University of Groningen

ON THE REJECTION ABILITY REQUIRED IN MULTIPLE HYPOTHESIS TECHNIQUES

Sako, H.; Kagehiro, T.; Fujisawa, H.

Published in:
EPRINTS-BOOK-TITLE

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2004

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Sako, H., Kagehiro, T., & Fujisawa, H. (2004). ON THE REJECTION ABILITY REQUIRED IN MULTIPLE HYPOTHESIS TECHNIQUES. In *EPRINTS-BOOK-TITLE* s.n..

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 13-02-2023

ON THE REJECTION ABILITY REQUIRED IN MULTIPLE HYPOTHESIS TECHNIQUES

HIROSHI SAKO, TATSUHIKO KAGEHIRO AND HIROMICHI FUJISAWA

Hitachi Central Research Laboratory
1-280 Higashi-Koigakubo, Kokubunji, Tokyo 185-8601, JAPAN
E-mail: sakou@crl.hitachi.co.jp

Abstract: The so-called multiple hypothesis technique is applied to solve a recognition problem that can be divided into at least two sub-problems. The principle of the technique is to solve the sub-problems by recognisers, a pre-recogniser and a post-recogniser, and to allow the pre-recogniser to leave several possible solutions to the post-recogniser. The pre-recogniser uses several hypotheses based on information or *a priori* knowledge. The post-recogniser tries to solve its assigned sub-problem by using the solutions from the pre-recogniser and different *a priori* knowledge. Therefore, there must be co-operation between the recognisers in order to achieve better total performance in solving the recognition problem. In this study, required abilities of the pre- and the post-recognisers are analysed in order to attain better recognition performance. This analysis gives guidelines for two special factors: number of outputs of the pre-recogniser and required recognition rate of each recogniser. These guidelines are applied to an actual mail-address reading system using a multiple hypothesis technique.

1 Introduction

It is important to understand the characteristics of a recogniser when unexpected patterns are input into a recogniser, which was trained or tuned to recognise only expected patterns in the character domain. This is because multiple hypothesis techniques and classifier combination techniques [1], which have been widely used in practical document-reading systems, must handle unexpected patterns. The multiple hypothesis technique can be applied to a recognition problem that can be divided into at least two sub-problems, such as the separation of hand-written characters and the recognition of separated hand-written characters [2]. This technique allows a pre-recogniser, which is assigned to the first sub-problem, to suggest several answers so as to include a correct one if it cannot give only one correct answer. A post-recogniser assigned to the second sub-problem processes each answer by using additional information and hopefully outputs a correct answer as the result of processing the correct input from the pre-recogniser. Therefore, the pre-recogniser's performance, the number of outputs of the pre-recogniser and the post-recogniser's selection ability are the main factors in determining the total performance of the multiple hypothesis technique, and the post-recogniser must have the ability to reject unexpected inputs because it may receive unexpected inputs from the pre-recogniser.

The next section gives a brief overview of the multiple hypothesis technique, which has been applied to the Japanese-mail-address reading system [3] developed

In: L.R.B. Schomaker and L.G. Vuurpijl (Eds.), Proceedings of the Seventh International Workshop on Frontiers in Handwriting Recognition, September 11-13 2000, Amsterdam, ISBN 90-76942-01-3, Nijmegen: International Unipen Foundation, pp 123-132

by our group, and explains what kind of difficulty the technique can solve. Section 3 analyses the relationship between the total correct acceptance rate and the rejection rate of unexpected inputs mathematically, and it also explains the relationship between the number of outputs, *n*, of the pre-recogniser, its accumulated correct detection rate (the rate of a correct output to *n* outputs) and the correct rejection rate of unexpected inputs in the post-recogniser. It is emphasised that there must be some relationship between these rates so as to increase the total correct acceptance rate while the increase of the error acceptance rate is being restrained. Section 4 presents a numerical simulation and experiments using actual data processed by the mail-address reading system.

2 Multiple hypothesis techniques

This section introduces several examples of the multiple hypothesis technique used in the Japanese-mail-address reading system, whose aim is to read a recipient address. This technique is useful not only for mail-address reading systems but also for other pattern-recognition systems.

2.1 Variation in mail features

Variation in the following features must be taken into account in order to develop a system to read the address on mail.

- *Illustrations*: Japanese mail often has various pre-printed advertisement illustrations, including characters and logos, that are obstacles in detecting an address-block composed of address lines and a recipient's name. At the same time, the location of the address-block varies. Therefore, these obstacles make it difficult to detect the correct address-block of some mail. A multiple hypothesis technique could help to solve this problem.
- *Orientation*: The orientation of the address on mail (i.e. a letter or postcard) is often different. The direction of characters is therefore also different. In this situation, a multiple hypothesis technique could be useful.
- *Characters and words*: The address is expressed by printed characters of a certain font or by hand-written characters. These are obviously different according to the writers even if they write the same address. Moreover, the characters' spaces and dimensions depend on the writer or the printer. In addition to this variety, Japanese addresses are expressed by Chinese characters (Kanji), Japanese characters (Hiragana and Katakana), the Roman alphabet, Arabic numbers and other special symbols. All these characters make up about 3,000 categories in total. These varieties of characters present two main problems. One is the difficulty of segmenting or separating characters in the address line one by one because a Kanji character is often composed of several sub-parts similar to the other characters. The other problem is how to recognise a character out of the 3,000 categories composing address words, such as city names and town names, at a high recognition rate. A multiple

hypothesis technique could also help to solve these problems.

2.2 Multiple hypothesis techniques to cope with mail variety

The mail-address reading system uses the multiple hypothesis technique in order to solve the problems mentioned above. Figure 1 shows the structure of the Mail-Address Reading System (MARS).

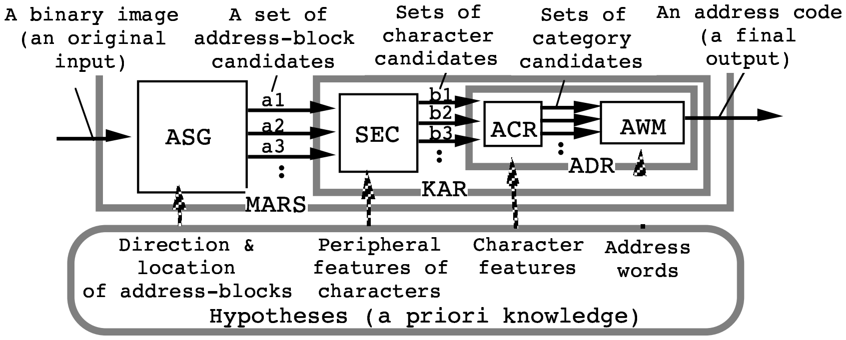
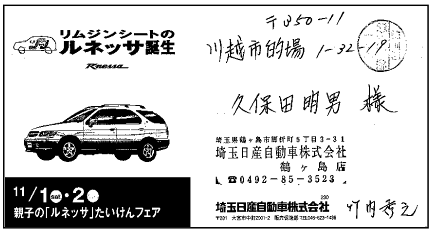


Figure 1. Structure of the mail-address reading system

The purpose of the pre-recogniser, named ASG (Address SeGmentation), in MARS is to detect the address-block from a binary input image. Matching between *a priori* knowledge about typical address layouts and locational information of key-objects such as stamps can determine several possible outputs but cannot determine only one correct output because of the variety of illustrations on each piece of mail. The output of ASG is thus a set of address-block candidates detected according to hypotheses based on *a priori* knowledge. A mail image and its address-block candidates are shown in Figure 2(a).

The address-block candidates are input to the post-recogniser, named KAR (Kanji Address Recogniser), whose purpose is to read a recipient address in the block. KAR examines each address-block candidate. The major components of KAR are a separator (SEC) of the characters in the lines of the block and a recogniser (ADR) of those characters and address words. Clearly, it is very difficult for the separator to detect one correct result because of the structure of a Kanji character. Therefore, several sets of separated character candidates, as shown in Figure 2(b), are determined by SEC according to hypotheses based on statistical information of the peripheral features of Kanji characters [4].

The pre-recogniser (ACR) in ADR examines the objects in each separation candidate by character recognition, which has been trained in advance by using the feature information of characters. Usually, because the recogniser does not always give a correct answer as the first candidate, several category candidates for each object are determined as shown in Figure 2(c). The post-recogniser (AWM) matches the strings expressed by the combinations of category candidates with *a* *priori* knowledge of existing address words (in order to search out address words in the strings) while it checks the compatibility between the end position of the previous word and the start position of the next word [5].



〒350-11
川越市的場 1-32-19
久保田明男 様

埼玉県鶴ヶ島市脚折町5丁目3-31
埼玉日産自動車株式会社
鶴ヶ島店
☎0492-85-3523

230
埼玉日産自動車株式会社 竹内秀之
〒331 大宮市中釘2001-2 販売促進部 TEL048-623-1436

3 Analysis

In the following sub-sections, first, two recognisers for solving a recognition problem are considered. One is a pre-recogniser, for example, ASG, which processes an original input (a binary image in Figure 1) and sends n outputs, of which one or none are correct. The other is a post-recogniser, for example, KAR, which receives n inputs from the pre-recogniser and process them to give a final output (an address code in Figure 1). The characteristics of the recognisers are defined as follows:

- $a(i)$: accumulated correct detection rate in the pre-recogniser, which is the value of the number of correct results divided by the total number of inputs. A result is defined as correct when one correct output exists (between the first and the ith output) for one input, where $0 \le a(i) \le 1$, $i \le n$ and $a(0)=0$.
- p: correct rejection rate in the post-recogniser, which is the value of the number of rejection results divided by the total number of unexpected inputs. A result is defined as correct when the recogniser correctly rejects an unexpected input, where $0 \le p \le 1$.
- r_c: correct acceptance rate in the post-recogniser, which is the value of the number of correct results divided by the total number of expected inputs. A result is defined as correct when the recogniser correctly recognises an expected input, where $0 \le r_c \le 1$.
- r_e: error acceptance rate in the post-recogniser, which is the value of the number of incorrect results divided by the total number of expected inputs. A result is defined as incorrect when the recogniser incorrectly recognises an expected input, where $0 \le r_e \le 1 - r_c$.

Next, two processing procedures, *Proc. A* and *Proc. B*, in the post-recogniser are considered, and they are typical in a multiple hypothesis scheme.

- *Proc. A*: The post-recogniser takes the first acceptance and stops processing the next input.
- *Proc. B*: The post-recogniser processes all n inputs, and the case of more than two acceptances is regarded as a rejection of the original input.

3.1 Formulation

Figure 3 shows a kind of tree whose branches express whether the inputs to the post-recogniser are correct, errors, or rejects. The ith line indicates the case that the ith input is correct. The $(n+1)$th line indicates that all inputs are errors. This tree expression helps to formulate total correct acceptance rate R_c, total error acceptance rate R_e and total rejection rate R_r in *Proc. A* and *Proc. B*. Suffixes a and b are added to distinguish between *Proc. A* and *Proc. B*.

(1) *Proc. A*

$$R_{ca} = r_c \bullet \sum_{i=1,n} \{(a(i) - a(i-1)\} \bullet p^{i-1},$$

$$R_{ea} = 1 - R_{ca} - R_{ra},$$

$$R_{ra} = (1 - r_c - r_e) \bullet a(n) \bullet p^{n-1} + \{1 - a(n)\} \bullet p^n.$$

(2) *Proc. B*

$$R_{cb} = r_c \bullet a(n) \bullet p^{n-1},$$

$$R_{eb} = a(n) \bullet \{r_e \bullet p^{n-1} + (n - 1) \bullet (1 - r_c - r_e) \bullet (1 - p) \bullet p^{n-2}\} + \{1 - a(n)\} \bullet n \bullet (1 - p) \bullet p^{n-1},$$

$$R_{rb} = 1 - R_{cb} - R_{eb}.$$

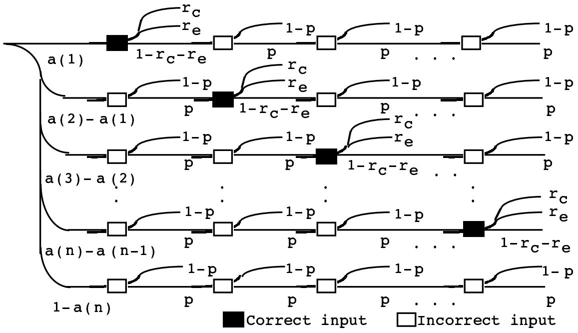


Figure 3. Tree expression for correct and error inputs

3.2 Comparison between correct acceptance rates in Proc. A and Proc. B

The following inequation is always true according to any parameter.

$$\begin{aligned} R_{ca} &= r_c \bullet \textstyle\sum_{i=1,n} \{(a(i) - a(i - 1)\} \bullet p^{i-1} \\ &= r_c \bullet \textstyle\sum_{i=1,n-1} \{a(i) \bullet (1 - p) \bullet p^{i-1}\} + r_c \bullet a(n) \bullet p^{n-1} \\ &\geq r_c \bullet a(n) \bullet p^{n-1} = R_{cb}. \end{aligned}$$

This inequation shows that *Proc. A* (the procedure that the post-recogniser takes the first acceptance and stops processing the next input) always gives a better total correct acceptance rate regardless of the ordering of inputs.

3.3 Changes of total acceptance rate and rejection rate

In this sub-section, the changes of total correct acceptance rate, total error acceptance rate and total rejection rate by increasing *n* in *Proc. A* are analysed. The purpose of this analysis is to judge whether the increase of $a(n)$, *i.e.*, $a(n+1) \geq a(n)$, can also increase the total correct acceptance rate.

$$\Delta R_{ca} = R_{ca}(n + 1) - R_{ca}(n) = r_c \bullet \{(a(n + 1) - a(n)\} \bullet p^n \geq 0.$$

$$\Delta R_{ra} = R_{ra}(n + 1) - R_{ra}(n) = p^{n-1} \bullet f(p),$$

where $f(p) = \{1 - a(n + 1)\} \bullet p^2 + [(1 - r_c - r_e) \bullet a(n + 1) - \{1 - a(n)\}] \bullet p - (1 - r_c - r_e) \bullet a(n)$. Because $f(p) \leq 0$ for $0 \leq p \leq 1$, then $\Delta R_{ra} \leq 0$.

$$\Delta R_{ea} = R_{ea}(n + 1) - R_{ea}(n) = - p^{n-1} \bullet g(p),$$

where $g(p) = \{1 - a(n + 1)\} \bullet p^2 + [r_c \bullet \{a(n + 1) - a(n)\} + (1 - r_c - r_e) \bullet a(n + 1) - \{1 - a(n)\}] \bullet p - (1 - r_c - r_e) \bullet a(n)$. Because $g(p) \le 0$ for $0 \le p \le 1$, then $\Delta R_{ea} \ge 0$.

The above inequations prove that increasing the number of inputs can increase total correct acceptance rate. However, total error acceptance rate also increases. The value $(\beta \bullet \Delta R_c - \Delta R_e)$, where β is a constant, must therefore be introduced in order to examine whether the increase of ΔR_c is much bigger than that of ΔR_e. This evaluation is equivalent to rejection-error performance examination to check whether $(\Delta R_r + \mu \bullet \Delta R_e)$ decreases when $\mu = 1 + (1 / \beta) >> 1$. In *Proc. A*,

$\beta \bullet \Delta R_{ca} - \Delta R_{ea} = p^{n-1} \bullet h(p)$,

where $h(p) = \{1 - a(n + 1)\} \bullet p^2 + [(1+\beta) \bullet r_c \bullet \{a(n + 1) - a(n)\} + (1 - r_c - r_e) \bullet a(n + 1) - \{1 - a(n)\}] \bullet p - (1 - r_c - r_e) \bullet a(n)$. Let p_0 be the bigger root of the quadratic equation $h(p) = 0$, named BICE (Balance of Increases in Correct and Error rates) Equation, then

$\beta \bullet \Delta R_{ca} - \Delta R_{ea} \ge 0$ for $p_0 \le p \le 1$, when $\beta \bullet r_c - r_e \ge 0$,

$\beta \bullet \Delta R_{ca} - \Delta R_{ea} < 0$ for $0 \le p \le 1$, when $\beta \bullet r_c - r_e < 0$.

The evaluation of $(\beta \bullet \Delta R_c - \Delta R_e)$ leads to the following conclusions.

(1) Only when $\beta \bullet r_c - r_e \ge 0$ AND the correct rejection rate p is bigger than p_0, the total performance becomes better as $a(n)$ increases. Here, p_0 is the lower limit of p obtained by $h(p) = 0$ and is therefore a function of $a(n)$, r_c, r_e and β.

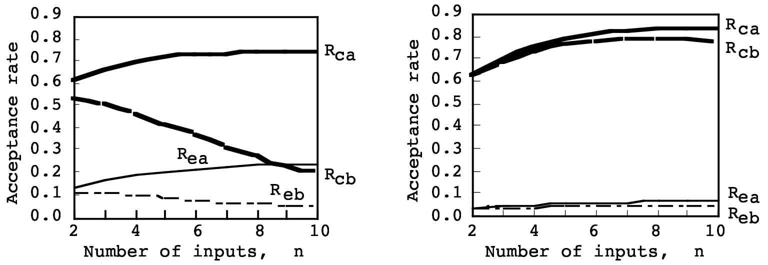
(2) When $\beta \bullet r_c - r_e < 0$, the total performance is degraded whatever the increase of $a(n)$.

4 Experiments and discussion

4.1 Numerical simulation

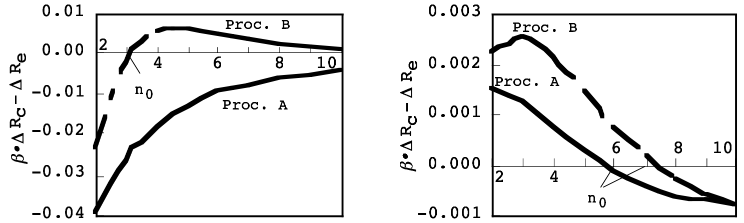
To compare *Proc. A* with *Proc. B* numerically, $a(n)$ is assumed to be given by $1/\{1+\exp(-\gamma \bullet n)\}$ ($\gamma = 0.1$, 0.5 and 1.0) and is shown in Figure 6, and p is assumed to be 0.85 or 0.99, while other parameters are fixed as $r_c = 0.85$ and $r_e = 0.05$. Parameter β is fixed as 0.11, which corresponds to $\mu = 10.0$. The condition $\beta \bullet r_c - r_e \ge 0$ is satisfied because $\beta \bullet r_c - r_e = 0.11 \bullet 0.85 - 0.05 = 0.044 \ge 0$. Figures 4, 5 and 7 show the changes of the rates according to $n = 2$ to 10. The main points from Figures 4-7 are summarised as follows:

- According to the value of the correct rejection rate p in the post recogniser, the total acceptance rate changes drastically, especially when *Proc. B* is used, as shown in Figures 4(a) and 4(b).
- Figure 4 shows that R_{ca} keeps the highest rate and increases monotonously as n increases. And R_{cb} decreases when n exceeds a certain number. The reason *Rcb* decreases is because the condition $p \ge a(n) / a(n+1)$, which can be derived from the inequation $\Delta R_{cb} = R_{cb}(n+1) - R_{cb}(n) \ge 0$, is not satisfied after n exceeds this certain number. When p becomes bigger as shown in Figure 4(b), all acceptance rates become similar.



(a) p = 0.85 (γ = 0.5) (b) p = 0.99 (γ = 0.5)

Figure 4. Acceptance rate vs. *n*



(a) p = 0.85 (γ = 0.5) (b) p = 0.99 (γ = 0.5)

Figure 5. ($\beta \bullet \Delta R_c$ - ΔR_e) vs. *n*

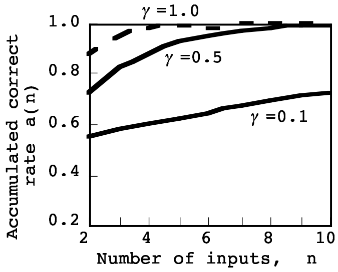


Figure 6. $a(n)$ when γ =0.1, 0.5 and 1.0

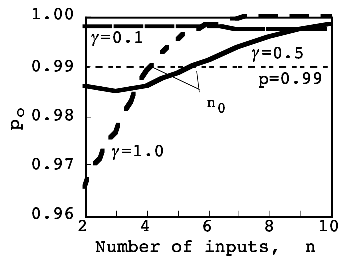


Figure 7. p_0 from *Proc. A* vs. *n*

- In Figure 5, it is important to pay attention to n_0, which is the value of *n* when ($\beta \bullet \Delta R_c$ - ΔR_e) crosses the *n* axis, because n_0 expresses the balance point of $\beta \bullet \Delta R_c$ and ΔR_e. In *Proc. B*, the increase of *n* becomes effective to increase ($\beta \bullet \Delta R_c$ - ΔR_e) after n_0(= 3) when p= 0.85 and is effective until n_0(= 7) when p= 0.99. In *Proc. A,* the increase of *n* is effective until n_0(= 6) when p= 0.99.
- In *Proc. A*, p_0 has an important role because the increase of *n* is meaningful in the sense that $\Delta R_{ca} \geq 0$ *and* ($\beta \bullet \Delta R_{ca}$ - ΔR_{ea}) ≥0 until n_0, the value of *n* that satisfies $p \geq p_0$. As $a(n)$ increases (as γ becomes bigger in Figure 6), p_0 becomes steep and n_0 becomes smaller in Figure 7. As shown in the figure,

when p= 0.99, n_0 is about 6 or 4 when γ = 0.5 or 1.0. This result is reasonable because it indicates that higher $a(n)$ requires lower n_0. When γ = 0.1, p_0 is around 0.998. Therefore, in this case, p must be increased to more than 0.998 from 0.990 in order to raise the total performance.

4.2 Analysis using actual data in the mail-address reading system

In this sub-section, ASG and KAR are selected as examples of a pre-recogniser and a post-recogniser in order to examine their required abilities in the actual mail-address reading system using the multiple hypothesis technique.

Figure 8 shows the accumulated correct detection rate $a(n)$ of the address-block in ASG. The correct rejection rate p of unexpected inputs (wrong address-blocks), the correct read rate r_c of expected inputs (correct address-blocks) and the error read rate r_e in KAR are 0.997, 0.843 and 0.030, respectively. They were measured by experiments using about 1,000 pieces of mail. Figure 9 shows p_0 versus n, which is obtained from BICE Equation $h(p) = 0$. Using these rates, $\beta \bullet r_c - r_e = 0.11 \bullet 0.843 - 0.030 = 0.063 \geq 0$, and n_0 (the upper limit of the number of the address-block candidates) must be four from the figure. At this limit, the best performance, R_{ca} =0.788 and R_{ea} =0.031, can be estimated by the formulas under the criterion ($\beta \bullet \Delta R_c - \Delta R_e$) $\geq$ 0. The measured rates at this limit are R_{ca} =0.804 and R_{ea} =0.028, which coincide well with the above calculated rates.

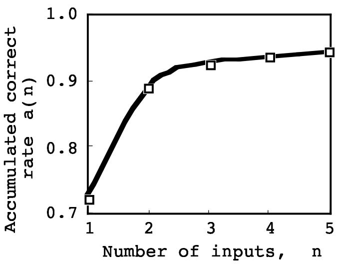


Figure 8. $a(n)$ of ASG

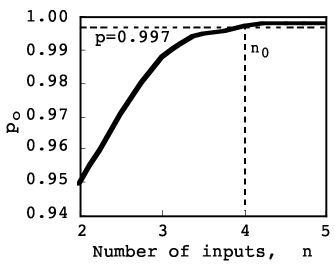


Figure 9. p_0 of ASG

5 Summary

The relationship between the performance of a pre-recogniser and the correct rejection rate of a post-recogniser in a system using a multiple hypothesis technique was mathematically analysed, and the actual recognition rates in a mail-address reading system were numerically simulated. The relationship was analysed mainly from the viewpoint of two special factors: number of outputs of the pre-recogniser and required performance of each recogniser. The main findings of the analysis are summarised below.

- When the criterion $\beta \bullet \Delta R_c - \Delta R_e \geq 0$ is used, the number of outputs n of the pre-recogniser has a certain range determined by the accumulated correct detection rate in the pre-recogniser, the correct rejection rate p and other characteristics of the post-recogniser. And when the post-recogniser takes the first acceptance, BICE Equation $h(p)= 0$, which is derived from this analysis, can give the relationship between rate p and number of outputs n, and can help to find the required performance of p, p_0, and the maximum number of n, n_0.
- Regardless of the ordering of the inputs (outputs of the pre-recogniser), the procedure in which the post-recogniser takes the first acceptance gives a better total correct acceptance rate than the procedure in which the recogniser processes all inputs.

It is thus concluded that the best recognition performance of an actual mail-address reading system using the multiple hypothesis technique can be achieved by using the parameters determined by the recognition performance analysis. This analysis can also be applied to other pattern-recognition systems, as well as mail-address reading systems, in order to increase their performance.

6 Acknowledgements

Many thanks should go to H. Ogata, N. Furukawa, H. Ikeda and M. Koga of the Central Research Laboratory, Hitachi, Ltd., for their discussions on this topic.

References

1. Dengel A. et al., Techniques for improving OCR results, *Handbook of Character Recognition and Document Image Analysis* (Eds. Bunke H. and Wang P. S. P.), (World Scientific, 1997) pp. 227-258.
2. Fujisawa H., Nakano Y. and Kurino K., Segmentation Methods for Character Recognition: From Segmentation to Document Structure Analysis, *Proc. of the IEEE*, **80**(7) (1992) pp. 1079-1092.
3. Fujisawa H., Koga M., Ogata H., Kagehiro T. and Sako H., Recognition Strategies for Japanese Mailpiece Recipient Address, *Proc. of 2nd Int. Conf. Multimodel Interface*, (1999) pp. III69-74.
4. Koga M., Mine R., Sako H. and Fujisawa H., Segmentation Method for Hand-Written Japanese Characters Using Peripheral Feature Analysis, *Proc. of International Conference on Pattern Recognition '98*, (1998) pp. 1137-1141.
5. Marukawa K., Koga M., Shima Y. and Fujisawa H., Error Correction for Recognition of Handwritten Kanji Names Using a Name Knowledge Base, *Proc. of IAPR Workshop on Machine Vision Applications '92*, (1992) pp. 303-307.