Research Article
A New Method of Diagnosing Constitutional Types Based on Vocal and Facial Features for Personalized Medicine

Bum Ju Lee, Boncho Ku, Kihyun Park, Keun Ho Kim, and Jong Yeol Kim

Division of Constitutional Medicine Research, Korea Institute of Oriental Medicine, 1672 Yuseongdae-ro, Yuseong-gu, Deajeon 305-811, Republic of Korea

Correspondence should be addressed to Jong Yeol Kim, ssmed@kiom.re.kr

Received 21 May 2012; Accepted 30 May 2012

Academic Editor: Sabah Mohammed

Copyright © 2012 Bum Ju Lee et al. 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.

The aim of the present study is to develop an accurate constitution diagnostic method based solely on the individual’s physical characteristics, irrespective of psychologic traits, characteristics of clinical medicine, and genetic factors. In this paper, we suggest a novel method for diagnosing constitutional types using only speech and face characteristics. Based on 514 subjects, the area under the receiver operating characteristics curve (AUC) values of classification models in age and gender groups ranged from 0.64 to 0.89. We identified significant features showing statistical differences among three constitutional types by performing statistical analysis. Also, we selected a compact and discriminative feature subset for constitution diagnosis in each age and gender group. Our method may support the direction of improved diagnosis prediction and will serve to develop a personal and automatic constitution diagnosis software for improvement of the effectiveness of prescribed medications and development of personalized medicine.

1. Introduction

Due to the development of medicine and advances of biotechnology and information technology, the midpoint of medical treatment has shifted away from common treatments of a certain disease to personalized medicine [1–8]. Consistent with this paradigm, there has been an explosion of interest in alternative oriental medicines and in a fusion of oriental and western medicines [9–16]. One of the core research areas of personalized medicine in western and oriental medicine is to understand the psychological characteristics, morphological traits, genetic characteristics, and constitution of individuals. The human constitution has been researched in western and oriental medicine for a long time. For example, Hippocrates suggested that the human constitution could be attributed to four kinds of substances (blood, phlegm, choler, and black bile) [4]. Wang classified humans into seven constitutional groups using physiological and physical status [4, 17]. Similarly, Lee classified humans into four Sasang constitutional types as TE (Teaeumin), TY (Taeyangin), SE (Soeumin), and SY (Soyangin) based on physiological, psychological, and physical characteristics [2, 11–13, 15].

Personal constitution diagnosis is important for several reasons. Firstly, people have vulnerability to particular diseases according to their individual psychological characteristics, genetic characteristics, and morphological traits. Therefore, risk factors for particular diseases can be identified according to an individual’s constitution in the early stages of disease progression [18]. Secondly, drug response to prescribed medicine varies with personal constitution [11]. As such, the efficiency of prescribed medicine can be improved if we know the patient’s constitution.

Many studies on Sasang constitution have been conducted. For the association of Sasang constitutions and diseases and the difference of constitutional types, many researchers introduced constitution analysis methods [1, 4, 5, 8, 12, 13, 15, 18–23]. Song et al. [22] classified TE and SE among the constitutional types using skin elasticity of the hand and proved that elasticity of the TE type was higher than that of the SE type. Their constitution diagnosis measured hand skin elasticity based on a questionnaire survey of thickness and elasticity of the skin. The limitations of the study were that experiments were performed in only TE and SE types and not in the TY and SY types, and elasticity measurements were performed only on the back of
2. Materials and Methods

2.1. Data Preparation. Speech and facial feature extraction from 514 subjects in several hospitals and the Korea Institute of Oriental Medicine.
of Oriental Medicine in the Republic of Korea was carried out. Constitutional types of all subjects were determined by specialists and drug responses [21]. Speech record configurations were no resonance, noise intensity from 40 to 50 dB, room temperature of 20 °C ± 5 °C and humidity of 40% ± 5%, Sennheiser e-835s microphone, Blaster Live 24-bit external sound card, and GoldWave recording program. Distance of the mouth of subjects and the microphone was 4–6 cm, and features were extracted using five vowels (A, E, I, O, U) and one sentence. The extracted features consisted of pitch, average ratio of pitch period, jita (absolute Jitter), MFCC (Mel frequency cepstral coefficients) [29, 30], and so forth.

We took photographs from the side and front of the subject’s face using a digital camera with a ruler (Nikon D700 with an 85 mm lens). Based on an identified feature point from a side- and front-face image, we obtained features such as distance, distance ratios, angle, and area from forehead, nose, mouth, face shape, and eye [20]. Doctors designated the feature points (Figure 1). Height and weight of subject were measured by a digital scale (LG-150; G Tech International Co., Ltd, Republic of Korea). A total of 82 features were used in this study (29 features from speech signals, 51 features from face, and weight and height features). All feature measurements were done based on self-made tool using MATLAB on Window XP. The specific content of the extracted features was described in Table 1.

Since the face and speech signals are influenced by age and gender [21, 31], experimental data were divided into five categories based on age and gender: Female-20 (women aged 18–20 years), Female-25 (women aged 25–30 years), Female-30 (women aged 30–35 years), Female-40 (women aged 40–45 years), and Female-50 (women aged 50–60 years).
Table 2: Comparisons using AUC values in 5 groups (Hybrid: combining face and speech data, FF: using full features, FS: using feature selection).

<table>
<thead>
<tr>
<th>Method</th>
<th>Female-20</th>
<th>Female-30</th>
<th>Female-50</th>
<th>Male-30</th>
<th>Male-50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hybrid-FS</td>
<td>0.89</td>
<td>0.64</td>
<td>0.75</td>
<td>0.72</td>
<td>0.78</td>
</tr>
<tr>
<td>Hybrid-FF</td>
<td>0.67</td>
<td>0.59</td>
<td>0.67</td>
<td>0.67</td>
<td>0.69</td>
</tr>
<tr>
<td>Face-FS</td>
<td>0.77</td>
<td>0.67</td>
<td>0.69</td>
<td>0.73</td>
<td>0.71</td>
</tr>
<tr>
<td>Face-FF</td>
<td>0.62</td>
<td>0.59</td>
<td>0.63</td>
<td>0.63</td>
<td>0.62</td>
</tr>
<tr>
<td>Speech-FS</td>
<td>0.69</td>
<td>0.57</td>
<td>0.58</td>
<td>0.67</td>
<td>0.73</td>
</tr>
<tr>
<td>Speech-FF</td>
<td>0.6</td>
<td>0.48</td>
<td>0.54</td>
<td>0.39</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Table 3: Detailed performance evaluation of experiments using feature selection in all groups.

<table>
<thead>
<tr>
<th>Group</th>
<th>Measure</th>
<th>TE</th>
<th>SE</th>
<th>SY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female-20</td>
<td>Sensitivity</td>
<td>0.89</td>
<td>0.77</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>Specificity</td>
<td>0.93</td>
<td>0.95</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>F-measure</td>
<td>0.87</td>
<td>0.83</td>
<td>0.88</td>
</tr>
<tr>
<td>Female-30</td>
<td>Sensitivity</td>
<td>0.63</td>
<td>0.4</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>Specificity</td>
<td>0.78</td>
<td>0.84</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>F-measure</td>
<td>0.6</td>
<td>0.44</td>
<td>0.52</td>
</tr>
<tr>
<td>Female-50</td>
<td>Sensitivity</td>
<td>0.87</td>
<td>0.17</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>Specificity</td>
<td>0.79</td>
<td>0.97</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>F-measure</td>
<td>0.8</td>
<td>0.26</td>
<td>0.72</td>
</tr>
<tr>
<td>Male-30</td>
<td>Sensitivity</td>
<td>0.81</td>
<td>0.84</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Specificity</td>
<td>0.74</td>
<td>0.78</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>F-measure</td>
<td>0.72</td>
<td>0.75</td>
<td>0.14</td>
</tr>
<tr>
<td>Male-50</td>
<td>Sensitivity</td>
<td>0.89</td>
<td>0.5</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>Specificity</td>
<td>0.81</td>
<td>0.92</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>F-measure</td>
<td>0.81</td>
<td>0.58</td>
<td>0.68</td>
</tr>
</tbody>
</table>

20–29 years), Female-30 (women aged 30–49 years), Female-50 (women aged 50 years and over), Male-30 (men aged 30–49 years), and Male-50 (men aged 50 years and over). The Male-20 (men aged 20–29 years) category was excluded due to the lack of a minimum number of subjects. The numbers of each constitutional type in 5 groups are listed in Table 4 for statistical analysis.

2.2. Experiment Configurations. The goals of our experiment were to measure the ability to distinguish constitutional types, and to identify a more discriminative and compact feature set through feature selection. We conducted classification experiments of TE, SE, and SY constitution types with our five data sets according to the difference of age and gender, with and without feature selection. To investigate the differences of detailed performances of each feature types, speech feature set, face feature set, and hybrid feature set (combining face and speech) were used in this experiment.

We applied normalization (scale 0~1 value) to all data sets. The Wrapper approach using machine learning of LIBLINEAR [32] and the best-first search (forward) was used in feature subset selection. All experiments were performed using LIBLINEAR (L2-loss SVM dual type) in Weka software [33], and a 10-fold cross validation for a statistical evaluation of learning algorithm was performed. For optimal parameter selection (tuning), the value of the C parameter was obtained in the range of {0.01, 0.05, 0.1, 0.5, 1, 2, 4, 8, 16, 32, 64, 128, ..., 16348}, and other parameters were fixed as defaults. The area under receiver operating characteristic curve (AUC) was used as a major evaluation criterion. AUC is widely used to quantify the quality of a prediction or classification model in medical science, bioinformatics, medicine statistics, and biology [34–36]. We also evaluated performance using the sensitivity, specificity, and F-measure for detailed evaluation. Statistical analyses were conducted by SPSS version 19 for Windows (SPSS Inc., Chicago, IL, USA).

3. Results and Discussion

3.1. Comparison of Experimental Results. For brief summarization of performance evaluation, the AUC values for the 5 groups with and without feature selection method are showed in Table 2.

In experiments using full features without feature selection, the results indicated that the hybrid feature set (Hybrid-FF) performed better than the individual face and speech feature sets (Face-FF and Speech-FF), except for performance by the face feature set in the Female-30 group. The AUC values of age and gender classification models using hybrid feature set without feature selection ranged from 0.59 to 0.69%.

After application of feature subset selection, the remaining number of features was small, whereas the AUC values of constitution classification were greater than that of
Table 4: Statistical analysis by the one-way ANOVA test and the post-hoc Scheffe’s test.

<table>
<thead>
<tr>
<th>Group</th>
<th>Feature</th>
<th>Class</th>
<th>N</th>
<th>Mean</th>
<th>Std.</th>
<th>F</th>
<th>Sch.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female-20</td>
<td>Height</td>
<td>TE</td>
<td>18</td>
<td>162.2</td>
<td>5.334</td>
<td></td>
<td>A, B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SE</td>
<td>22</td>
<td>163.8</td>
<td>5.018</td>
<td>3.689</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SY</td>
<td>23</td>
<td>159.7</td>
<td>4.979</td>
<td></td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>Weight</td>
<td>TE</td>
<td>18</td>
<td>62.26</td>
<td>7.328</td>
<td></td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SE</td>
<td>22</td>
<td>52.09</td>
<td>4.689</td>
<td>19.85</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SY</td>
<td>23</td>
<td>51.62</td>
<td>5.776</td>
<td></td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>uF2</td>
<td>TE</td>
<td>18</td>
<td>−0.613</td>
<td>0.750</td>
<td></td>
<td>A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SE</td>
<td>22</td>
<td>−0.531</td>
<td>1.146</td>
<td>4.975</td>
<td>A</td>
</tr>
<tr>
<td>Female-30</td>
<td>FD43_143</td>
<td>TE</td>
<td>57</td>
<td>133.0</td>
<td>5.985</td>
<td></td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SE</td>
<td>53</td>
<td>124.4</td>
<td>7.085</td>
<td>23.55</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SY</td>
<td>73</td>
<td>127.2</td>
<td>7.100</td>
<td></td>
<td>A</td>
</tr>
<tr>
<td>Female-50</td>
<td>Weight</td>
<td>TE</td>
<td>57</td>
<td>61.78</td>
<td>8.052</td>
<td></td>
<td>A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SE</td>
<td>53</td>
<td>52.09</td>
<td>4.689</td>
<td></td>
<td>A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SY</td>
<td>73</td>
<td>51.62</td>
<td>5.776</td>
<td></td>
<td>A</td>
</tr>
<tr>
<td>Male-30</td>
<td>Weight</td>
<td>TE</td>
<td>32</td>
<td>79.34</td>
<td>9.244</td>
<td></td>
<td>C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SE</td>
<td>31</td>
<td>64.45</td>
<td>7.099</td>
<td>26.29</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SY</td>
<td>22</td>
<td>71.22</td>
<td>7.868</td>
<td></td>
<td>B</td>
</tr>
<tr>
<td>Male-50</td>
<td>iJITA</td>
<td>TE</td>
<td>26</td>
<td>0.464</td>
<td>1.282</td>
<td></td>
<td>A, B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SE</td>
<td>18</td>
<td>1.793</td>
<td>4.813</td>
<td>3.456</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SY</td>
<td>25</td>
<td>−0.364</td>
<td>1.144</td>
<td></td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>oJITA</td>
<td>TE</td>
<td>26</td>
<td>0.099</td>
<td>1.023</td>
<td></td>
<td>A, B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SE</td>
<td>18</td>
<td>1.762</td>
<td>4.564</td>
<td>3.558</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SY</td>
<td>25</td>
<td>−0.163</td>
<td>1.004</td>
<td></td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>SITS</td>
<td>TE</td>
<td>26</td>
<td>−0.730</td>
<td>1.182</td>
<td></td>
<td>A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SE</td>
<td>18</td>
<td>0.252</td>
<td>0.763</td>
<td>5.088</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SY</td>
<td>25</td>
<td>−0.042</td>
<td>1.123</td>
<td></td>
<td>A, B</td>
</tr>
<tr>
<td></td>
<td>RSF60_120_240_480</td>
<td>TE</td>
<td>26</td>
<td>−0.148</td>
<td>1.033</td>
<td></td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>FDV81_50</td>
<td>TE</td>
<td>26</td>
<td>32.42</td>
<td>2.788</td>
<td></td>
<td>B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SE</td>
<td>18</td>
<td>29.93</td>
<td>2.744</td>
<td>6.426</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SY</td>
<td>25</td>
<td>29.86</td>
<td>2.953</td>
<td></td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>FArea03</td>
<td>SE</td>
<td>18</td>
<td>4135</td>
<td>417.1</td>
<td>14.23</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SY</td>
<td>25</td>
<td>4331</td>
<td>502.4</td>
<td></td>
<td>A</td>
</tr>
</tbody>
</table>

(N: number of samples, Std.: standard deviation, Sch.: Scheffe’s test, *P < 0.05, **P < 0.01, and ***P < 0.001).
the full feature set in all the experiments. For example, the 
AUC values of classification models using hybrid feature 
set ranged from 0.64 to 0.89%. The values of classification 
models using feature selection showed improvements of 0.22 
in the Female-20 group and 0.09 in Male-50.

When comparing the face (Face-FS), speech (Speech-
FS), and hybrid feature sets (Hybrid-FS), the performance 
of Hybrid-FS method was higher than that of Face-FS and 
Speech-FS in the Female-20, Female-50, and Male-50 groups. 
However, performance of Face-FS method was higher than 
that of Hybrid-FS and Speech-FS in the Female-30 and 
Male-30 groups. Thus, it is preferable to use the Face-FS 
method in these two groups. The theoretical performance 
of the Hybrid-FS method using feature selection is better 
than or equal to that of Face-FS and Speech-FS, because the 
hybrid feature set includes all the speech and face features. 
However, realistically, many of the feature selection methods 
may not ensure better performance because a greater number 
of features add difficulty in building a classifier and lead 
to the curse of dimensionality and an NP problem [36–
38]. Detailed performance evaluation of experiments using 
feature selection was showed in Table 3.

3.2 Statistical Analysis of Meaningful Features. For the 
statistical analysis of features obtained from feature selection 
in Hybrid-FS method, we carried out a one-way ANOVA 
test, and for the post-hoc analysis, we performed the Scheffe’s 
multiple comparison test. The ANOVA test indicated that 
there was a significant difference among constitutional types. 
All features showing $P$ values of $<0.05$ are shown in Table 4.

In this statistical analysis, we did not found features 
that covered a broad range of applicability in predicting 
constitutional types in the age- and gender-specific classification. 
However, we identified several features with an obvious 
propensity for classifications of constitutional type.

In facial feature analysis using results from the Scheffe’s 
multiple comparison test, there was a statistically significant 
difference in weight among constitutional types. Weights 
between TE and the other types were significantly different 
in the Female-20 group ($F = 19.85, P = 0.0000$). Weights 
between TE and the other types in the Female-50 group were 
significantly different ($F = 25.64, P = 0.0000$). Weights among 
TE, SE, and SY were significantly different in the Male-30 
group ($F = 26.298, P = 0.0000$). There was a significant 
difference in $P_{D43,143}$ between TE and the other types in the 
Female-30 group ($F = 23.558, P = 0.0000$) and between 
TE and SE in the Female-50 group ($F = 5.384, P = 0.0374$). 
$FArea03$ between TE and the other types was significantly 
different in the Male-50 group ($F = 14.231, P = 0.0000$).

For speech features, there was a significant difference in 
SITS between TE and SE in the Male-50 group ($F = 5.088, 
P = 0.0088$). The uF2 value between SY and the other 
types was significantly different in the Female-20 group 
($F = 4.975, P = 0.0100$), and $aRF2_F1$ between TE and SY 
was significantly different in the Female-20 group ($F = 3.772, 
P = 0.0286$).

3.3 Limitations and Future Work. Constitution diagnosis is 
very difficult in this area, because constitution type decisions 
are dependent on qualitative judgments of doctors and 
inspectors [22]. We think that it is possible to develop an 
accurate diagnosis method or standardization for constitu-
tion diagnosis after collecting more diagnosis information by 
doctors or inspectors.

Our experimental and statistical analysis showed an 
important and useful feature for better diagnosis based on 
differences in age and gender. Since diagnosis performances 
and selected features differ according to age and gender, it 
makes constitution diagnosis difficult in the real world. Until 
now, features showing an obvious propensity for constitution 
diagnosis are not yet sufficient, and data sets to achieve diag-
noses that are more accurate are insufficient. Accordingly, 
more research for constitution diagnosis is needed.

In future work, we will investigate the relationship 
between constitution and improvement of the effectiveness 
of medications and explore the role of constitutional types in 
certain disease. We think that this is very important research 
in clinical medicine, because the efficiency of prescribed 
medicine can be improved if we know the patient’s consti-
tution. For instance, Jeong et al. [9, 10] investigated changes 
in cytokine production in the acute stage of SY constitution 
CI (cerebral infarction) patients after oral administration of 
Yangkyuk-Sanwha-Tang water, and revealed that Yangkyuk-
Sanwha-Tang had a good effect on anti-inflammatory 
cytokines and a good CI treatment effect. The results of these 
studies may help to improve the effectiveness of prescribed 
medications in Sasang constitutional types.

4. Conclusions

This study describes a novel prediction method for consti-
tution diagnosis as an essential prerequisite for personalized 
medicine or alternative medicine. We demonstrated the 
possibility and usefulness of constitution diagnosis using the 
combination of face and speech feature sets in age- and 
gender-specific groups, identified a compact and discrim-
inative feature subset, and included supporting statistical 
analysis of significant features. Our results could be used 
for developing an automatic constitution diagnostic tool for 
improving the effectiveness of prescribed medications and 
could be used in the fields of speech and face recognition.

Acknowledgments

This work was supported in part by National Research 
Foundation of Korea (NRF) Grant funded by the Korea 
government (MEST) (20110027738).

References

[1] R. P. Ebstein, J. Benjamin, and R. H. Belmaker, “Person-
ality and polymorphisms of genes involved in aminergic 

individualized medicine: psychological and physical traits of 
Sasang typology,” Journal of Alternative and Complementary 


Submit your manuscripts at http://www.hindawi.com