Neuro-Fuzzy Quantification of Personal Perceptions of Facial Images Based on a Limited Data Set

Luis Diago, Tetsuko Kitaoka, Ichiro Hagiwara, and Toshiki Kambayashi

Abstract—Artificial neural networks are nonlinear techniques which typically provide one of the most accurate predictive models perceiving faces in terms of the social impressions they make on people. However, they are often not suitable to be used in many practical application domains because of their lack of transparency and comprehensibility. This paper proposes a new neuro-fuzzy method to investigate the characteristics of the facial images perceived as *Iyashi* by one hundred and fourteen subjects. *Iyashi* is a Japanese word used to describe a peculiar phenomenon that is mentally soothing, but is yet to be clearly defined. In order to gain a clear insight into the reasoning made by the nonlinear prediction models such as holographic neural networks (HNN) in the classification of *Iyashi* expressions, the interpretability of the proposed fuzzy-quantized HNN (FQHNN) is improved by reducing the number of input parameters, creating membership functions and extracting fuzzy rules from the responses provided by the subjects about a limited dataset of 20 facial images. The experimental results show that the proposed FQHNN achieves 2–8% increase in the prediction accuracy compared with traditional neuro-fuzzy classifiers while it extracts 35 fuzzy rules explaining what characteristics a facial image should have in order to be classified as *Iyashi*-stimulus for 87 subjects.


I. INTRODUCTION

In an attempt to understand what motivates people to buy consumer products product manufacturers often rely on questionnaires and polls to investigate the consumers preferences and why they purchase or avoid certain items [1], [2]. The major problem with this approach is that, as shown by a variety of studies [1], consumer attitudes and behaviors are frequently formed and influenced by non-conscious emotional and cognitive processing [2], which is not subject to detection via traditional survey or interview methods. Here we assume that product-related stimuli (e.g., paintings or photographs) are shown as questionnaires to the consumers and at the same time a machine automatically recognizes and analyzes interactive signals while consumers evaluate what they felt about products in photographs. Then, using simple (e.g., “yes/no”) answers, a comprehensive analysis of those results could play an important role during analysis of consumer state of mind (cognitive and emotional reaction, both conscious and un-conscious) at the time of review, selection or use of a product or product-related stimuli.

In Japan, *Iyashi* 1 is a popular buzzword today, referring to anything that is physically or mentally soothing [3], [4]. *Iyashi* goods—books, music, pictures, incense, and aromas, bath salts, and plants—abound, offering to heal the physical and psychological stress of the workplace and of daily life in general [4]. Although there are instances whereby the English word *Healing* is used to explain the Japanese word *Iyashi*, it is not appropriate since there are healthy people who also consume *Iyashi* products [3]. In society the expression *Iyashikei* is frequently used to describe laypersons who simply help people to relax. In this paper, a new neuro-fuzzy method is proposed to investigate this peculiar phenomenon.

Psychological researchers use diverse methods to investigate emotions [5]–[7]. These procedures range from imagery inductions to film clips and static pictures. One of the most widely used stimulus sets is the International Affective Picture System (IAPS) [6], a set of static images based on a dimensional model of emotion [8]. The image set contains various pictures depicting mutilations, snakes, insects, attack scenes, accidents, contamination, illness, loss, pollution, puppies, babies, and landscape scenes, among others. However, while many samples are desirable for estimating the response of a person more accurately (e.g., how much the person likes a product), in a real world situation, only a small number of samples needs to be obtained because of the efforts required for the persons to provide their responses from many samples. Hence in this paper, we use a small dataset to teach the machine to classify the facial images in the same way that people perceive them.

Facial expressions are our primary means of communicating emotion [8]–[10], and that is why the majority of efforts in affective computing concern automatic analysis of facial displays [11]–[14]. Instead of analyzing facial expressions of an individual to determine their emotional state, using an artificially constructed database of faces, Brahnam and Nanni [15] showed that machines are also capable of perceiving faces as they relate to the social impressions they make on people. However, when photos of facial expressions (or avatars) are presented to a person the machine classification problem

1We use the word *Iyashi* as the stimulus given to a person which changes its internal condition to a better emotional state (see [3] for details).
becomes ill-posed [16], [17] because data dimensionality (i.e., number of features observed in facial expressions) becomes higher than the number of samples. Many classifiers such as artificial neural networks (ANN) [18] and support vector machines (SVM) [19] have been proposed to solve this ill-posed classification problem. Although both nonlinear techniques typically provide the most accurate predictive models in the analysis of facial expressions [11]–[14], [20], they are often not suitable to be used in many practical application domains because of their lack of transparency and comprehensibility. There are few studies that discuss the ways of making nonlinear prediction models more interpretable and transparent [21], [22]. This paper focuses on the machine analysis of personal perceptions, which is an application field where validation of the underlying models is required. A new method for the quantification of qualitative judgments and evaluations of facial expressions is proposed in order to gain a clear insight into the reasoning made by nonlinear prediction models, such as holographic neural networks (HNN) [20], [23]–[25].

First, the proposed approach introduces fuzzy quantification theory II [26] to select the most important facial parameters and to build a membership function (MF) describing the classifications made by a subject on a small data set of facial expressions. Later, the fuzzy reasoning is used to provide transparency to HNN models extracting linguistic rules and to answer what characteristics a facial image should have in order to be classified as Iyashi-stimulus.

The rest of this paper is organized as follows. Section II includes a brief review about emotion recognition technologies and affective rating of facial expressions. Previous neuro-fuzzy methods and HNN are also reviewed. Section III shows the proposed fuzzy-quantized holographic neural network (FQHNN). In Section IV, a limited data set of facial images is evaluated by hundreds of subjects and the proposed FQHNN is compared with neuro-fuzzy classifiers proposed in the literature [20], [27], [28] predicting subjects impressions of the faces. The meaning of Iyashi expressions according to the rules extracted from the computational models is also clarified. The results are discussed in Section V and the main conclusions and future works are presented in VI.

II. RELATED WORKS

A. Emotion Recognition from Facial Expressions

Automatic emotion recognition is inherently a multidisciplinary exercise involving different research fields, including psychology, linguistics, computer vision, speech analysis, and machine learning. Most of the vision-based affect recognition studies (see survey in [29]) focus on facial expression analysis [11]–[14] due to the importance of the face in emotion expression and perception.

Face recognition is a relatively mature sub-field of facial information processing, with several companies offering access control and security systems to end users [30]. There are mainly two types of approaches for facial feature extraction [11]–[14] geometric feature-based methods and appearance-based methods. The geometric methods present the shape and location of facial components (including mouth, eyes, brows, nose, etc.) while in appearance-based methods, image filters, such as Gabor wavelets, are applied to either the whole face or specific region in a face image to extract a feature vector. The choice of the appropriate approach is a question that remains in discussion among researchers.

If appearance-based methods are used with small-sample datasets then data dimensionality becomes higher than the number of samples and the classification problem becomes ill-posed. One way of solving ill-posed classification problems is to reduce the data dimensionality. However, if the data dimensionality is reduced, the models obtained using appearance-based methods will yield rules that are meaningless because the parameters describing the features of the face (e.g., Gabor wavelets [31], scale-invariant features and semantic factors [32]) are not interpretable. On the contrary, if we use geometric parameters, the rules are easier to interpret because the parameters themselves have a clear meaning (e.g., size of the eyes, mouth, the separation between the eyes and eyebrows). Since this paper focuses on the comprehensibility of nonlinear classifiers, facial expressions are represented by geometric facial features. Table I shows areas \( p_1, \ldots, p_7 \) and distances \( p_8, \ldots, p_{20} \) computed from feature points marked on the faces (See Fig. 1 and [3]). Note that in Table I, the parameters \( p_5, p_8, \) and \( p_9 \) are used for normalization and only 17 normalized parameters \( x_1, \ldots, x_{17} \) are used to represent the geometric facial features.

Facial expression classification is the last stage of a system to classify the facial features of the observed person in terms of categories about emotional states. Current feelings that emerge when looking at certain images are heavily influenced by contextual information. However, as was mentioned in the introduction, most of the research papers on this area used the IAPS [6] and the self-assessment Manikin [5] to measure the subject’s emotion response to the picture. Recently, after Psychology experts [7] classified a subset of 396 images from the IAPS into ten separate emotional categories, several researches [31], [32] used the emotional labels as the ground
truth to predict the emotions as perceived by humans while looking at pictures. There was a broad consensus among the subjects in [7] on the categorization of the emotions perceived, and as such these categories were used as labels for classification. If there were no clear definitions or consensus about certain categories of emotion perceived, these categories were not used as labels for classification, for example, pleasure, life satisfaction, positive emotions, etc. In this paper, instead of using the categories of emotion perceived such as anger, disgust, fear, sadness, we used linguistic terms [33] to investigate the characteristics of the facial images perceived as *Iyashi* by the subjects. Hence the selected facial images are called *Iyashi* expressions. Several applications can be developed based on the knowledge of the *Iyashi* expressions perceived by a subject. Mental commit robots [34] are an example of an *Iyashi*-based medical applications to provide psychological, physiological, and social effects in human beings through physical interaction.

### B. Neuro-Fuzzy Methods

In the last decade, various neuro-fuzzy systems have been developed ([35], [36] and reference therein). They combine the natural language description of fuzzy systems and the learning properties of neural-networks. Some of them are known by their abbreviations like ANFIS [37], ANFC [28], and NEFCLASS [27]. In general, the research field is divided into: linguistic fuzzy modeling that is focused on interpretability, mainly using the Mamdani model [38], and precise fuzzy modeling that is focused on accuracy, mainly using the Takagi-Sugeno-Kang model [39], [40]. The literature contains a large amount of work using fuzzy neural networks to find nonlinear relationships between input-output data and most of them are based on a connectionist approach using the back-propagation (BP) learning algorithm. In this paper, instead of using traditional neural networks modeling [27], [28], [37], we focused on HNN models [20], [23], [24].

The holographic method, developed by Sutherland [23], proceeds considerably beyond the neural paradigms. It could also be used to implement the above paradigms as well as to serve for fuzzy processing. HNN are based on the fact that the input-output relationships are linearized according to a mapping of inputs into outputs at the complex plane. HNN are also a kind of complex-valued neural networks (CVNN) that have gained popularity in recent years [25]. However, unlike CVNN the governing equations of HNN formed a non-connectionist approach in which a large number of input-output associations could be enfolded onto a single memory element. In 1998, Khan [24] developed the concept of learning with dynamic attention and suggested the use of multidimensional complex numbers but leaves open the way to handle these. Recently, in our previous works [20], we proposed the idea of using fuzzy quantification to expand the networks to higher dimensions without the need to empirically adjust the terms for higher-order statistics [23] to increase the accuracy of HNN. This paper focuses on the interpretability instead of the accuracy of HNN models.

In recent years, research in the neuro-fuzzy community began focusing on the trade-off between interpretability and accuracy [21], [22], [41]. There is no well-established definition for interpretability of a fuzzy system hence the researches are based on the taxonomy of interpretability of fuzzy systems introduced in [42]. Since the proposed method is based on neural networks, we focus only on the symbolic interpretation of trained ANN which mainly focuses on the extraction of fuzzy rules [43]–[45]. After Andrews *et al.* [43], [44] proposed a taxonomy for categorizing the numerous contributions in this area, layered neural networks are used for automated extraction of fuzzy IF-THEN rules stored in the nodes of the network or fuzzyfying the whole learning process in which the nodes of the network handle fuzzy weights. The weights are updated within a fuzzy learning process [45]. However, for HNN the situation is completely different. The HNN have no nodes nor weights and therefore algorithms developed so far cannot be used directly by HNN. Andrews classification scheme has been extended to extract rules from SVM [46] but the authors state that SVM rule extraction is still in its infancy, certainly compared to ANN rule extraction. This paper is a continuation of our previous works [3], [20], [47] in which we introduced fuzzy quantification to build an MF describing how subjects classify *Iyashi*-stimulus into fuzzy groups. In the following, the idea of FQ is extended and a new fuzzy-quantized HNN (FQHNN) is proposed to gain a clear idea of what it means for a face to be *Iyashi* according to the computational models.

### III. FUZZY-QUANTIZED HNN

Fig. 2 shows the architecture of the proposed FQHNN. A novel FQ module (denoted as FQ) is added to increase the interpretability of HNN models. In the following, the proposed architecture is explained in more details.

#### A. HNN

Assume two real vectors $x$ and $y$ represented as input vector $x = [x_1, x_2, \ldots, x_k]^T$ and output vector $y = \ldots$
{y_1, y_2, \ldots, y_m}^T$. In this paper, the dimensions of $x$ and $y$ are $k = 17$ and $m = 1$, respectively. If we have a set of $n = 20$ input-output vectors, they can be represented as matrix $X$ and $Y$ according to (1) and (2), respectively:

$$X = \begin{pmatrix}
    x_1^T & \cdots & x_n^T
  \end{pmatrix}
  = \begin{pmatrix}
    x_{11} & x_{12} & \cdots & x_{1k} \\
    x_{21} & x_{22} & \cdots & x_{2k} \\
    \vdots & \vdots & \ddots & \vdots \\
    x_{n1} & x_{n2} & \cdots & x_{nk}
  \end{pmatrix}$$

(1)

$$Y = \begin{pmatrix}
    y_1^T & \cdots & y_m^T
  \end{pmatrix}
  = \begin{pmatrix}
    y_{11} & y_{12} & \cdots & y_{1m} \\
    y_{21} & y_{22} & \cdots & y_{2m} \\
    \vdots & \vdots & \ddots & \vdots \\
    y_{n1} & y_{n2} & \cdots & y_{nm}
  \end{pmatrix}.$$ 

(2)

Each element of $X$ and $Y$ is changed into angles $\theta_{ai} (a = 1, \ldots, n; i = 1, \ldots, k)$ and $\phi_{aj} (a = 1, \ldots, n; j = 1, \ldots, m)$ by transformation/mapping functions $f_x$ and $f_y$:

$$\theta_{ai} = f_x(x_{ai}),$$

(3)

$$\phi_{aj} = f_y(y_{aj}).$$

(4)

Some commonly used types of data conversion in $f_x$ and $f_y$ are linear conversion (real values from a known finite range) and sigmoid or arctan conversion (real values from an unknown or infinite range). Next, angles are mapped on the complex plane by the exponential function:

$$s_{ai} = \lambda_{ai} e^{i\phi_{ah}},$$

(5)

$$r_{aj} = \gamma_{aj} e^{i\phi_{aj}}.$$ 

(6)

In (5) and (6), $i$ is the imaginary unit of a complex number ($i^2 = -1$), $\lambda_{ai}$ and $\gamma_{aj}$ are the magnitudes of the complex numbers used by Khan [24] to learn with dynamic attention. Through the operations represented in (3)–(6), arrays $X$ and $Y$ are mapped onto the complex plane and named stimulus $S$ and response $R$, respectively:

$$S = \begin{pmatrix}
    s_{11} & s_{12} & \cdots & s_{1k} \\
    s_{21} & s_{22} & \cdots & s_{2k} \\
    \vdots & \vdots & \ddots & \vdots \\
    s_{n1} & s_{n2} & \cdots & s_{nk}
  \end{pmatrix},$$

$$R = \begin{pmatrix}
    r_{11} & r_{12} & \cdots & r_{1m} \\
    r_{21} & r_{22} & \cdots & r_{2m} \\
    \vdots & \vdots & \ddots & \vdots \\
    r_{n1} & r_{n2} & \cdots & r_{nm}
  \end{pmatrix}.$$ 

(7)

If the relationship between $S$ and $R = [r_1, \ldots, r_m]$ is stored in a matrix $H = [h_1, \ldots, h_n]$ called holographic memory, then the difference between $R$ and the product $S \cdot H$ is minimized by computing $H$ as follows:

$$H = (S^* \cdot S)^{-1} \cdot S^* \cdot R.$$ 

(8)

Here, the symbol $*$ denotes the conjugate transpose of the matrix $S$. The matrix $H$ is determined using input $S$ and output $R$ during learning (training) and the output $V$ for a new input $U$ unused during the training is predicted by $V = U \cdot H$. To get the real vectors $v$ for the new output $V$, it is necessary to reverse the mapping in (4).

Note that the original architecture proposed by Sutherland [23] includes an additional expansion of the input field to higher order combinatorial product terms computed by $\prod_{i=1}^{K} \lambda_i e^{i\phi}$. In our previous works [20], the optional expansion step has been substituted by a FQ step (named “FQ expansion” in Fig. 2) to avoid the trial and error approach defining the number of combinations of unique K-th order product terms that can be generated from a data field of size $k$ and to increase the generalization accuracy of HNN classifiers. In this paper, we also use the external domain of each parameter $x_i$ ($i = 1, \ldots, k$) as real numbers in the range $-\infty < x_i < +\infty$ with an arbitrary distribution $f_i(x)$. We divide the external domain of the variable $x_i$ into $l_i$ categories $C_{il}$ ($l = 1, \ldots, l_i$) and compute new distribution functions for each category by:

$$\int_{-\infty}^{\infty} f_i(x) dx = \int_{\tau_1}^{\tau_2} f_i(x) dx = \cdots = \int_{\tau_{l-1}}^{\infty} f_i(x) dx.$$ 

(9)

After the correct selection of the boundaries $\tau_1, \tau_2, \ldots, \tau_{l_i-1}$, the new subsets $C_{il}$ are obtained from the original sequence $x_i$ according to the following syntactic rules [33]:

$$\begin{align*}
  \text{if } & -\infty < x_i \leq \tau_1 \text{ then } x_i \in C_{i1} \\
  \text{if } & \tau_1 < x_i \leq \tau_2 \text{ then } x_i \in C_{i2} \\
  & \quad \vdots \\
  \text{if } & \tau_{l_i-1} < x_i < +\infty \text{ then } x_i \in C_{il_i}.
\end{align*}$$ 

(10)

The generalization accuracy is increased using above preprocessing method [20]. However, the non-transparency of the resulting models is perhaps the greatest deficiency of HNN from an engineering perspective. When one tries to interpret a resulting HNN model, the only information one may withdraw is the saliency of the input variables i.e., a measure of their contribution to the final output. Although HNN models serve for fuzzy processing they are not necessarily comprehensible, especially because the learning algorithm aims at reaching a maximum accuracy. As the results in [20] are sub-optimal, this paper improves the predictions by selecting subsets of input parameters [48] and focuses on increasing the interpretability of the models by FQ [26].

**B. FQ**

The objective of FQ theory [26] as it is used in this paper is to improve the interpretability of HNN models expressing several fuzzy groups for the word *Iyashi* in terms of face parameters that take the form of values on $[0, 1]$. We try
to express, as best as is possible, using a linear equation of category weight $c_{il}$ of category $C_{il}$, the structure of the external standard fuzzy groups for the word Iyashi on the real number axis

$$y(x_a) = \sum_{i=1}^{K} \sum_{l=1}^{l_i} c_{il} \mu_{C_{il}}(x_a), \quad a = 1, \ldots, n.$$  \hspace{1cm} (11)

So it is determining $c_{il}$ that gives the best separation of the external standard fuzzy groups on the real number axis $y_a = y(x_a)$. The degree of separation of the fuzzy groups is defined as the fuzzy variance ratio, $\eta^2$, which is the ratio of the total variation and variation between fuzzy groups [26]. It is possible to find category weight $c = [c_{11}, c_{12}, \ldots, c_{1k_1}]$ in (11), from eigenvector $c$ which maximizes fuzzy variance ratio $\eta^2$ solving the generalized eigenvalue problem [26], [47]

$$A c = \eta^2 B c.$$  \hspace{1cm} (12)

The values of $\eta^2$ satisfying (12) are the generalized eigenvalues and the corresponding values of $c$ are the generalized right eigenvectors.

In (12), $A$ and $B$ (nearly) share a null-space, which means that the eigenvalues and eigenvectors are not uniquely defined (or extremely ill-conditioned) [47]. In [47], we have proposed a perturbation of the eigenvalue problem by using the Moore–Penrose pseudoinverse that keeps the error bounds of the solution within the error bounds of the widely used for “Generalized Upper TRIangular Form” (GUPTRI) algorithm [49]. The solution obtained by the proposed method [47] is more stable and accurate than the GUPTRI algorithm [49] but the final MFs for the output groups of one subject are overlapped because the number of categories for the input parameters is not optimal and the input-MFs are not tuned. In this paper, we use a simpler solution based on the approach proposed by Okuhara et al. [50] in a cultivation environment selection (see details in the Appendix) to solve above-mentioned problems.

C. Improving HNN by FQ

Modifying the optimization algorithm proposed in [50], the HNN interpretability is improved from three perspectives: reducing the number of input parameters, creating output MFs, and extracting fuzzy rules.

1) Reducing the Number of Input Parameters: Feature selection algorithms used to reduce data dimensionality may be divided into filters, wrappers, and embedded approaches [48]. Filters select subsets of input as a preprocessing step, independently of the chosen predictor. Wrappers utilize the learning machine of interest as a black box to score subsets of input according to their predictive power. Embedded methods perform input selection in the process of training and are usually specific to given learning machines. Based on the proposed FQHNN, two filters called $hnn$-filter and $fq$-filter are developed. The $hnn$-filter ranks the variables using the level of input-output correlation measured directly from the magnitude $\rho_i$ of the complex value $h_i = \rho_i e^{i\theta_i}$ stored in the holographic memory matrix $H$. The $fq$-filter sorts the variables based on the category weight $c = [c_1, \ldots, c_K]$ in (11). In both cases, after the “FQ expansion” step (see Section III-A) the final number of input fuzzy groups will be $K = \sum_{i=1}^{k} l_i$. The top $K_5$ features in the descending order of $\rho_i$ (or $|c_i|$) are selected as the new $K_5$ features ($1 \leq K_5 \leq k$).

In order to find the optimal number of categories ($l_i$) for each parameter, the generalized eigenvalue problem in (12) is solved by tuning the MFs $\mu_{il}(x_i) \subset N(b_i, \sigma_i)$ and $\mu_{G_i}$ for input and output fuzzy groups, respectively. The optimization algorithm proposed in [50] is modified to deal with some drawbacks when the multiplicity of the biggest eigenvalue is bigger than one and to avoid getting trapped in a local optimum. Whereas $c_{il}, \sigma_i$ and $b_i$ are randomly initialized in step 2) of optimized-FQ [50] (see Appendix), we only use random initialization for $c_{il}$ and $\sigma_i$ while $b_i$’s are initialized from the mean of the data in each category $C_{il}(l \neq 1, l \neq l_i)$ and minimum and maximum value for category $C_{i1}$ and $C_{il_i}$, respectively. Also, in order to deal with the greatest eigenvalue with multiplicity bigger than one, step 4) is modified to minimize the relative normwise backward error using a few iterations with the Lanczos conjugate-gradient method as in [51]. The relative normwise backward error for a computed finite eigenvalue, $\eta^2$, and its corresponding eigenvector, $c$, in (12) is evaluated by the expression [52]: $\Gamma(\hat{c}, \hat{\eta}^2) = (||\hat{\eta}^2BC - \hat{A}||_2/(||\hat{\eta}^2||_2||B||_2 + ||A||_2)||\hat{c}||_2)$.

Note that tuning the MFs to solve the eigenvalue problem in (12) does not impose any semantic constraint for MF-optimization [53] but only maximizes the separation between the output fuzzy groups. The main goal of the optimized-FQ is to select the optimal number of categories ($l_i$) for FQHNN initialization.

2) Creating Output MFs: The simple representation of the information contained in the classifier by output MFs adds a lot to its interpretability [54]. To create MFs for FQHNN outputs, the magnitudes $\rho_i$ of the complex values $h_i$ stored in the holographic memory matrix $H$ are rescaled and used as weights $c_i$ for the linear equation in (11). Since the magnitudes of complex values in $H$ are only positive, the phases $\varphi_i$ of the complex numbers are used to determine the sign of each coefficient $c_i$. Note that the FQ in (9) and (10) uses crisp input functions and the distribution function of the input parameters is considered by the mapping function.

The only variable part of a solution based on HNN is preprocessing of input/output data in (3)–(6) by mapping functions $f_x$ and $f_y$. Sutherland [23] suggested using a sigmoid

$$\theta_i = \frac{2\pi}{1 + e^{-f(x_i, b_i, \sigma_i)}}$$  \hspace{1cm} (13)

to obtain an ideal symmetrical distribution of the input data that has a normal distribution. The symmetry achieved by the mapping functions is measured as a summation over the set of stimulus vectors [23]. Generally $f(x_i, b_i, \sigma_i) = A|x_i - b_i/\sigma_i|^2 : A, z \in \mathbb{R}$. Here, $x_i$ is the $i^{th}$ input stimulus element and $b_i$ and $\sigma_i$ are the mean and standard deviation of stimulus field distribution. Since the query pattern and the encoded patterns must be mapped using the same function, the response pattern may not be retrieved correctly if the query pattern significantly differs from the encoded pattern, e.g., they have a different mean and standard deviation. For this reason,
Khan [24] uses the following spiral mapping function:
\[ \theta_i = \left( k_{\text{spread}} \frac{x_i 2\pi}{(x_i)_{\text{max}} - (x_i)_{\text{min}}} \right) \mod 2\pi - \pi \] (14)
for achieving a symmetrical distribution of stimulus pattern elements. This function transforms the input data into \(-\pi\) to \(+\pi\) range. Here, \(k_{\text{spread}}\) is the spreading coefficient, which is chosen to be any number which is not a factor of \((x_i)_{\text{max}} - (x_i)_{\text{min}}\), otherwise the mapped values will not be unique. The parameters \(A = 3\), \(z = 1\), and \(k_{\text{spread}} = 7\) are used in the experimental section because they achieved the highest symmetry in (13) and (14), respectively.

3) Extracting Fuzzy Rules: After selecting the most important features and creating input-output MFs, the “black-box” (also called “pedagogical”) technique [43]–[45] in which the extraction of rules is done only through the exploration of the relationships between inputs and outputs is applied for the first time in the context of HNN. In this paper, the method proposed by Wang and Mendel [55] is used to generate an initial set of rules. The Wang–Mendel method [55] generates one rule for each input-output data pair. A small set of rules will be generated because the dataset used in this research is very small. The number of rules can be reduced even more using the importance degree of each rule [55] or inducing cooperation among rules [56]. Note that the MFs of input groups in the proposed approach are crisp and the importance degree of each rule will be determined by the MFs of output groups. In the next section, the classification accuracy of the proposed method is compared with ANFC [28], NEFCLASS [27], and previous HNN [20]. ANFC [28] is an adaptive neuro-fuzzy classifier based on ANFIS [37] in which the rule weights are adapted by the number of rule samples. The implementation of ANFC [28] uses the k-means algorithm to initialize the fuzzy rules. For that reason, the user should give the number of cluster for each class. NEFCLASS [27] is also a neuro-fuzzy classification model in which the fuzzy rules describing the data are of the form: \(R^i: \text{if } x_1 \text{ is } A^i_1 \text{ and } x_2 \text{ is } A^i_2 \text{ and } \ldots \text{ and } x_n \text{ is } A^i_n \text{ then } \ldots \text{ pattern } (x_1, x_2, \ldots, x_n) \text{ belongs to class C where } A^i_1, \ldots, A^i_n \text{ are fuzzy sets.}
To create a rule base for a NEFCLASS system using small datasets, Nauck and Kruse [27] recommend choosing “best” rule learning within NEFCLASS [27] because there are classes which have to be represented by a larger number of rules than other classes.

IV. Simulation Results

The main goal of the simulations is to obtain a reliable representation of the MFs describing subject’s evaluations of facial images and to gain an insight into the reasoning made by HNN classifiers predicting the degree of Iyashi perceived by the subjects in a set of facial images. The extraction of fuzzy rules by the above neuro-fuzzy classifiers is used to clarify the meaning of Iyashi expressions.

A. Psychological Experiments

One hundred and fourteen people between 15 and 70 years old (102 Japanese and 12 non-Japanese, 47 females, and 67 males) participated in the experiment. We are aware that the subjects’ answers are completely subjective, however, we decided to select the following method for engineering simplicity. The participants rated each stimulus on the scale ‘0’- NO, ‘1’- DON’T KNOW, ‘2’- YES to express whether or not they perceive an Iyashi-stimulus from the faces in Fig. 1. Note that the three groups used in the classification, give freedom to the subjects to use the second group (DON’T KNOW) for the case in which the question does not match with their perception of Iyashi-stimulus. On average, each experiment took less than 5 minutes. The time for evaluating each stimulus image was set to a maximum of 10 seconds. In each trial, the stimulus was selected at random from the set of 20 images and presented for five seconds. The total time for evaluation of each image was also recorded. If the time exceeds 10 seconds, the image is returned to the set of images in order to be presented one more time. The images evaluated within the maximum time interval of 10 seconds are removed from the evaluation set.

Fig. 3 shows the frequency distribution of Iyashi degree for each face number according to 114 subjects. The black, gray, and white bars correspond to evaluation 0 (No), 1 (Don’t Know) and 2 (Yes), respectively.
$l_i = 3$ categories (small, medium, and large) denoted as $C_{l_i} = S$, $C_{l_i} = M$, and $C_{l_i} = L$. The MFs of 114 subjects are constructed from $K = 51$ fuzzy groups computed from $x_i$ and $M = 2$ fuzzy groups associated with the evaluations in Fig. 3. $G_r$ values were defined as follows: images evaluated as 0 were assigned $G_1 = 0$ and $G_2 = 1$, images evaluated as 2, $G_1 = 1$ and $G_2 = 0$ and for those assessed as 1, $G_1$ and $G_2$ were 0.5.

1) Optimized-FQ: First, using the evaluations provided by the 114 subjects, we demonstrate the importance of the proposed optimization comparing the distribution of the biggest eigenvalue $\eta^2$ and the distribution of the 20 images on the real axis solving (12) with GUPTRI and optimized-FQ. The default parameters ($\epsilon = 10^{-8}$, gap = 1000) are used in GUPTRI and $a = 0.001$ and $\delta = 10^{-8}$ are used in optimized-FQ. In comparison, the data in Fig. 4 show only $N = 98$ of the 114 subjects which included images in the three groups. Fig. 4(a) and (c) shows the results of GUPTRI without parameter tuning of MFs and Fig. 4(b) and (d) shows the results of optimized-FQ.

In Fig. 4(a), the values of $\eta^2$-GUPTRI are very small ($\text{mean} = 0.02$, $\text{std} = 0.04$) in comparison with the values of $\eta^2$-optimized-FQ in Fig. 4(b) ($\text{mean} = 0.36$, $\text{std} = 0.05$). Fig. 4 also shows an example of how the 20 sample images are distributed across the $y_a$-axis (e.g., subject 31 from the set of 98 subjects). In Fig. 4(c), without tuning of MFs (i.e., using GUPTRI algorithm only), images evaluated as 1, are grouped around $y_a = 0$, and the images in the other two groups are scattered on both sides of the axis $y_a = 0$. For none of the subjects, the class distribution is balanced. For example, subject 31 evaluated 45% of the 20 sample images as 2, 45% as 0 and 10% as 1. After tuning the input MFs, all images evaluated as 0 moved to the right of the axis $y_a = 0$ and all images evaluated as 2 moved to the left side. The images evaluated as 1 remained on the axis $y_a = 0$. The result in Fig. 4(b) and (d) shows that tuning the input MFs increases the degree of separation between the output fuzzy groups ($\eta^2$) which could facilitate the predictions made by nonlinear classifiers.

\begin{center}
\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig4.png}
\caption{(a) $\eta^2$-GUPTRI, (b) $\eta^2$-optimized-FQ, (c) Subject-31($\eta^2 = 0.0091$), and (d) Subject-31($\eta^2 = 0.4319$). Distribution of biggest eigenvalue $\eta^2$ for $n = 98$ subjects (upper part) and example of how the 20 sample images are distributed across the $y_a$-axis (lower part) solving the generalized eigenvalue problem with GUPTRI (left) and optimized-FQ (right).}
\end{figure}
\end{center}

\begin{center}
\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig5.png}
\caption{MFs obtained by optimized-FQ from the evaluations given by subject 100. Note that no interpretability constraints are imposed on the MFs because they are only used to quantify the proposed FQHNN (see details in the text).}
\end{figure}
\end{center}

2) Non-Random Initialization: The optimized-FQ has the severe drawback of requiring many iterations to maximize the fuzzy variance ratio $\eta^2$ which has long, narrow valley structures when using the matrices obtained from the limited dataset. To see how the proposed initialization influences the behavior of the algorithm, we run the optimized-FQ with random initialization [50] by using the proposed initialization procedure explained in Section III-C.1. Except for the subject 77, there is no significant difference in the values of $\eta^2$ for both methods. In the case of subject 77, the optimized-FQ is trapped in a local optimum using a random initialization which does not happen if the proposed initialization algorithm is used. The number of iterations needed by optimized-FQ is considerably reduced using the proposed initialization. The average number of iterations was 19 317 for random initialization and 3887 for the proposed procedure. The main conclusion coming from the above simulations is that the proposed initialization can accelerate the solution of the generalized eigenvalue problem.

C. FQHNN Interpretability

1) Selecting the Most Important Face Parameters: The importance of facial parameters is automatically determined by FQHNN from the coefficients $c_j$ in (11). The parameters $b_i$ and $\sigma_i$ of the MFs describing the fuzzy sets of facial parameters are also automatically adjusted by optimized-FQ. However, the optimization imposes no interpretability constraints [41], [42], [53] on the MFs of input fuzzy groups because we are only interested in using FQ to define the number of input groups to quantify the proposed FQHNN.

Fig. 5 shows an example of MFs obtained by optimized-FQ from the evaluations given by subject 100. Although three categories ($l_i = 3$) were initially used to define the fuzzy sets for each facial parameter, after tuning the MFs, some parameters are described only by one or two groups because the $\sigma_i$ values are very small (for example $x_5$, $x_7$, $x_8$, and $x_{10}$). In Fig. 5, only very narrow lines are observed for such parameters. The number of remaining sets (not narrow sets) is used as the optimal number of categories for HNN quantification. Note that the new values of $l_i$ are used by the HNN for FQ expansion without either tuning the parameters.
of the input mapping functions or tuning the boundaries of the FQ in (9)–(10).

2) Visualization of Output MFs: As designing a good HNN in fact means choosing an appropriate preprocessing method, we compared the sigmoid function with the spiral function proposed by Khan [24] in order to find a transformation function that enables HNN MFs for output sets to be similar to the functions obtained by optimized-FQ. Fig. 6 shows the comparison for the same subject 100 as discussed in the previous section. The upper part shows the MF obtained by optimized-FQ. The MFs for HNN constructed from the sigmoid function are shown in the middle and the bottom shows the functions obtained for the HNN with the spiral function. Although in none of the cases the functions are identical to the functions obtained by optimized-FQ, the sigmoid function offers better results. As we saw in the previous sections, the MFs according to the groups “0” and “2,” are distributed on both sides of the MF of the group “1” for optimized-FQ and the spiral function. On the contrary, by using the sigmoid function the samples of the groups “0” and “2” are overlapped and the samples in group “1” do not remain close to the axis $y_a = 0$. This is because the sigmoid function assumes that the parameters follow a Gaussian distribution, which is not correct for the images in the dataset. So, as proposed by Khan [24], the spiral function offers better results when the distribution of parameters is not Gaussian.

D. Classification Accuracy and Rule Extraction

In order to give a basis for comparison, a leave-one-out cross-validation (LOOCV) method was tested for classification and rules extraction with four neuro-fuzzy classifiers: ANFC [28], NEFCLASS [27], HNN [20], and the proposed FQHNN. In the LOOCV method, an image from the dataset is eliminated and the classifier is trained to reproduce the associations with the remaining images. After training, the classifier is tested for the prediction of the removed image. The process is repeated for all the images in the dataset.

First, we analyze the influence of the mapping functions in the HNN classifier. In the previous section, we concluded that using the spiral function for mapping input parameters provides better MFs of output fuzzy groups than using the sigmoid function. Table II shows the classification results for the same subject as discussed in the previous section. Row 1 shows the results of the HNN using the sigmoid function (sigmoid – HNN) and row 2 shows the result of HNN using the spiral function (spiral – HNN). We explored the seventeen parameters in Table I using the hnn-filter (see Section III-C.1). Column 2 shows the number of selected features ($K_s$) that provided the highest classification accuracy for the removed image. The final column shows the total percent of correct retrieval (PCR) that measures the performance of the HNN classifiers. PCR is calculated as $PCR = \frac{\text{number of correct retrievals}}{n} \cdot 100 \%$ where $n$ is the number of retrieved patterns and “correct” means that the prediction made by HNN equals the evaluation of the subject. Columns 3–5 show the PCR for each class. As shown in Table II, the generalization accuracy using the spiral function is 5% higher than using sigmoid. The number of selected features using spiral is also smaller than using sigmoid. Note that using sigmoid, the class 0 could never be classified correctly which coincides with the results in Fig. 6. This demonstrates the advantages of using spiral over sigmoid for predicting the evaluations of subject 100. The last two rows show the results of the proposed FQHNN. In both cases, the spiral mapping function was applied for input mapping function $f_x$. The hnn-filter was also applied in the method shown in row 3 and the fq-filter was applied in row 4 (fq-FQHNN). For the FQHNN the input to the network is fuzzyfied using the proposed quantification algorithm with $l = 3$ categories (small, medium, and large) and the holographic memory is computed using (8). Table II shows that the generalization accuracy of the proposed FQHNN is 5–10% higher than the generalization accuracy of previous HNN classifiers [20]. The highest PCR = 60 is achieved by using fq-FQHNN but $K_s = 14$ is still very high for understanding the HNN model compared with hnn-FQHNN ($K_s = 3$).

Finally, we compared the above results with ANFC [28] and NEFCLASS [27] for $N = 87$ subjects which included more than two images in each group. This restriction is imposed by ANFC [28] which uses 2 rules for each class and needs at least 1 image in each class to create the classifier. In the NEFCLASS model, the number of rules was determined automatically by training fuzzy sets with three bell-shaped MFs which must overlap and keep the order. The learning rate was 0.01 and the training proceeded for at most 300 cycles. Table III shows the PCR of four neuro-fuzzy classifiers. The first and the sixth column shows the ID of each subject. ANFC [28], HNN [20], and NEFCLASS [27] are represented.

---

**Table II**

<table>
<thead>
<tr>
<th>Classifier</th>
<th>$K_s$</th>
<th>Class 0</th>
<th>Class 1</th>
<th>Class 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>sigmoid-HNN</td>
<td>15</td>
<td>0</td>
<td>67</td>
<td>50</td>
<td>45</td>
</tr>
<tr>
<td>spiral-HNN</td>
<td>12</td>
<td>20</td>
<td>89</td>
<td>17</td>
<td>50</td>
</tr>
<tr>
<td>hnn-FQHNN</td>
<td>3</td>
<td>60</td>
<td>78</td>
<td>17</td>
<td>55</td>
</tr>
<tr>
<td>fq-FQHNN</td>
<td>14</td>
<td>40</td>
<td>100</td>
<td>17</td>
<td>60</td>
</tr>
</tbody>
</table>

Fig. 6. MFs describing output fuzzy sets from HNN trained with evaluation of Subject 100.
in columns 2–4 and 7–9 by letters A, H, and N, respectively. The results with the proposed FQHNN are shown in columns 5 and 10. The fq-filter was applied for FQHNN and the spiral function is used in both holographic classifiers. The highest PCR obtained for each subject is highlighted in bold. The last function is used in both holographic classifiers. The highest FQHNN rule was created for each subject after pruning and trimming the FQHNN rule base using the NEFCLASS model. Table IV shows a set of 35 rules extracted from the computational models representing the evaluations of each subject. The rule base is divided into two sets: the rule set A which includes 21 rules with a single facial parameter in the antecedent and the rule set B which includes 14 rules with all parameters in the antecedent. PCR for the N = 41 subjects generating the rule set A was higher than PCR for the N = 46 subjects generating the rule set B (46.95 ± 9.6 versus 39.77 ± 9.35). In Table IV, columns 2 and 6 (IF-part) show the antecedent of each rule and columns 3 and 7 (THEN-part) show the consequent. Columns 4 and 8 show the number of times the rule is repeated according to N = 87 subjects which could indicate the degree of each rule (D). The rule set A is easier to interpret than rule set B. In the rule set A, only eleven parameters (x1, x2, x3, x5, x6, x8, x10, x11, x13, x14, x15, x17) are used to explain the meaning of subjects evaluations. The most repeated parameters are x5 (11 times), x6 (8 times), x13 (7 times), and x11 (4 times). According to the HNN models, the classification made by the subjects is based primarily on the upper face (i.e., x3,x5-the size of the eyebrows, x3-the distance from the right eye to the eyebrow and x11-the size of the forehead). In Table IV, there are conflicting rules explaining that the rules are subject specific and cannot be mixed if there is no consensus. For example, the three rules A-3, A-12, and A-16 use x5 in the antecedent and they are conflicting. The same applies to rules A-4 and A-17 which use x6 in the antecedent. In the rule set B, the most repeated rules in each class are also in conflict (e.g., B-1, B-5, and B-11). We are currently studying the relationship between the obtained rule base and the age, sex, and nationality of the subjects.

V. DISCUSSION

A first point for discussion is that the naturally distinguishable classes in this paper are defined by three terms (range 0–2) used in the context of the word *Iyashi*. A more flexible vocabulary of 16 candidate words or phrases-terms (none, very little, a small amount, a little bit, a bit, some, a moderate amount, a fair amount, a good amount, a considerable amount,
TABLE IV
RULES EXTRACTED AFTER PRUNING AND TRIMMING THE FQHNN RULE
BASE CREATED FROM THE EVALUATIONS OF N = 87 SUBJECTS

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(N = 41 subjects using a single parameter in the antecedent)</td>
<td>(N = 46 subjects using all parameters in the antecedent)</td>
</tr>
<tr>
<td><strong>No.</strong></td>
<td><strong>IF</strong></td>
</tr>
<tr>
<td>1</td>
<td>x1 is S</td>
</tr>
<tr>
<td>2</td>
<td>x3 is S</td>
</tr>
<tr>
<td>3</td>
<td>x5 is M</td>
</tr>
<tr>
<td>4</td>
<td>x6 is M</td>
</tr>
<tr>
<td>5</td>
<td>x10 is M</td>
</tr>
<tr>
<td>6</td>
<td>x11 is S</td>
</tr>
<tr>
<td>7</td>
<td>x11 is M</td>
</tr>
<tr>
<td>8</td>
<td>x13 is M</td>
</tr>
<tr>
<td>9</td>
<td>x14 is S</td>
</tr>
<tr>
<td>10</td>
<td>x17 is S</td>
</tr>
<tr>
<td>11</td>
<td>x17 is M</td>
</tr>
</tbody>
</table>

| **No.** | **IF** | **THEN** | **D** |

a sizeable amount, a large amount, a substantial amount, a lot, an extreme amount, and a maximum amount) has been used in the literature to cover the interval 0–10 in a context-free situation [57]. However, an open issue is whether or not such context-independent results can be applied when the terms are used within a specific context. Also other common impressions (such as anger, disgust, sadness, awe, etc) from common images in the IAPS [6] (such as flowers, snakes, etc) have been predicted in the literature [31], [32]. The highest PCR was about 74% using an SVM classifier with 20 semantic factors and a dictionary of 500 emotional words. In our experiments, each subject represented one ground truth and the highest PCR obtained by FQHNN in Table III was 75%. Despite a different context, the PCR for FQHNN is 1% higher than the PCR for SVM in [31] and [32] and the proposed FQHNN is also more interpretable mainly because the features used in [31] and [32] are meaningless.

For the neuro-fuzzy community, the greatest impact of this paper lies in the fact that the word *Iyashi* (like other such as pleasure, well-being or satisfaction) does not have a consensus among people. In such situation, this paper showed that it is worth considering the FQHNN for learning and knowledge mining on small datasets. Selecting the most important parameters from the FQHNN avoids having to solve the generalized eigenvalue problem and it allows obtaining output MFs. As discussed in the experiments, the spiral mapping function provides reliable output MFs compared with *sigmoid*, which is the most widely used function in the context of HNN. However, note that we have not tuned the mapping function in the HNN model. Tuning within the HNN model, while ensuring the maximum separation of the classes is still an open issue.

The proposed method also provides the HNN with explanatory properties. Not just because it provides a simple representation of the information contained in the classifier by output MFs, but also because it can extract rules in a simple way. The rules can be further reduced through any of the existing techniques.

VI. CONCLUSION

In this paper, we have presented a new method that uses qualitative discrimination analysis to describe subjective levels of induced emotions in a hundred and fourteen subjects. The proposed approach can be used with any type of stimulus and the number of interpretation categories can also be changed within the context of other emotions. The proposed approach finds the parameters of the MFs describing the personal perception of faces by an optimization approach seeking to maximize the output class separation. Then, for the optimized fuzzy approach and HNN, the MFs obtained by both methods were compared through simulations. The simulation results showed that tuning of parameters allows output groups to be farther apart and facilitates the predictions made by non-linear classifiers. Also, the number of iterations for the optimization problem was considerably reduced if initialization is performed as proposed in this paper. The number of input sets and the importance of each parameter was automatically determined by the proposed neuro-fuzzy approach. Although the proposed approach provides HNN with explanatory properties, it should be further explored how to do the tuning with HNN while keeping the maximum class separation. Also whether the change in the quantification levels by using semantic constrained optimization improves the interpretability of output MFs is a topic for further research. As a way to validate the MFs obtained by the proposed approach and to increase the
generalization accuracy of the classifiers, we will focus on an active learning approach based on obtained MFs.

**APPENDIX**

The approach introduced in [50] to solve the generalized eigenvalue problem in (12) represents the fuzzy variance ratio \( \eta^2 \) as the Rayleigh quotient

\[
\eta^2 = \frac{c^T(C_G^*G + C^*)G(C_G^*G + C^*)c}{c^T(C + C^*)G(C + C^*)c}.
\]

(15)

The maximum value of this Rayleigh quotient corresponds to the higher eigenvalue in (12). Here, matrices \( C, C^*_G, \) and \( C^* \), represented as the \((Mn, Mn)\) diagonal matrix \( G \) in (12), and the \((M, N)\) diagonal matrix \( \mu \), obtained by maximizing the fuzzy variance ratio \( \eta^2 \), respectively, are computed from the MFs \( \mu_{C_i}(x) \) of input groups, the fuzzy mean \( \overline{p}_i \) within each group for the membership value of category \( C_i \), and total fuzzy mean \( \overline{p}_i \). The \((Mn, Mn)\) diagonal matrix \( G \) is formed from membership value \( \mu_G \), as

\[
G = \begin{bmatrix}
\mu_{G_1}(x_1) & 0 & \cdots & 0 \\
0 & \mu_{G_1}(x_n) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \mu_{G_M}(x_1) \\
0 & 0 & \cdots & \mu_{G_M}(x_n)
\end{bmatrix}.
\]

(17)

Since matrices \( C, C^*_G, C^* \), and \( G \) depend on \( \mu_C \) and \( \mu_G \), the value of \( \eta^2 \) is also a function of \( c \) and the parameters defining \( \mu_C \) and \( \mu_G \). In general, if weights \( c_i \) and the parameters defining \( \mu_C \) are denoted as \( \eta_i \), then the values of \( \theta_i \) are obtained by maximizing the fuzzy variance ratio \( \eta^2 \), through e.g., the steepest descent method as

\[
\frac{\partial \theta_i}{\partial \theta} = \frac{\partial \eta^2}{\partial \theta} = \frac{(\partial B_G \partial \theta) - (\partial T \partial \eta^2)}{T^2} = \frac{\partial B_G \partial \theta}{\partial \theta} - \eta^2 \frac{\partial T \partial \eta^2}{\partial \theta}.
\]

(18)

Here, \( \alpha \) is a number representing the step size of the method. In [50] normal cumulative distribution functions \( \mu(x) = \int_{-\infty}^{x} 1/\sqrt{2\pi} \exp(-y^2/2\sigma^2) dy \) with mean \( b \) and standard deviations \( \sigma \), and MFs

\[
\mu_C_i(x) = \exp\left\{ -\frac{(x - b_i)^2}{2\sigma_i^2} \right\}, \quad i = 1, 2, \ldots, K
\]

(19)

\[
\mu_C_i(x) = 1, (x < b_i) \mu_C_i(x) = 1(x > b_K)
\]

are used to describe the fuzzy sets associated categories \( C_i \) with \( \theta \in \{c, b, \sigma\} \) and their approach is summarized in the following steps.

1) For the dependent variable \( y \), derive the corresponding mean and standard deviation of each item \( r \), compute the fuzzy integral normal density function, and find \( S^r_r = (\mu_G(r, y))/\sum_{n=1}^{N} \mu_G(r, y) \) and \( S_{a} = (\sum_{r=1}^{M} \mu_G(r, y))/\sum_{n=1}^{N} \sum_{r=1}^{M} \mu_G(r, y) \) from obtained \( \mu_G(r, y) \) and observed data \( y_a \).

2) For explanatory variables \( x_i \), consider the MFs for each item \( i \) corresponding to each category \( C_i \) and randomly initialize the value of parameters \( \theta_i \in \{c_i, b_i, \sigma_i\} \).

3) Obtain initial value of \( \eta^2_0 \) from matrices \( C, C^*_G, C^* \), and \( G \) computed from observation data \( x_a, \mu_C(x_a), \overline{p}_i, \) and \( \overline{p}_i \).

4) Change the weights \( c_i \) as

\[
c_i(t + 1) = c_i(t) + \alpha \left( \frac{\partial B_G}{\partial c_i} - \eta^2 \frac{\partial \overline{p}_i}{\partial c_i} \right)
\]

(20)

and find the variance ratio \( \eta^2 \) from (15). Here, \( y_a \) is updated using (11), \( T = \sum_{r=1}^{M} \sum_{n=1}^{N} (y_a - S_{a} \mu_G(y_a))^2 \mu_G(y_a) \) and partial derivatives are computed as

\[
\frac{\partial B_G}{\partial c_i} = 2 \sum_{a=1}^{M} \sum_{r=1}^{n} \mu_G(y_a) \mu_C(x_a) \times \left[ \frac{\partial \mu_G}{\partial c_i} \sum_{i=1}^{K} \mu_G(x_a) \right] \]

\[
\times \left[ \sum_{a=1}^{n} (S^r_a - S_{a} \mu_G(x_a)) \right] \]

\[
\times \left[ \sum_{i=1}^{K} \mu_C(x_a) \right] = 2 \sum_{r=1}^{M} \sum_{a=1}^{n} \mu_G(y_a) \mu_C(x_a) \times \left[ \sum_{i=1}^{K} \mu_C(x_a) \right] \]

\[
\times \left[ \sum_{a=1}^{n} (S^r_a - S_{a} \mu_G(x_a)) \right] \]

\[
\times \left[ \sum_{i=1}^{K} \mu_C(x_a) \right] \]

\[
\times \left[ \sum_{a=1}^{n} (S^r_a - S_{a} \mu_G(x_a)) \right] \]

5) Change the parameters of \( \mu_C_i(x_a) \) as

\[
b_i(t + 1) = b_i(t) + \alpha \left( \frac{\partial B_G}{\partial b_i} - \eta^2 \frac{\partial \overline{p}_i}{\partial b_i} \right)
\]

(21)

\[
\sigma_i(t + 1) = \sigma_i(t) + \alpha \left( \frac{\partial B_G}{\partial \sigma_i} - \eta^2 \frac{\partial \overline{p}_i}{\partial \sigma_i} \right)
\]

(22)

and compute a new variance ratio \( \eta^2_+ \) from new matrices \( C, C^*_G, C^* \), and \( G \) computed from observation data \( x_a, \mu_C(x_a), \overline{p}_i, \) and \( \overline{p}_i \). Here, \( y_a \) is updated using (11), \( T \) is computed as in the previous step and partial derivatives are computed as

\[
\frac{\partial B_G}{\partial b_i} = 2c_i \sum_{a=1}^{M} \sum_{r=1}^{n} \mu_G(y_a) \mu_C(x_a) \times \left[ \sum_{i=1}^{K} \mu_C(x_a) \right] \]

\[
\times \left[ \sum_{a=1}^{n} (S^r_a - S_{a} \mu_G(x_a)) \right] \]

\[
\times \left[ \sum_{i=1}^{K} \mu_C(x_a) \right] \]

\[
\times \left[ \sum_{a=1}^{n} (S^r_a - S_{a} \mu_G(x_a)) \right] \]
\[
\frac{\partial T}{\partial b_i} = 2c_i \sum_{r=1}^{K} \sum_{a=1}^{n} \mu_G_i(y_a) \\
\times \left\{ \sum_{i=1}^{n} \frac{\mu_C_i(x_{ai}) - \sum_{a'=1}^{n} S_{a'} \sum_{i=1}^{K} \mu_C_i(x_{a'i})}{\sigma_i^2} \right\} \\
\times \left\{ \frac{x_{ai} - b_i}{\sigma_i} - \sum_{a'=1}^{n} S_{a'} \frac{x_{ai} - b_i}{\sigma_i^2} \mu_C_i(x_{a'i}) \right\} \\
\frac{\partial B_g}{\partial \sigma_i} = 2c_i M \sum_{r=1}^{K} \sum_{a=1}^{n} \mu_G_i(y_a) \\
\times \left\{ \sum_{i=1}^{n} \frac{\mu_C_i(x_{ai}) - \sum_{a'=1}^{n} S_{a'} \sum_{i=1}^{K} \mu_C_i(x_{a'i})}{\sigma_i^2} \right\} \\
\times \left\{ \frac{(x_{ai} - b_i)^2}{\sigma_i^3} \mu_C_i(x_{ai}) \right\} \\
\times \left\{ \frac{-\sum_{a'=1}^{n} S_{a'} (x_{ai} - b_i)^2}{\sigma_i^3} \mu_C_i(x_{a'i}) \right\}. \\
\text{(23)}
\]

6) If \(|\eta_{i+1}^2 - \eta_i^2| < \delta\) then quit, otherwise, go back to 4.

**ACKNOWLEDGMENT**

The authors would like to thank the five reviewers for all their valuable comments.

**REFERENCES**


Tetsuko Kitaoka received the B.S. degree in education, psychology and human sciences from Aoyama Gakuin University, Tokyo, Japan, in 1984. She received the Ph.D. degree in mechanical engineering from the Tokyo Institute of Technology (TIT), Tokyo, Japan, in 2009.

She was a Popular Counselor accumulating vast experience for more than 20 years. In 2004, she was a Researcher with the Department of Mechanical Sciences and Engineering, Graduate School of Science and Engineering, TIT. She is currently an Associate Professor with Gender Equality Center, TIT, in the fusion of psychology, arts, and engineering for the development of the Iyashi engineering field.

Dr. Kitaoka founded the Iyashi Research Committee within the Japan Society Mechanical Engineering in 2008.

Ichiro Hagiwara received the B.S. and M.S. degrees in applied mathematical engineering from Kyoto University, Kyoto, Japan, in 1970 and 1972, respectively. He received the Ph.D. degree in mechanical engineering from the University of Tokyo, Tokyo, Japan, in 1990.

He was a Researcher with the Research Center of Nissan Motor Co. Ltd., Nissan, Japan. Since 1996, he has been a Professor with the Department of Mechanical Sciences and Engineering, Graduate School of Science and Engineering, Tokyo Institute of Technology (TIT). He has published over 250 journal papers and presented many invited lectures in the research areas. His current research interests include CAD/CAM/CAE, sound and vibration, and Origami engineering.

Dr. Hagiwara is a fellow of the American Society of Mechanical Engineers, the Japan Society Mechanical Engineering, the Japan Automobile Standards Association, the Japan Society of Starting Machinery Ltd., Sumitomo, Japan, since 1990.

He was a Researcher with the Department of Mechanical Sciences and Engineering, Graduate School of Science and Engineering, Tokyo Institute of Technology (TIT). He has received the Ph.D. degree in mechanical engineering from the University of Tokyo, Tokyo, Japan, in 2008.

He has been a Researcher with Sumitomo Construction Machinery Ltd., Sumitomo, Japan, since 2006. His current research interests include facial expression recognition and image processing.