Examinee Identification in e-Test using Press Localized Arc Pattern Method*

Shin'ichi KIKUCHI*1, Takehiro FURUTA*2 and Takako AKAKURA*2

*1 Graduate School of Engineering, Tokyo University of Science, 1-3, Kagurazaka, Shinjuku-ku, Tokyo, 162-8601 Japan
*2 Faculty of Engineering, Tokyo University of Science, 1-3, Kagurazaka, Shinjuku-ku, Tokyo, 162-8601 Japan

Received for publication, April 26, 2010

We propose a method of authenticating an examinee by using answers written by the examinee on a pen tablet, to address the problem that it is easy to impersonate someone taking an examination by e-testing over the Internet, and have evaluated the validity of the method. First of all, we confirmed by experiments that a localized arc pattern method for identifying a writer from characters written on paper can be applied to characters written on a pen tablet. We also propose a press localized arc pattern method that incorporates pen pressures that can be acquired from the pen tablet, and have confirmed its validity by experiments. We applied the proposed method to answers given when examinees take an examination by e-testing, not just by character input, and evaluated the validity of the method. As a result, the proposed method judged that the examinee was the authorized examinee 86.1% of the time, with a 3.3% error rate in which another examinee was judged to be the authorized examinee. From this, it is clear that the proposed press localized arc pattern method has validity in examinee authentication.

Key words: e-test, examinee authentication, biometrics

1. INTRODUCTION

Social factors such as the falling birthrate, aging society, and decreasing population have led to expectations that higher education agencies will have a role in providing support for skills development and re-education to members of the public. However, the conventional types of coursework provided by general higher education agencies in the past have many temporal and spatial restrictions from the viewpoint of ordinary members of the public, making it difficult for them to continue studying.

At the same time, the introduction of e-learning is preceding at various different education sites. If the higher education agencies could also enable a learning environment that can be used effectively to enable learning whenever and wherever the students require, which are the advantages of e-learning, we consider this will be an effective means for solving the problems that members of the public have with studying.

However, tuition leading to unit accreditation by e-learning is provided by 20% of the higher education agencies (National Institute of Multimedia Education 2007). In addition, examinees do not just attend courses through e-learning—they must also take examinations at universities.

In addition to university education, there is an increase in other qualification examinations that make use of computers, such as the Kanken (nation-wide Japanese writing tests) CBT and TOEFL iBT. However, although computers are used in these examinations, the examinees have to travel to education sites to take the examinations.

If we could provide spatial freedom under this current situation, we would be closer to implementing an environment that would make it easier for members of the public to study for examinations for unit accreditation.

Against this background, e-testing that provides examinations over the Internet is attracting attention. It is thought that e-testing, which presents few spatial restrictions on examinees in remote locations such as the countryside, will be extremely effective. However, many education agencies require IDs and passwords (National Institute of Multimedia Education 2007) and, since it is simple to beat the
system by impersonation or cheating during examinations, the introduction of e-testing is not progressing. Fuwa (2006) and Fuwa, et al., (2007) have cited problems relating to the authentication of individuals in e-testing, and write that the only alternative at the moment is to trust the examinees.

There has recently been much research into biometrics for personal authentication using the biological characteristics of humans, as research into personal authentication that is not limited to e-testing (Biometrics Security Consortium 2006). There are two types of biometrics: one uses physical characteristics and the other uses behavioral characteristics. The former type is equivalent to features such as fingerprints, palmprints, faces, and iris patterns, and the latter type is equivalent to features such as voice-prints and signatures.

At the same time, there has been a great deal of research aimed at text creators and examinees as research relating to e-testing, including proposals for methods of creating text tailored to examinees by item response theory (Songmuang and Ueno 2009, Osawa 2001), a proposal for a method of selecting problems according to examinees’ comprehension state by Bayes’ estimations (Katou and Akabori 2001), and a proposal for a method of evaluating the degree of comprehension of examinees by using skills maps (Shimoyama et al. 2007). There have been reports on the possibility of confirming the identity of people by using handwriting information that can be acquired by a pen tablet, by Fujimori, et al., (2008), but there has not been much research at the moment into personal authentication for e-testing.

For the above reasons, our aim is to construct an environment that allows higher education agencies or the like to provide e-learning systems that will enable even people in remote locations to study courses freely, and also provide e-testing for examinations relating to that coursework. In this paper, we focus on e-testing and, in particular, discuss a personal authentication method that is essential for the construction of an environment that enables e-testing. More specifically, of the problems of impersonation and cheating in e-testing, we focus on impersonation and propose an examinee authentication method based on characters written by a pen tablet, and evaluate the validity of that method. The examination method we studied is multiple-choice examination.

Impersonations that can be considered in e-testing can be divided into two main categories: one in which another examinee takes the examination from start to finish, and one in which the other examinee takes only part of the examination. It should be noted, however, that it is difficult to consider another examinee taking the examination for only one or two questions, so impersonation involving another examinee taking the examination for a large number of questions is the subject of this study. In this case, the aim of the study is to ensure that an impersonation that extends over a large number of questions does not result in another examinee being judged to be the authorized examinee, and the authorized examinee is judged to be the correct person when taking the examination.

2. PROPOSAL FOR EXAMINEE AUTHENTICATION METHOD TO PREVENT IMPERSONATIONS IN E-TESTS

2.1. Proposed examinee authentication method for e-tests

With the current form of e-testing in which an ID and password are given at the start of the examination, impersonation is simple (Figure 1). Various tricks could be imagined, such as the examinee inputting his ID and password, then a third party taking the examination instead. One way of preventing impersonation during examinations that could be considered would be to ask for authentication by ID and password repeatedly. However, frequent demands to input ID and password during the examination are not feasible. In addition, it is also simple for the third party to learn the ID and password before the start of the examination.

Against that problem, biometrics that makes impersonation difficult is thought to be valid. Among biometrics, handwritten character authentication could be applied to the handwritten answers of the examinee, and this is thought to present no impediment to the examinee during the examination. In this study, we aim to prevent impersonation by performing examinee authentication that uses the handwritten answers of the examinee that were written on a pen tablet (Figure 2).

When considering handwritten character authentication, impersonation from the time of registration of handwritten character data would be meaningless, so the registration of such data presents a challenge. The e-testing that is the subject of this paper is assumed to be linked to coursework provided by an e-learning system of a higher education agency, as described previously,
which focuses on learners who have registered. In other words, the recorded data is acquired not just for e-testing, but also for use in the e-learning system, so we assume it will be acquired during registration and will be monitored and further recorded data will be acquired during the coursework. It is difficult to believe that an impersonation could continue through registration and all of the coursework, so we consider that impersonation is unlikely to occur at the recording stage with the e-testing system that is the subject of this study, unlike with a one-off examination such as a qualification examination.

2.2. Investigation of examinee authentication method by handwritten answers
Authentication by handwritten characters included authentication using static information and authentication using dynamic information (Biometrics Security Consortium 2006). Static information is configurational information of data such as handwriting. This is mainly used in the authentication of signatures on checks and handwriting appraisal. In contrast, dynamic information is information on the writing process, such as x- and y-coordinates, pen pressure, and the angle and position of the pen, which can be acquired from a pen tablet. Dynamic signature authentication is used in the personal computer security field.

In this study, we envision e-testing in which an examinee uses a pen tablet to answer multiple-choice questions. Since there has been research into dynamic signature authentication that uses dynamic information that can be acquired by a pen tablet, and there are also suitable products available, we consider using dynamic information in this study too. Dynamic signature authentication involves judging the degree of similarity between dynamic information on a recorded signature and dynamic information on a signature that has just been input. Since Japanese names are constructed of a number of complicated characters (Chinese characters) and the owners of those names are used to writing them, they enable authentication because they are characterized in exhibiting individual differences that are highly repeatable. On the other hand, the simple single characters used for decision choices do not have those features, so it is considered difficult to use their dynamic information alone for authentication. Fujimori, et al., (2008) report that, as the result of identifications of single character units from dynamic information when the traditional Japanese syllabary order of “i-ro-ha-ni-ho-e-to” was written, it was difficult to obtain features that would identify actual people.

In this study, we propose an examinee authentication method that uses both dynamic information and static information. First of all, we apply a localized arc pattern method, which is a writer identification method that uses static information (Yoshimura et al. 1991), to e-testing using pen tablets, and investigate the action of examinee authentication. The localized arc pattern method focuses on the ways in which the individuality of a character written on paper is apparent in static information which is the proportions of straight lines and curved lines of the strokes, to perform writer identification. This static information is considered to be apparent even in a character written on a pen tablet. Thus the first part of this study was to perform experiments to check whether or not the localized arc pattern method can be applied to single characters written on a pen tablet. In addition, we consider it valid to combine dynamic information that can be acquired by using a pen tablet, and so investigate the incorporation of dynamic information into the localized arc pattern method.

Of the types of impersonation listed in Table 1, it is difficult to imagine that trained forgeries could be used in e-testing linked to coursework, which is the subject of this study, and also that it is not feasible to produce facsimile forgeries while writing answers under time constraints. In addition, since pen tablets are used, we consider it possible to respond to this problem to a certain degree by acquiring details such as writing speed and number
of re-writings, which is why this paper focuses on simple forgeries as basic research.

3. APPLICATION OF LOCALIZED ARC PATTERN METHOD TO PEN TABLET DATA

In this study, we envision e-testing in which an examinee uses a pen tablet to answer to multiple-choice questions. In this case, we performed experiments to check whether or not the localized arc pattern method is valid for a single character written using a pen tablet (pen tablet data). We first discuss the localized arc pattern method that is valid for identification with a single character written by hand on paper (Yoshimura et al. 1991). We then check whether or not the localized arc pattern method can be applied when pen tablets were used to write the “ā” row of hiragana (Japanese syllabics), which are often used as decision choices in multiple-choice examinations in a similar way as ABCDE in English, as evaluation experiment 1.

3.1. Localized arc pattern method

3.1.1. Feature quantities

The localized arc pattern method (Yoshimura et al. 1991) focuses on a single character written on paper by an unknown writer Q, to identify the writer of that character with reference to the same character that writers P<i>(i=1, 2, ..., a) had recorded previously (called “recorded data” in this paper). In this case, the sample written by pen or pencil is captured by a scanner and turned into a binary image, then is used to identify the writer.

We focus on the way that the individuality of the character is apparent in the proportions of straight lines and curved lines of the strokes, as features for identifying the writer. More specifically, we first consider nxn binary images and define “model patterns nxn” for a number of patterns p from among all of the binary images. Yoshimura, et al., (1991) define for n = 2, 3, 5. As shown by way of example in Figure 3, nine binary images are selected from all of the binary images of the 2×2 regions, and also similar images are combined to produce a model pattern 2×2 that defines five patterns.

Next, we divide the character region that has been turned into a binary image into d×d regions. We then investigate the frequencies that exist within the model pattern nxn for each of the divided regions. The thus-investigated frequencies form a feature vector of p (=p×d×d) dimensions. The example shown in Figure 4 demonstrates a case in which the character region is divided into 3×3 regions.

3.1.2. Writer identification

We perform principal component analysis with respect to p-dimensional feature vectors of the

<table>
<thead>
<tr>
<th>Type of Impersonation</th>
<th>Information Provided</th>
<th>Training Required</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple forgery</td>
<td>None</td>
<td>No</td>
<td>In this study, characters written by other people are used as simple forgeries</td>
</tr>
<tr>
<td>Facsimile Forgery</td>
<td>Character shape</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Trained Forgery</td>
<td>Character shape Writing method</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Types of impersonation and information provided in examinee authentication by handwritten answers

![Fig. 3. Model patterns nxn (n = 2)](image)

![Fig. 4. Character region divided into nxn (d = 3) regions](image)
recorded data, in order to determine which of the registrants $P_i$ is the unknown writer $Q$, from one character written by the writer $Q$, and create a feature space that is effective for identification centered on the principal components. We calculate an adjusted Mahalanobis distance (Kimura et al. 1996) between the feature vectors of the recorded data in this feature space and the character written by the writer $Q$, from the following equation:

$$D(P_i, Q) = \sum_{i=1}^{q} (z^i_0 - z^i_q)^2 / \lambda_k + \sum_{k=q+1}^{p} (z^i_0 - z^i_q)^2 / \lambda_q$$

where $z^i_0$ and $z^i_q$ are the $k$th principal component scores of the writer $Q$ and the registrants $P_i$, respectively, $q$ ($1 \leq q \leq p$) is the number of eigenvalue truncations, and $\lambda_1, \lambda_2, \ldots, \lambda_p$ are eigenvalue. We check this distance for all of the registrants, and determine that the writer $Q$ is the registrant $P_i$ that produces the minimum value.

3.2. Experiment applying localized arc pattern method to pen tablet data (evaluation experiment 1)

3.2.1. Experiment outline

In this evaluation experiment, we check whether or not it is valid to apply the localized arc pattern method when the "**" row of characters, which are often used as decision choices in multiple-choice examinations, have been written on pen tablets.

The conditions for this evaluation experiment were as follows:
- Test subjects: 20 people
- Pen tablet: WACOM Intuos3 A6 Wide
- Number of writings: Sequence of 15, with the five characters "**" to "**" being one sequence

Note that some practice time was provided beforehand, to ensure that the test subjects were familiar with the use of pen tablets.

The evaluation was done with reference to the evaluation method of Yoshimura, et al., (1991), to investigate (1) the model pattern $n \times n$ and the number of character division regions $d \times d$, then (2) determine the average identification rates for the five characters of the "**" row as the result of investigating the number of eigenvalue truncations $q$. The recorded data was also similar to that of Yoshimura, et al., (1991), with eight items for each one character, with the remaining seven items being the data items for identification. Note that we calculate the identification rate by $(100 \times$ the number of determination matches/the number of data items for identification), when the result of the determination of the data items for identification was a match for one of the registrants.

3.2.2. Appropriate combinations of model pattern $n \times n$ and number of character division regions

We investigated combinations of the model pattern $n \times n$ ($n = 2, 3, 5$) and the number of character division regions $d \times d$ ($d = 2, 3, 4$). We assumed that the number of eigenvalue truncations $q$ is 30 and $q = 15$ only when $p < 30$.

The average identification rate for each character type and the average identification rate for the "**" row, for each of the 20 test subjects, is shown in Table 2. The term $Sn$ of $SnDd$ in Table 2 denotes the model pattern $n \times n$ and $Dd$ denotes the number of character division regions $d \times d$. For example, S2D2 denotes the combination of the model pattern $2 \times 2$ and the number of character division regions $2 \times 2$. When we look at the average identification rate for "**", we see that S2D3 gives the best result. Combinations with S3 or S4 gave low identification rates, even though they have larger numbers of features. Ishii, et al., (1998) say that the identification rate will drop in

<table>
<thead>
<tr>
<th>Combination</th>
<th>$p \times q$</th>
<th>$d$</th>
<th>い</th>
<th>う</th>
<th>え</th>
<th>お</th>
<th>Average*</th>
</tr>
</thead>
<tbody>
<tr>
<td>S2D2</td>
<td>5×2×2=20</td>
<td>15</td>
<td>80.0</td>
<td>80.7</td>
<td>72.9</td>
<td>75.7</td>
<td>87.9</td>
</tr>
<tr>
<td>S2D3</td>
<td>5×3×3=45</td>
<td>30</td>
<td>81.4</td>
<td>71.4</td>
<td>77.9</td>
<td>84.3</td>
<td>85.0</td>
</tr>
<tr>
<td>S2D4</td>
<td>5×4×4=80</td>
<td>30</td>
<td>63.6</td>
<td>63.6</td>
<td>70.0</td>
<td>76.4</td>
<td>66.4</td>
</tr>
<tr>
<td>S3D2</td>
<td>25×2×2=100</td>
<td>30</td>
<td>66.4</td>
<td>73.6</td>
<td>65.0</td>
<td>75.7</td>
<td>87.9</td>
</tr>
<tr>
<td>S3D3</td>
<td>25×3×3=225</td>
<td>30</td>
<td>67.8</td>
<td>66.4</td>
<td>71.4</td>
<td>80.7</td>
<td>71.4</td>
</tr>
<tr>
<td>S3D4</td>
<td>25×4×4=400</td>
<td>30</td>
<td>58.6</td>
<td>52.9</td>
<td>62.9</td>
<td>66.4</td>
<td>68.6</td>
</tr>
<tr>
<td>S5D2</td>
<td>67×2×2=268</td>
<td>30</td>
<td>58.6</td>
<td>75.7</td>
<td>66.4</td>
<td>80.0</td>
<td>85.0</td>
</tr>
<tr>
<td>S5D3</td>
<td>67×3×3=603</td>
<td>30</td>
<td>77.1</td>
<td>65.0</td>
<td>71.4</td>
<td>83.6</td>
<td>77.8</td>
</tr>
</tbody>
</table>

**"Average" indicates the average identification rate for the "**" row.
character identification when the number of features increases, because of the mixing in of groups of features that are highly correlated. Thus we consider that the identification rate dropped in this evaluation experiment because the increased number of features resulted in the mixing in of groups of features that are highly correlated. In the next section, our analysis is fixed on S2D3.

3.2.3. Number of eigenvalue truncations $q$
In the previous section, the number of eigenvalue truncations is fixed at $q = 15$ or $q = 30$. In this case, we investigate changes in the identification rate when $q$ was varied from 10 to 45 in steps of 5, to see the effects of changes in $q$ on S2D3.

These results are shown in Table 3. When $q$ is between 15 and 45, the average identification rate for the "し" row shows good results at approximately 80%. Thus, in a similar manner to previous research (Yoshimura et al., 1991), we found that it was best to set the number of eigenvalue truncations $q$ to 15 or more. In addition, the identification rates for "し", "へ", and "け" were on the order of 80%, but in contrast the identification rates for "い" and "ぞ" were on the order of 70%. This is considered to be due to differences in complexity of those characters.

3.2.4. Summary of evaluation experiment 1
From evaluation experiment 1, we determined that the localized arc pattern method can be applied to decision-choice characters ("し" to "ぞ"), by selecting suitable parameters.

The above results suggested that use of the localized arc pattern method is valid for single characters written by using pen tablets. However, the identification rates cannot be said to be sufficient for implementing examinee authentication in e-testing. That is why we investigated the incorporation of dynamic information obtained during writing, which is thought to be difficult to impersonate, to the localized arc pattern method.

4. PROPOSAL FOR PRESS LOCALIZED ARC PATTERN METHOD

In this section, we first discuss local pattern matching (Nakagawa et al., 2002), when proposing a localized arc pattern method that incorporates dynamic information. Based on that, we then propose a localized arc pattern method that incorporates pen pressure. We also discuss an experiment that checked whether or not the proposed method is valid (evaluation experiment 2).

4.1. Local pattern matching
Nakagawa, et al., (2002) define patterns that take into account densities that depend on pixel count, and propose local pattern matching that uses principal component analysis to search for similar images. To confirm the validity of local pattern matching, they prepared two copies of grayscale diagrams of 15 images that are not the same, and report that it is possible to achieve 77% of correct answers as a result of searching for images that are similar to the question image, using level 1 (pixel count 1) and level 2 (pixel count 2) patterns where the density differences between pixels was only 1.

4.2. Proposal for press local pattern matching
To substitute image density for pen pressure, which is one type of dynamic information. For example, writers exhibit individual differences in their pen pressure strengths, such that the pen pressures for straight lines are strong and the pen pressures for curved lines such as pot-hooks are weak. Based on that concept, we propose a press local pattern matching (press method) that incorporates pen pressure.

With the localized arc pattern method, a binary image is created during the image processing in which the presence of writing is coded as 1 and the absence of writing is coded as 0. This is used as the basis for defining a model pattern $n \times n$. With the proposed method, we use pen pressures (at levels 0 to 1023) that can be acquired from a pen tablet, and

<table>
<thead>
<tr>
<th>$q$</th>
<th>&quot;し&quot;</th>
<th>&quot;い&quot;</th>
<th>&quot;う&quot;</th>
<th>&quot;え&quot;</th>
<th>&quot;お&quot;</th>
<th>Average*</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>75.7</td>
<td>71.4</td>
<td>70.7</td>
<td>86.4</td>
<td>81.4</td>
<td>77.1</td>
</tr>
<tr>
<td>15</td>
<td>80.0</td>
<td>73.6</td>
<td>74.3</td>
<td>85.7</td>
<td>81.4</td>
<td>79.0</td>
</tr>
<tr>
<td>20</td>
<td>81.4</td>
<td>74.3</td>
<td>77.1</td>
<td>82.8</td>
<td>84.3</td>
<td>80.0</td>
</tr>
<tr>
<td>25</td>
<td>82.1</td>
<td>71.4</td>
<td>76.4</td>
<td>84.3</td>
<td>84.3</td>
<td>79.7</td>
</tr>
<tr>
<td>30</td>
<td>81.4</td>
<td>71.4</td>
<td>77.9</td>
<td>84.3</td>
<td>85.0</td>
<td>80.0</td>
</tr>
<tr>
<td>35</td>
<td>83.6</td>
<td>70.7</td>
<td>76.4</td>
<td>84.3</td>
<td>85.0</td>
<td>80.0</td>
</tr>
<tr>
<td>40</td>
<td>82.8</td>
<td>70.0</td>
<td>76.4</td>
<td>82.8</td>
<td>84.3</td>
<td>79.3</td>
</tr>
<tr>
<td>45</td>
<td>82.1</td>
<td>69.3</td>
<td>75.7</td>
<td>82.8</td>
<td>83.6</td>
<td>78.7</td>
</tr>
</tbody>
</table>

*"Average" indicates the average identification rate for the "し" row.
divide the pressures into four groups. These groups are 0 for places where there is no writing, 1 for writing where the pen pressure is weak (level 1 to 341), 2 for writing where the pen pressure is normal (levels 342 to 682), and 3 for writing where the pen pressure is strong (levels 683 to 1023). These define the level $n$ ($n = 1, 2$) of patterns that incorporate pen pressures corresponding to the pixel count. Note that sudden changes in pen pressure, such as one from a weak pressure 1 to a strong pressure 3, were not seen to occur when we looked at the data of evaluation experiment 1, so they were omitted. The identification is done in a similar manner to that of Section 3.1.2.

4.3. Proposal for press localized arc pattern method

In this study, the distances we use in the determination also have an additional adjusted Mahalanobis distance, in both the case in which the localized arc pattern method is applied to pen tablet data and the press method. For a single character, we take these as the adjusted Mahalanobis distance $D_1 (P_i, Q)$ for when the localized arc pattern method is applied to pen tablet data and the adjusted Mahalanobis distance $D_2 (P_i, Q)$ of the press method. Distances to which those two distances have been added are checked for all the a registrants, in a similar manner to that of Section 3.1.2, and registrant $P_i$ that gives the minimum value is judged to be $Q$. We call this method the press localized arc pattern method (hybrid method).

4.4. Evaluation experiment on validity of proposed method (evaluation experiment 2)

4.4.1. Experiment outline

In this evaluation experiment, we check whether or not the proposed method is valid for the "は" row, from the identification rate when the localized arc pattern method is applied to the "は" row written on a pen tablet, the identification rate with the press method, and the identification rate with the hybrid method. Using the data of evaluation experiment 1, we select parameters for the press method in (1) and (2), and perform evaluations in (3).

1. Investigation of a suitable combination of level $n$ and the number of character division regions $d \times d$.

2. Investigation by identification rates when the number of eigenvalue truncations $q$ was varied from 10 to 50 in steps of 5, from combinations of the results of (1).

3. Evaluation of the validity of the proposed method, by comparing the identification rate of the press method when parameters have been set based on the results of (2), as well as the identification rate when the localized arc pattern method has been applied to the pen tablet data (results of evaluation experiment 1), and also the identification rate of the hybrid method when the parameters have been set on the basis of the above two identification methods.

4.4.2. Suitable combinations of level $n$ and number of character division regions $d \times d$

We investigate combinations of either level 2 or combined levels 1 and 2 with character division regions $d \times d$. The number of eigenvalue truncations was fixed at $q = 30$.

The average identification rate for each character written by 20 test subjects and the average identification rate for the "は" row are shown in Table 4. L2 denotes level 2 and L1.1.2 denotes a combination of level 1 and level 2. Dd denotes the number of character division regions $d \times d$ ($d = 2, 3, 4$), and P3 denotes pen pressure divided into three stages. The combination that gives the best results for the average identification rate for the "は" row is L2D2P3. The identification rates were lower for other combinations, even though the numbers of features were higher than those for L2D2P3. It is thought that the identification rate dropped because the increased number of features resulted in the mixing in of groups of features that are highly correlated, in a similar manner to evaluation experiment 1. In the next section, our analysis is fixed on L2D2P3.
Table 4. Identification rates (%) due to different combinations of model pattern $n=n$ and number of character division regions $d=x$.

<table>
<thead>
<tr>
<th>Combination</th>
<th>$p$</th>
<th>$q$</th>
<th>ござ</th>
<th>さ</th>
<th>な</th>
<th>に</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2D2P3</td>
<td>28×2×2=112</td>
<td>30</td>
<td>80.7</td>
<td>77.8</td>
<td>73.6</td>
<td>85.7</td>
<td>86.4</td>
</tr>
<tr>
<td>L2D3P3</td>
<td>28×3×3=252</td>
<td>30</td>
<td>75.0</td>
<td>70.0</td>
<td>77.1</td>
<td>82.8</td>
<td>84.3</td>
</tr>
<tr>
<td>L2D4P3</td>
<td>28×4×4=448</td>
<td>30</td>
<td>64.3</td>
<td>65.0</td>
<td>65.7</td>
<td>71.4</td>
<td>74.3</td>
</tr>
<tr>
<td>L1L2D2P3</td>
<td>31×2×2=124</td>
<td>30</td>
<td>77.8</td>
<td>77.8</td>
<td>72.9</td>
<td>80.7</td>
<td>88.6</td>
</tr>
<tr>
<td>L1L2D3P3</td>
<td>31×3×3=279</td>
<td>30</td>
<td>75.7</td>
<td>64.3</td>
<td>67.1</td>
<td>81.4</td>
<td>74.3</td>
</tr>
<tr>
<td>L1L2D4P3</td>
<td>31×4×4=496</td>
<td>30</td>
<td>64.3</td>
<td>64.3</td>
<td>62.9</td>
<td>70.7</td>
<td>69.3</td>
</tr>
</tbody>
</table>

"Average" indicates the average identification rate for the "ござ" row.

Table 5. Changes in identification rate (%) when number of eigenvalue truncations $q$ is varied.

<table>
<thead>
<tr>
<th>$q$</th>
<th>ござ</th>
<th>さ</th>
<th>な</th>
<th>に</th>
<th>Average*</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>72.9</td>
<td>75.7</td>
<td>76.4</td>
<td>72.1</td>
<td>83.6</td>
</tr>
<tr>
<td>15</td>
<td>80.7</td>
<td>77.8</td>
<td>72.1</td>
<td>84.3</td>
<td>85.7</td>
</tr>
<tr>
<td>20</td>
<td>80.7</td>
<td>77.8</td>
<td>72.9</td>
<td>85.0</td>
<td>87.1</td>
</tr>
<tr>
<td>25</td>
<td>80.0</td>
<td>77.8</td>
<td>75.0</td>
<td>85.7</td>
<td>87.1</td>
</tr>
<tr>
<td>30</td>
<td>80.7</td>
<td>77.8</td>
<td>73.6</td>
<td>85.7</td>
<td>86.4</td>
</tr>
<tr>
<td>35</td>
<td>81.4</td>
<td>77.8</td>
<td>76.4</td>
<td>85.7</td>
<td>85.7</td>
</tr>
<tr>
<td>40</td>
<td>80.7</td>
<td>76.4</td>
<td>75.7</td>
<td>85.7</td>
<td>85.7</td>
</tr>
<tr>
<td>45</td>
<td>79.3</td>
<td>75.0</td>
<td>73.6</td>
<td>84.3</td>
<td>84.3</td>
</tr>
<tr>
<td>50</td>
<td>79.3</td>
<td>73.6</td>
<td>74.3</td>
<td>82.9</td>
<td>84.3</td>
</tr>
</tbody>
</table>

"Average" indicates the average identification rate for the "ござ" row.

4.4.3. Number of eigenvalue truncations $q$

In the previous section, the number of eigenvalue truncations $q$ was fixed at 30. In this case, we investigated changes in the identification rate when the number of eigenvalue truncations $q$ was varied from 10 to 50 in steps of 5, to see the effects of changes in $q$ on L2D2P3.

The results of this experiment are shown in Table 5. Except when $q=10$, the average identification rate for the "ござ" row was on the order of 80%, with a good result of 81.4% when $q=35$. We found that it was best to set the number of eigenvalue truncations $q$ to 15 or more, in a similar manner to that found by previous research (Yoshimura et al. 1991) and in evaluation experiment 1.

4.4.4. Comparison of identification rates of identification methods

We evaluate the validity of the proposed method by comparing the identification rate when the localized arc pattern method is applied to pen tablet data (evaluation experiment 1), the identification rate of the press method, and the identification rate of the hybrid method.

The parameters of the localized arc pattern method when applied to pen tablet data was S2D3 with $q=35$, as determined by evaluation experiment 1. The parameters for the press method were L2D2P3 with $q=35$, as determined by Sections 4.4.2. and 4.4.3. From this, the parameters of the hybrid method were S2D3 with $q=35$ and L2D2P3 with $q=35$.

The results of comparing the identification rates of these identification methods are shown in Figure 7. A comparison of the identification rate when the localized arc pattern method is applied to pen tablet data and the identification rate of the press method in Figure 7 shows that the press method produces a better result for "ござ" than the localized arc pattern method. In addition, the average identification rate for the "ござ" row is slightly better at 81.4% with the press method than the 80.0% of the localized arc pattern method. The average identification rate of the hybrid method for the "ござ" row is 90.1%, which is a markedly improved identification rate over those of the other two identification methods. From this, it was considered that the proposed method is valid, and the hybrid method is particularly effective. Note that each identification method exhibits low identification rates for "ござ" and "ざ", in contrast to the high identification rates for "ござ", "ざ", and "ざ". This is thought to be due to differences in complexity of the characters, in a similar manner to that found with evaluation experiment 1.

![Figure 7](image-url)
4.4.5. Summary of evaluation experiment 2

We propose a press method and a hybrid method and have evaluated the validity of those identification methods.

As the result of investigations into suitable combinations of level \( n \) and number of character division regions \( d \times d \), we found that L2D2P3 is a suitable combination as described in Section 4.4.2.

As the result of investigations into identification rates when the number of eigenvalue truncations \( q \) was varied from 10 to 50 in steps of 5, using the combination L2D2P3 that was the result of Section 4.4.2., we found it was best to set the number of eigenvalue truncations \( q \) to 15 or more, in a similar manner to that found by previous research (Yoshimura et al., 1991) and in experiment 1, as described in Section 4.4.3.

We compared the identification rates of the press method set to the parameters obtained in Section 4.4.3., the identification rates when the localized arc pattern method was applied to pen tablet data (the results of evaluation experiment 1), and also the identification rates of the hybrid method when parameters were set on the basis of the above two methods, as described in Section 4.4.4. As a result, it was considered that the proposed method is valid, and the hybrid method is particularly effective. However, that does not mean that the validity was confirmed when the proposed method is applied to answers when examinees were asked to take examinations. In Section 5, we describe an evaluation experiment conducted under taking an examination.

5. EVALUATION EXPERIMENT UNDER TAKING AN EXAMINATION (EVALUATION EXPERIMENT 3)

5.1. Experiment outline

In this evaluation experiment, by applying the proposed method to answers obtained when an examinee answered a number of questions under e-testing conditions, we confirm that the answers of the authorized examinee are judged to be those of that person. In addition, when a simple forgery occurred in which another examinee answered a number of questions, we checked that the examinee was not judged to be the authorized examinee.

This evaluation experiment was performed under the following conditions used in e-testing:
- Examinees: 6 subjects (A–F)
- Examination: 40 five-choice ("A" to "E") English grammar problems created for university entrance examinations
- Date of experiment: December 9, 2008
- Time limit: 30 minutes
- Pen tablet: WACOM Intuos3 A6 Wide
- Date of collection of recorded data: at least one month before the examination (October 21, 2008, and November 4, 2008)

Since it is difficult to imagine that an impersonation could occur for just one question, we used values to which adjusted Mahalanobis distances were added for each block of questions from problems 1 to 5, problems 2 to 6, ..., to problems 36 to 40, for a total of 36 blocks. The evaluation compared the three identification methods according to the following objectives:

1. Confirmation as to whether or not the answers of the authorized examinee can be judged to be those of that person.
2. Confirmation as to whether or not another examinee is not judged to be the authorized examinee, when the answers of the other examinee are simple forgeries over a large number of questions.

Note that the operating system of the computer used for this experiment was Windows XP Professional SP2, the CPU was a 1.20-GHz Intel Pentium M processor, with 512 MB of main memory. Each written answer was turned into a 64x64 binary image. The program for performing the identifications was written in Java, and the times required for determining data by each of the following methods for each character (times required for calculating feature quantities by model patterns and computing the principal component scores) averaged at 0.95 seconds:
- Localized arc pattern method
- Press local pattern matching (press method)
- Press localized arc pattern method (hybrid method)

5.2. Experiment results

5.2.1. Identification when the authorized examinee was judged to be that person

We checked whether or not it is possible to determine that a person taking an examination is the authorized examinee, from answers obtained when the authorized examinee took the examination in practice. The evaluation compared the identification rates when the localized arc pattern method was applied to pen tablet data, the press method, and the hybrid method. The various parameters were: the localized arc pattern method as described in Section 4 was 52D3 with \( q = 35 \), the press method was L2D2P3 with \( q = 35 \), and
the hybrid method was S2D3 with $q = 35$ and L2D2P3 with $q = 35$.

The identification rates when the authorized examinee was identified correctly from among the examinees are shown in Table 6. The identification rate for each examinee with the hybrid method cannot always be said to be better than the localized arc pattern method or press method, but a glance at the average identification rate for the six examinees shows that it is the best result overall. In this paper, we investigate writing authentication, focusing on characters written during e-testing. For that reason, it is necessary to consider psychological influences such as the way that examinees tend to concentrate on the problems and answers rather than on the actual writing process during the examination, and the way in which they get impatient with the time constraints of examinations. Under such conditions, it is considered valid to use a number of methods in an integrated manner, rather than using just one identification method. A look at the results obtained by the localized arc pattern method and the press method for examinees C and E in Table 6 shows that the identification by one of the methods was extremely bad in each case. However, the other method maintained a good identification rate in each case, so that the identification rate by the hybrid method, which integrates those two methods, is around 80%. Similar trends can be seen for the remaining examinees.

From the above, we see that an average identification rate of 86.1% can be obtained for the six examinees by the hybrid method, which is better than the rates obtained by the other two methods, and which is considered useful when checking whether or not the person taking the examination is the authorized examinee.

### 5.2.2. Erroneous identification of another examinee as the authorized examinee

We assumed an impersonation in which, out of 40 questions, the authorized examinee answered the 20 questions in the first half and another examinee answered the 20 questions in the second half (simple forgery). In this case, we check whether or not the other examinee is not judged by the proposed method to be the authorized examinee in the 18 blocks of the second half. The evaluation was done by comparisons of the three identification methods, in a similar manner to that of Section 5.2.1. The parameters was also assumed to be the same as in Section 5.2.1.

The average error rates where another examinee was judged to be the authorized examinee are shown in Table 7. The values in this table are the error rates obtained when an examinee other than the person who answered the first half of questions was brought in to answer the 20 questions of the second half, together with averages. When the average erroneous identification rates for the six examinees are examined, we found an error of 5.0% when another examinee was judged to be the authorized examinee by the localized arc pattern method, but 3.3% by the press method and hybrid method. Thus it is considered that the proposed method is valid not from the aspect of judging that the authorized examinee is the person taking the examination, but also from the aspect of judging that another examinee is not the authorized examinee.

| Table 6. Identification rates (%) when the authorized examinee is identified correctly |
|---------------------------------|--------|--------|--------|--------|--------|--------|
|                                 | A      | B      | C      | D      | E      | Average* |
| Localized arc pattern method    | 94.4   | 50.0   | 30.6   | 97.2   | 97.2   | 100.0   | 78.2    |
| Press local pattern matching    | 75.0   | 75.0   | 100.0  | 52.8   | 94.4   | 82.9    |
| (press method)                  |        |        |        |        |        |         |
| Press localized arc pattern     | 86.1   | 69.4   | 77.8   | 100.0  | 83.3   | 100.0   | 86.1    |
| method (hybrid method)          |        |        |        |        |        |         |

**"Average" indicates the average identification rate for six examinees**

| Table 7. Error rates (%) when another examinee is judged to be the authorized examinee |
|---------------------------------|--------|--------|--------|--------|--------|--------|
|                                 | A      | B      | C      | D      | E      | Average* |
| Localized arc pattern method    | 0.0    | 7.8    | 0.0    | 11.1   | 4.4    | 5.0     |
| Press local pattern matching    | 2.2    | 0.0    | 1.1    | 0.0    | 7.8    | 3.3     |
| (press method)                  |        |        |        |        |        |         |
| Press localized arc pattern     | 0.0    | 3.3    | 1.1    | 0.0    | 8.9    | 6.7     |
| method (hybrid method)          |        |        |        |        |        | 3.3     |

**"Average" indicates the average erroneous identification rate for six examinees**
5.3. **Summary of evaluation experiment 3**

We asked examinees to answer a large number of questions under realistic conditions, and applied the localized arc pattern method, the press method, and the hybrid method to those answers, to evaluate the validity of the proposed method.

As a result of checking whether or not it is possible to judge that the answers of the authorized examinee are those of that person, the identification rate over 36 blocks by the hybrid method was 86.1%, which is better than those of the other identification methods, so the authorized examinee can be judged to be that person as described in Section 5.2.1.

When a simple forgery was performed in which another examinee answered a sequence of a number of questions, we checked that the other examinee was not judged to be the authorized examinee, as described in Section 5.2.2. As a result, the error rate of the hybrid method when another examinee was judged to be the authorized examinee (in the 18 blocks of the second half) was a good result of 3.3%, and we found that the other examinee was not judged to be the authorized examinee.

From the above, we consider that, in most cases, the answers of the authorized examinee are judged to be those of that person, and the answers of other examinees over a number of questions (simple forgery) were not judged to be those of the authorized examinee.

6. **CONCLUSIONS AND FUTURE CHALLENGES**

It is thought that e-testing with few spatial restrictions would be effective for examinees in distant locations, but there are problems in that it is simple to cheat the system by impersonation or cheating. Consequently, in this study we considered the application of the localized arc pattern method, which is a writer identification method that uses single characters written on paper, as an examinee authentication method that focused on impersonation, of those two problems, from consideration of the load on the examinee. As a result of evaluation experiment 1, it is thought that the localized arc pattern method can be applied to pen tablet data. However, the identification rate when the localized arc pattern method is applied to pen tablet data is not thought to be sufficient for application to e-testing, so we have proposed a press localized arc pattern method (hybrid method) which can acquire data from a pen tablet and which incorporates pen pressures during writing that are difficult to impersonate. The results of evaluation experiment 2 showed that whereas the proposed press localized arc pattern method had an identification rate of 90.1%, the identification rate when the localized arc pattern method alone was applied to pen tablet data was 80.0%. It is therefore considered that the proposed method is valid. In addition, by applying the proposed method to answers obtained when an examinee answered a number of questions under e-testing conditions, we have confirmed that the answers of the authorized examinee are judged to be those of that person and, when a simple forgery occurred in which another examinee answered a number of questions, we checked that the examinee was not judged to be the authorized examinee. The results of evaluation experiment 3 showed that the proposed press localized arc pattern method judged that the examinee was the authorized examinee 86.1% of the time, with a 3.3% error rate in which the answers of another examinee (simple forgery) were judged to be those of the authorized examinee. From this, it is clear that the proposed press localized arc pattern method has validity in examinee authentication.

A future challenge would be to prove the validity of the proposed method with an increased number of subjects, since the six examinees of evaluation experiment 3 was too few. In addition, this study only dealt with simple forgery, so we consider it necessary to perform further experiments that introduce facsimile forgery and trained forgery.

**ACKNOWLEDGEMENTS**

This work was supported in part by Scientific Research Grants in Aid for Challenging Exploratory Research from 2007 to 2009 (No. 19650245, representative researcher: Takako Akakura).

**REFERENCES**


Information and Communication Engineers Journal, J82-II(1) : 148–158 (in Japanese)