighted 10-ary tree.

When determining the element of the deepest component, a weight is assigned to each component of x_{30} , as shown in Figure 4, through the method of assigning weight to the appearance of each component according to its significance:

$$x_{30} = (m_0x_{3000}, m_1x_{3001}, m_2x_{3002}, \dots, m_9x_{3009}). \quad (1)$$

The sum of $\sum_{j=0}^9 m_jx_{300j}$ is transitioned into the first component of x_{20} . In addition, the second to tenth components of x_{20} are transitioned into the sum of the components of x_{31} to x_{39} , respectively. Subsequently, the sum of the components of x_{20} is transitioned into the first component of x_1 . The components of the weighted reverse 10-ary tree having a depth of 3 are derived as follows:

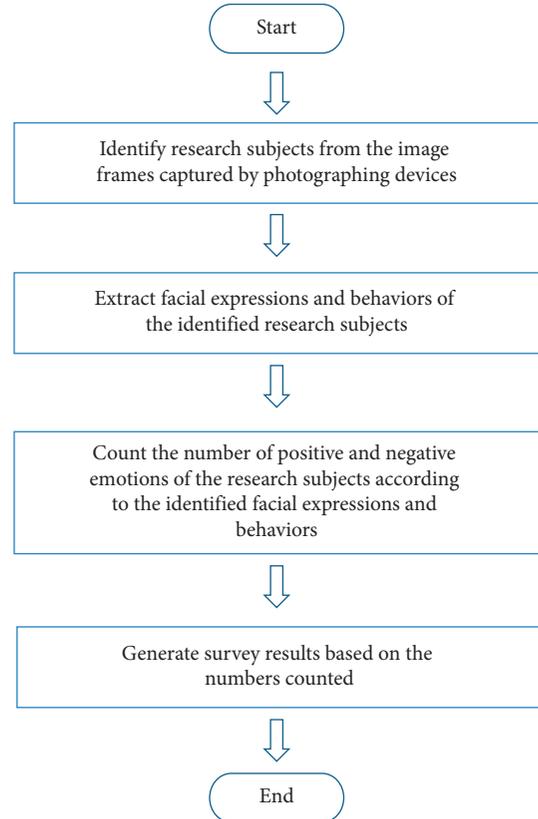


FIGURE 4: Flowchart of the non-face-to-face, contactless preference survey.

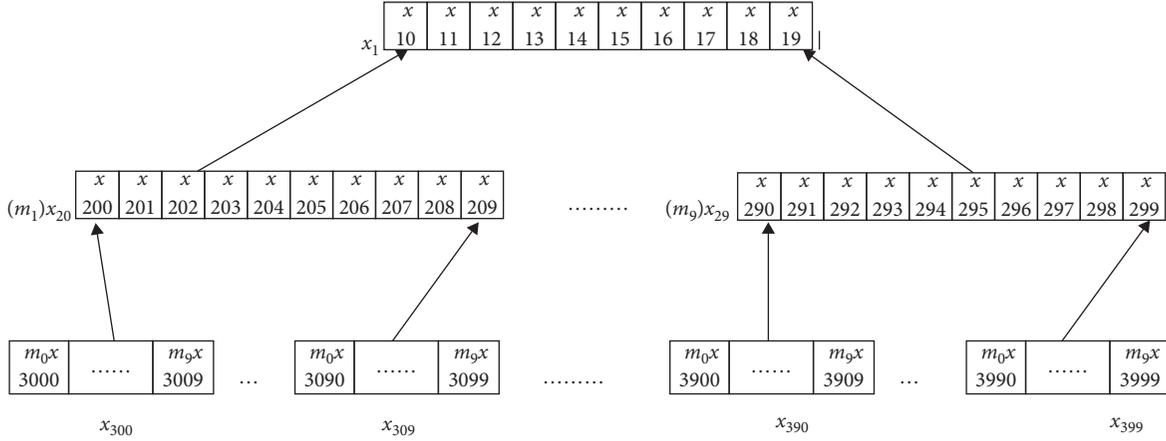


FIGURE 5: Weight of deep 3.

$$\begin{aligned}
 x_{10} &= \sum_{k=0}^9 x_{20k}, \\
 &= x_{200} + x_{201} + \cdots + x_{208} + x_{209}, \\
 &= \sum_{j=0}^9 m_j x_{300j} + \sum_{j=0}^9 m_j x_{301j} + \cdots + \sum_{j=0}^9 m_j x_{309j}, \quad m_j x_{30kj}, \\
 &= \sum_{k=0}^9 \sum_{j=0}^9 m_j x_{30kj}.
 \end{aligned} \tag{2}$$

$$\begin{aligned}
 x_{19} &= \sum_{k=0}^9 x_{29k}, \\
 &= x_{290} + x_{291} + \cdots + x_{298} + x_{299}, \\
 &= \sum_{j=0}^9 m_j x_{390j} + \sum_{j=0}^9 m_j x_{391j} + \cdots + \sum_{j=0}^9 m_j x_{399j}, \\
 &= \sum_{k=0}^9 \sum_{j=0}^9 m_j x_{39kj}.
 \end{aligned} \tag{3}$$

The weighted 10-ary tree is capable of assigning features to the component to be searched, and if the deepest component has strong features, even if the appearance frequency is low, such features can be reflected in the parent module. Figure 5 shows the weight of deep 3.

For example, suppose the facial expression of the research subject identified by the server in image frame #1 has a neutral emotion, and the behavior is “received as a promotional material,” which is determined to be associated with the positive emotion. If the research subject’s behavior in image frame #2 is verified to be “reading a business card,” a positive count of 1.5 can be generated by applying the weight to the positive emotion. Moreover, if the research subject refused to receive the business card, the situation shows a relatively weak negative emotion. In this case, counting can be performed after a weight of 0.8 is applied to the negative emotion.

4. Conclusion

With advances in smartphones, many users have begun to express a variety of opinions more easily through social media such as Twitter and Facebook. In addition, social media have become popular, and the Internet of Things has spread rapidly. As a result, multimedia content such as text, pictures, and videos has exponentially increased. As the amount of data has increased rapidly, storing, managing, and analyzing the massive scale of data using conventional methods has become difficult. Large-scale data have led to the big data phenomenon, and those now have an important influence across overall society, including economy, culture, and politics. Developments in information and communication networks have emphasized the problems of the existing public opinion survey system. Simultaneously, it demands research on new ways of conducting public opinion surveys. The proposed non-face-to-face preference survey analyzes facial expressions and changes in the behaviors of the research subjects, which are not directly expressed to the candidate, to determine the positive or negative emotions. In doing so, the positive or negative emotions toward the candidate can be determined and quantified. The support for the candidate can be checked in real time by either adding up the positive and negative emotions toward the candidate or by calculating the number of positive and negative emotions compared to the total number of research subjects. The biggest problem for the existing ARS public opinion survey system is that the data is collected with a response rate of less than 5%; hence, the reliability of the survey results can be doubted. The proposed method can resolve doubts regarding the reliability of the survey results.

The non-face-to-face, contactless preference survey system basically quantifies the emotional expressions and behaviors of citizens walking by at the campaign site. This system has the advantage of being capable of checking the responses of citizens in real time and avoiding the conventional telephone polling method. It is also possible to understand and collect changes in the support rate, which changes abruptly from six days prior to the election date, and the changes in the 5% support rate by

region and age, which changes daily. These changes can be monitored without having to conduct the existing ARS public opinion survey. In this study, the data structure was designed using the weighted reverse tree structure in the existing non-face-to-face, contactless preference survey system. In addition, published face recognition AI was used to collect the situations at the campaign site in more detail. The weighted reverse tree data structure was set up such that the user can assign the weight for each module to the existing reverse data structure. The system was designed so that a countervalue greater than one can be assigned for strong positives and strong negatives at the campaign site. As a result, more detailed public opinion analysis can be performed in the administrator mode. Furthermore, the system was designed to classify only the strong positives and strong negatives, separate from the neutral public opinion represented by weak positive and weak negative responses. The system was designed to collect data using the previously published facial recognition AI, so the ages, gender, and facial expressions of the responding citizens can be collected. The published facial recognition AI showed a recognition rate of over 99.5%, denoting a standard error of less than 5%. Hence, the results are reliable. The Naver Clova used herein supports all programming languages, such as C, Java, and Python. Hence, it can be highly utilized. In addition, existing tested local election application systems were utilized to apply face recognition algorithms. We then design and implement a mathematical analysis of the inverse tree structure. By applying the facial recognition algorithm and weighted reverse tree-type data structure, the non-face-to-face preference survey system showed superior scalability and ease of application compared to the existing preference survey system. The improved non-face-to-face, contactless preference survey system solves the problems of the election polls to a certain degree. Furthermore, it holds great significance in that it has provided a method for conducting non-face-to-face, contactless public opinion surveys in the post-coronavirus era. In this study, a published AI tool was used. However, the further development of the behavior and facial expression recognition by researchers is expected to broaden the direction of the non-face-to-face, contactless preference survey.

Data Availability

No data were used to support this study.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This work was supported by the Ministry of Education of the Republic of Korea and National Research Foundation of Korea (NRF-2019R1I1A3A01063132).

References

- [1] C. Kennedy, M. Blumenthal, S. Clement et al., "An evaluation of the 2016 election polls in the United States," *Public Opinion Quarterly*, vol. 82, no. 1, pp. 1–33, 2018.
- [2] W. Jennings and C. Wlezien, "Election polling errors across time and space," *Nature Human Behaviour*, vol. 2, no. 4, pp. 276–283, 2018.
- [3] A. M. Müller and S. Khoo, "Non-face-to-face physical activity interventions in older adults: a systematic review," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 11, p. 35, 2014.
- [4] K. Dang and S. Sharma, "Review and comparison of face detection algorithms," in *Proceedings of 2017 7th International Conference on Cloud Computing, Data Science & Engineering-Confluence*, pp. 629–633, Noida, India, January 2017.
- [5] T. Allison, A. Puce, D. D. Spencer, and G. McCarthy, "Electrophysiological studies of human face perception. I: potentials generated in occipitotemporal cortex by face and non-face stimuli," *Cerebral Cortex*, vol. 9, no. 5, pp. 415–430, 1999.
- [6] I. Masi, Y. Wu, T. Hassner, and P. Natarajan, "Deep face recognition: a survey," in *Proceedings of 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pp. 471–478, Parana, Brazil, November 2018.
- [7] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, and Le Song, "SphereFace: deep hypersphere embedding for face recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 212–220, Honolulu, HI, USA, July 2017.
- [8] F. Schroff, D. Kalenichenko, and P. James, FaceNet, A unified embedding for face recognition and clustering," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 815–823, Boston, MA, USA, June 2015.
- [9] T. Valentine, M. B. Lewis, and P. J. Hills, "Face-Space: a unifying concept in face recognition research," *Quarterly Journal of Experimental Psychology*, vol. 69, no. 10, pp. 1996–2019, 2016.
- [10] Y. Guo, L. Zhang, Y. Hu, X. He, and J. Gao, "MS-Celeb-1M: a dataset and benchmark for large-scale face recognition," in *Computer Vision—ECCV 2016*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds., Springer, Berlin, Germany, 2016.
- [11] J. Juchheim, P. Annighöfer, C. Ammer, K. Calders, P. Raunonen, and D. Seidel, "How management intensity and neighborhood composition affect the structure of beech (*Fagus sylvatica* L.) trees," *Trees*, vol. 31, no. 5, pp. 1723–1735, 2017.
- [12] T. Xue, J. Wu, Z. Zhang et al., "Freeman, seeing tree structure from vibration," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 748–764, Munich, Germany, September 2018.
- [13] Y. Zhang, X. Li, J. Wang, Y. Zhang, C. Xing, and X. Yuan, "An efficient framework for exact set similarity search using tree structure indexes," in *Proceedings of 2017 IEEE 33rd International Conference on Data Engineering (ICDE)*, pp. 759–770, San Diego, CA, USA, April 2017.
- [14] H. Nam, M. Baek, and B. Han, "Modeling and propagating cnns in a tree structure for visual tracking," arXiv preprint arXiv:1608.07242, 2016.
- [15] X. Lu, Y. Wang, X. Zhou, Z. Zhang, and Z. Ling, "Traffic sign recognition via multi-modal tree-structure embedded multi-task learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 4, pp. 960–972, 2017.