

The role of background in recognition of character pairs

Bachelor Thesis

J. Bouwman (j.bouwman.1@student.rug.nl, S1276948)
Supervised by: L. Schomaker and J. van Oosten

February 1, 2015

Abstract

Humans have little difficulty reading handwriting. There are some indications that background has an effect in recognizing familiar objects. Characters can be seen as familiar objects for humans as well. We explore the use of features from the background between character pairs as a means to successfully recognize the characters themselves. What is the role of background in recognition of character pairs? In the cases of machine font stimuli and handwritten stimuli we test if humans are capable of achieving successful recognition using background features. We show them extracted background blobs from character pairs that appear rarely, occasionally or frequently in Dutch text. Results indicate that humans can accomplish recognition with a average success rate of 95% for the machine font condition and 38% for the handwriting condition. This implies background between characters hold features suitable for future character recognition applications.

1 Introduction

Off-line handwriting recognition is a broad contemporary topic. Its usability has been demonstrated in several real-world modern appliances such as: recognizing handwritten ZIP codes during the sorting of letter mail by finding and recognising digits based on pixels and contour (Srihari & Kuebert, 1997); MONK, a search engine that retrieves words from scanned handwritten documents using a collection of algorithms and search techniques (van der Zant, Schomaker, & Valentijn, 2008; van Oosten & Schomaker, 2014); writer identification for authenticity or forensic purposes using contours of connected components

(Schomaker, Franke, & Bulacu, 2007).

Human handwriting recognition can become tedious and inefficient when expressed in time constraints, as opposed to automated handwriting recognition. Off-line unconstrained machine handwriting recognizers are prone to false rejection and high error rates during recognizing as Plamondon and Srihari (2000) sum up in a comprehensive survey of studies. In particular when processed documents contain all sorts of artefacts and noise due to degradation as is often the case in historical documents. For most applications decreasing the error rate and increasing performance is desirable. This research explores the concept of incorporating key features

from the background between characters as a means to recognize characters. If human recognition is successful, background features could hold potential for handwriting recognizing schemes in an attempt to reduce error rates and increase performance.

1.1 Background in human recognition

The question what role background plays in the recognition of characters by humans is still very much an open question. Characters can be described as objects which are being defined by their curvatures and edges. These edges and curvatures come into existence by a difference in brightness or structure between the character and the background. Without background, there would be no distinguishable character. It is already known that background can play a role in human vision when it comes down to completing objects, as illustrated by the Kanizsa triangle (Kanizsa, 1979) and the pacmen figure (Ringach & Shapley, 1996) in figure 1. Ringach and Shapley (1996) showed that human vision is sensitive for object unity and border completion and that humans use shape discrimination with backward masking to determine borders of objects.

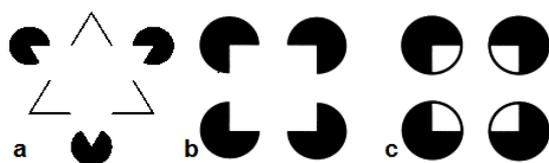


Figure 1: Background playing a role in object recognition. (a) depicts Kanizsa triangle from (Kanizsa, 1979). (b) and (c) depict the pacmen figure in modal and amodal completion from (Ringach & Shapley, 1996).

Furthermore, the recognition of borders

between familiar objects and background has an effect in assigning objects as either a figure or as background as Peterson and Rhodes (2006) showed in figure 2. Here a pineapple can be recognized in (A) and a woman in (C) although these objects both belong to the white background. Characters can be thought of as objects



Figure 2: Image from (Peterson & Rhodes, 2006) with depiction of a pineapple (A), a sea horse (B) and a woman (C) that shows familiarity of objects and background has an effect in recognition.

of human familiarity as well. Baker et al. (2007) showed that when native words or consonants are presented to a reader, a small region of the visual brain is more responsive than when shown other non-character visual stimuli. This prompts the idea that borders between a character and the background might aid in recognizing the character analogue to Peterson's study.

A second indication that background influences the recognition of characters can be found in the research done by Tinker and Paterson (1931), which showed that background color influences reading performance and that legibility is increased when the brightness-contrast between print and background is maximized.

A third factor is letter spacing or inter-character spacing that influences reading performance (Chung, 2002). Although let-

ter spacing influences the horizontal packing, allowing more or less letters to fit in the retina for each eye fixation, it also influences the shape of the background between character pairs which could influence reading performance.

These studies give indications that background influences character recognition by humans.

1.2 Background in machine recognition

Besides humans, machine character recognition schemes exploit information in the background. Particularly during the preliminary stage of segmenting words or characters from images. In general character recognition uses the following steps (Casey & Lecolinet, 1996):

1. Find the next character image.
2. Extract distinguishing attributes of the character image.
3. Find the member of a given symbol set whose attributes best match those of the input, and output its identity.

A range of heuristics already exist to improve handwritten character recognition at various stages. In an overview by Nath and Rastogi (2012) they name:

1. During preprocessing: noise removal, filtering, morphological operations, thresholding, skeletonization and normalization (skew, slant and size).
2. During segmentation: white space pitch, vertical projection analysis, connected component analysis and matching spatial features.
3. During features extraction: using pixels on diagonal as feature, partitioning image in parts, boundary tracing,

Kohonen network, spatial features of strokes and using the direction of pixels and neighbouring pixels.

4. During classification: template matching features, direct matching against prototypes, statistical decision functions, grammatical methods and neural networks.

So far exploiting background into heuristics is seen during the segmentation phase to generate potential cuts between characters. For example, the algorithm by Pal, Belaïd, and Choisy (2003) uses the principle that two touching digits have a large space of background between the digits, that they named a reservoir. This space is used to determine a cutting point around the reservoir, by analysing the position and size of that reservoir.

Another example comes from Chen and Wang (2000). They devised an algorithm which extract skeletons from the foreground and background of an image. The skeletons contain feature points for finding the segmentation path with connected handwritten numeral strings that touch one or more times.

More recently a method by Roy, Bhowmick, Pal, and Ramel (2012) uses the spatial features of background blobs to detect if there is a signature present in a document image and determine its location.

So far background features have not been used in a way to classify characters.

1.3 Research

Incorporating information from the background proves to be valuable for object recognition by machines, in particular during the process of character segmentation. Also, humans are influenced by the background when reading and recognizing char-

acters. In this study we will explore the possibility of using the background as a recognition feature by researching the question: what is the role of background when recognizing character pairs? We will try to answer this question by presenting stimuli to subjects that consist of the background space between two characters and test if the original character pair can be identified. If subjects can identify the character pairs correctly then the background holds features that may also be used to improve handwriting recognition schemes. Based on the notion that humans are influenced by background in several ways we hope to find that background also plays a role in successfully recognizing character pairs.

A few constraints and considerations were added to the experiment to help answer the research question. In this study we refrain from using characters with ascenders or descenders for two reasons. First of all, Schomaker and Segers (1999) showed that ascenders and descenders influence recognition during reading. These characters hold shape information in words that increase recognition rates and we want to eliminate this factor in this experiment. Secondly making background blobs poses additional complications when including the ascender or descender part of a character.

We want to test the feature usability of both machine font characters and handwritten characters. Machine font is absent of connecting strokes, irregular spacing, character size and individual writer differences and this makes a large difference in recognizing schemes.

We use character pairs with a low, medium and high frequency in the Dutch language for two reasons. Subjects may try to guess character pairs based on their

knowledge of a language and occurrences of certain character pairs in the language. Furthermore character familiarity might play a role in recognition. It is not without reason machine schemes use statistics and grammar rules to identify characters and perhaps the frequency will also play a role when it comes to using background as a recognition feature.

We corrected the human handwriting for slant and used normalisation on character size, so slant and size of characters are factored out.

2 Method

2.1 Subjects

This study had 36 native Dutch speaking inhabitants from the province Groningen, the Netherlands. All 36 participants completed the experiment. 17 female; age range: {23...59}; $\bar{x} = 27,8$. 19 male; age range: {25...71}; $\bar{x} = 33,5$. All the participants were asked if they were diagnosed with dyslexia and answered with a negative reply. Participation was on voluntary basis. Informed and written consent was obtained before participation in the experiment.

2.2 Material

For this study two sets of stimuli were developed. Each set contained the backgrounds between the following nine character pairs: 'o o', 'r e', 'v a', 'v u', 'c c', 'a e', 'm w', 'w v' and 'n x'. One set contained stimuli extracted from character pairs in the proportional Arial font. The other set from handwritten character pairs from the archive of the cabinet of the Dutch Queen (Kabinet der Koningin, 1903). Figure 11 (appendix) shows the words that were used from the archive of the cabinet

of the Dutch Queen to obtain the character pairs. In each set the character pairs consisted of three low, three medium and three high frequency occurrences of character pairs in the Dutch language. The frequency was determined using sentences from the Dutch CDBL Newspaper that is part of the Eindhoven Corpus (Uit den Boogaard, 1975). Figure 3 shows the distribution of frequencies of the character pairs in the corpus. Table 1 shows the frequencies of character pairs that were used in the experiment. The character pairs were chosen in such a way that for each of the categories: low, medium and high the frequencies of the character pairs were in close proximity to each other. For the low frequencies there was an additional consideration that the handwritten character pairs had to be available in the archive of the cabinet of the Dutch Queen.

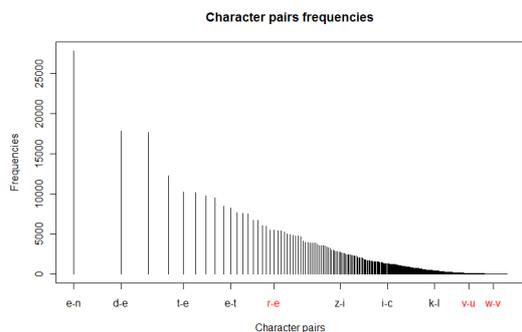


Figure 3: Distribution of all lower case character pairs in the CDBL Newspaper from the Eindhoven Corpus, with character pairs used in the experiment marked in red.

Each stimuli has been centric aligned on an image with a dimension of 600 by 600 pixels, with the median line of a character at a height of 400 pixels and the baseline at a height of 200 pixels. Each stimuli consisted of the region enclosed by the

Table 1: Frequencies in the CDBL Newspaper from (Uit den Boogaard, 1975) of the character pairs used in the experiment.

Low		Medium		High	
Pair	Freq.	Pair	Freq.	Pair	Freq.
n x	1	a e	68	v a	5.401
w v	10	v u	68	r e	5.454
m w	11	c c	70	o o	5.504

two characters, the baseline and the median line. This region was obtained by filling it using the water reservoir principle applied by Roy et al. (2012). A region is filled analogue to water pouring down, filling the region from the baseline to the median line as shown in Figure 4. Then the characters were completely dissolved into the background leaving the filled space as a black foreground.

In the case of the machine font stimuli set no kerning was applied; the width between two characters exists solely due to tabular width. Figure 5 shows the stimuli from the machine font set. In the



Figure 4: Filling of regions between characters using the Water reservoir principle from (Roy et al., 2012). The red lines show the ascender, median, base and descender line.

case of the handwritten character pair set; the characters were extracted from the light brown background using a colour threshold. If the characters were slanted the character pair was slightly rotated to fit the characters in a way that both characters were aligned with the baseline. Then the

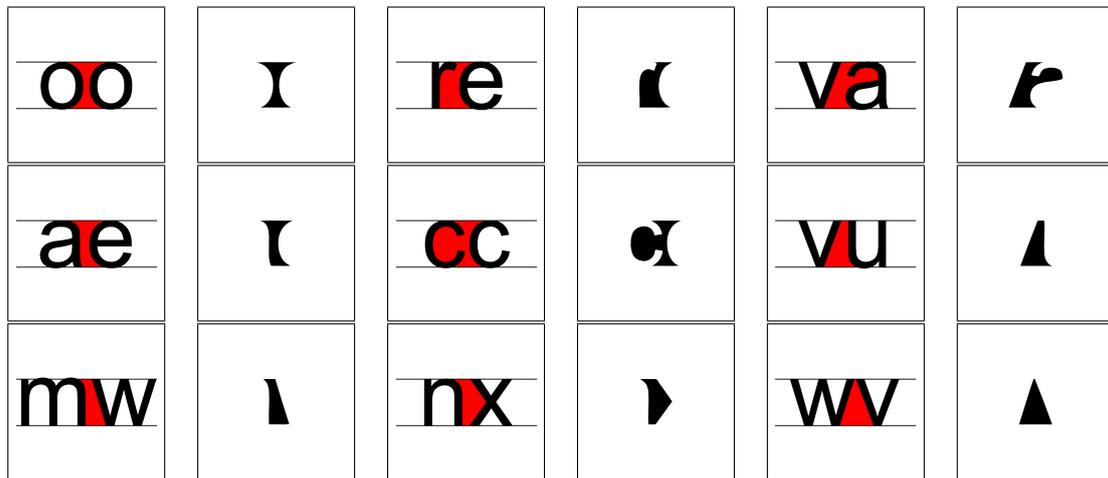


Figure 5: Nine character pairs in the Arial machine font condition. The red areas show the background between the characters, baseline and median line which is extracted to become a stimuli. The resulting stimuli is depicted next to it.

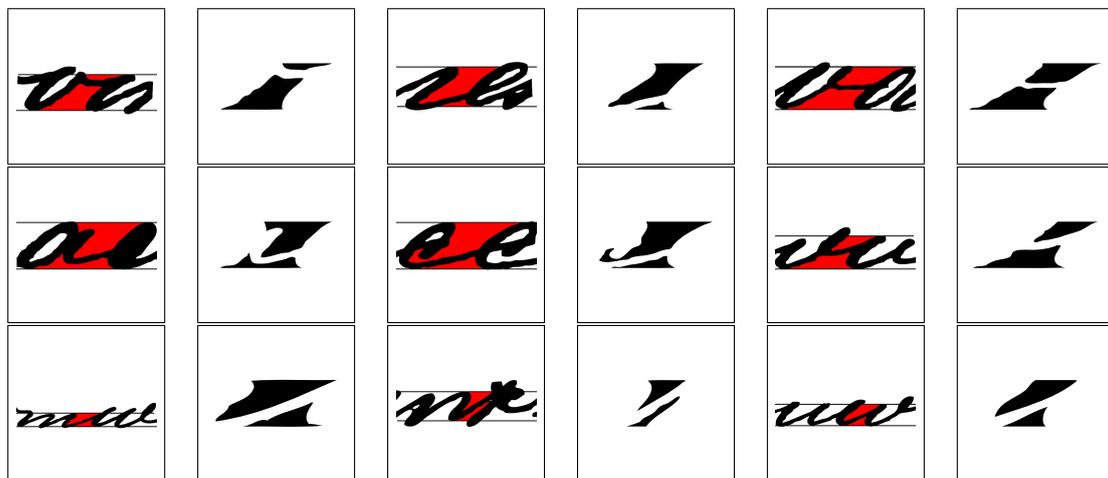


Figure 6: Nine character pairs in the handwriting condition. The red areas show the background between the characters, baseline and median line which is extracted to become a stimuli. The resulting stimuli is normalized for size and depicted next to it.

same water reservoir filling and dissolving paradigm was applied. Finally the images were normalized for size to fit between the baseline and median line. The handwritten stimuli are shown in Figure 6.

For each stimuli a response screen was made. The screen consisted of a multiple choice question with four character pairs. On one random determined row the same character pair as the stimuli was shown. The three other rows were filled with random predetermined character pairs made from the set of alphabetic characters except ascenders and descenders: $\{acemnor suvwxz\}$. Figure 8 shows an example response screen.

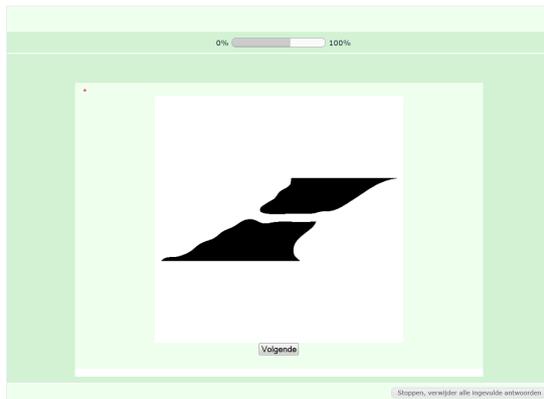


Figure 7: Screen from the experiment showing a stimuli.

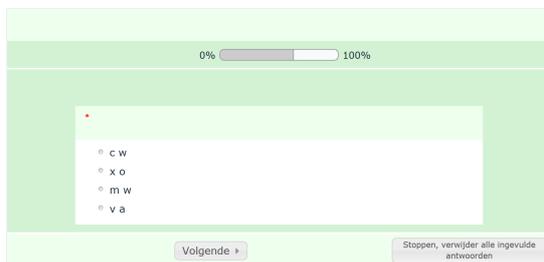


Figure 8: Screen from the experiment showing a multiple choice question.

2.3 Design

The experiment is set up with two dependent variables. A categorical variable letter-type that consists of the factors: machine font and handwriting. A ordinal variable frequency that consists of three factors: low, medium and high. For each variable the responses of three different pairs are summed. The independent interval variable is the percentage of correct answers of the summed character pair stimuli.

2.4 Procedure

The participants were given an on-line computer survey. Participants were shown instructions explaining to read carefully; that there was no time constraint; that they would see images and for each image they could press a next button resulting in the image disappearing and a multiple choice response screen emerging. The experiment consisted of two parts. Each part started with an example of four sentences (machine font / handwriting) so participants could get accustomed to the writing and continued with an example stimuli and correct answer so participants could get accustomed at the task at hand. This example stimuli was not used in the experiment itself. Each part had nine stimuli presented in random order. The first part with the machine font stimuli and the second part with the handwritten stimuli. Participants could not go back to an image when the response screen appeared. The time participants took to look at a stimuli and to fill in responses was recorded. At the end of the experiment the participants were shown a screen stating that the experiment was over and a debriefing about the experiment.

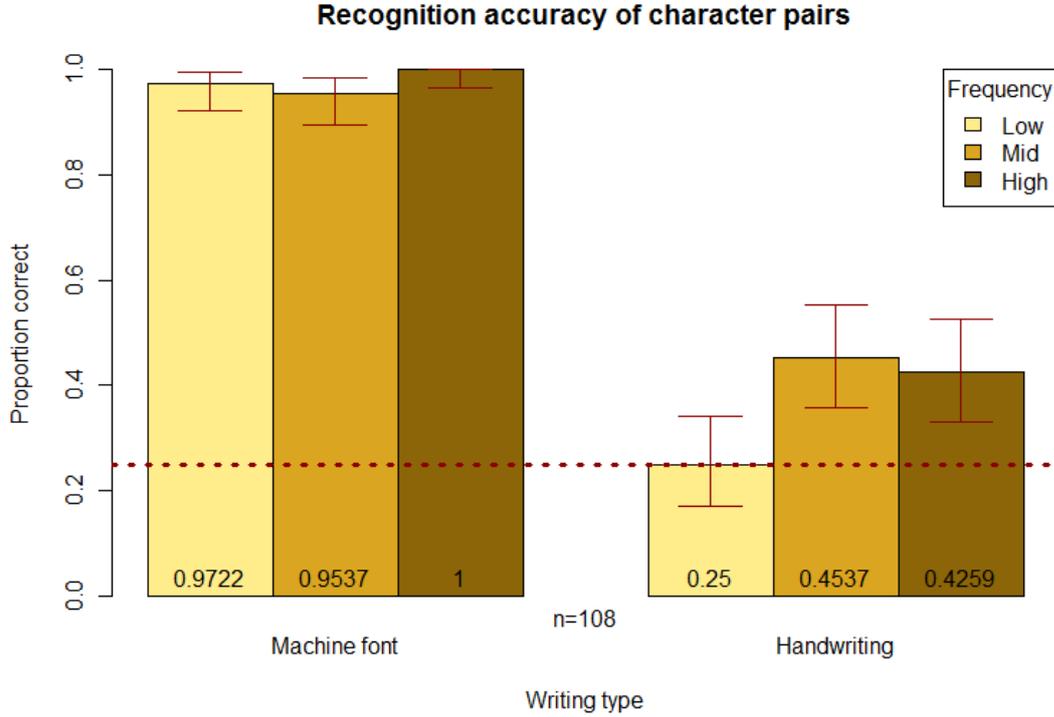


Figure 9: Human recognition accuracy as proportion (95% CI) of character pairs in the machine font and handwriting condition using background features. The red dotted line shows the base success proportion of 0.25.

3 Results

The total accuracy is 97.5% in the machine font condition and 37.7% in the handwriting condition. Figure 9 shows the recognition accuracy of character pairs given the writing type and frequency condition. The red line represents the baseline proportion of $p_0 = 0.25$ that participants would score when taking random guesses. Exact binomial tests were conducted, with sample size $n = 108$ and confidence level 0.95. Table 2 shows the number of correct responses x followed by the p-value under H_0 that the true probability of success is $p_0 = 0.25$. There was a significant effect at $p < 0.05$.

Table 2: Number of correct responses x and p-value in the condition writing type and frequency.

	Machine font	Handwriting
Low	105, $p = 2.2 \times 10^{-16}$	27, $p = 1.0$
Medium	103, $p = 2.2 \times 10^{-16}$	49, $p = 5.7 \times 10^{-6}$
High	108, $p = 2.2 \times 10^{-16}$	46, $p = 7.8 \times 10^{-5}$

In these cases we rejected H_0 and assumed H_A : the true probability of success isn't equal to 0.25. These results suggest that participants were able to correctly identify character pairs beyond ability of chance, except for the low frequency handwriting character pair condition.

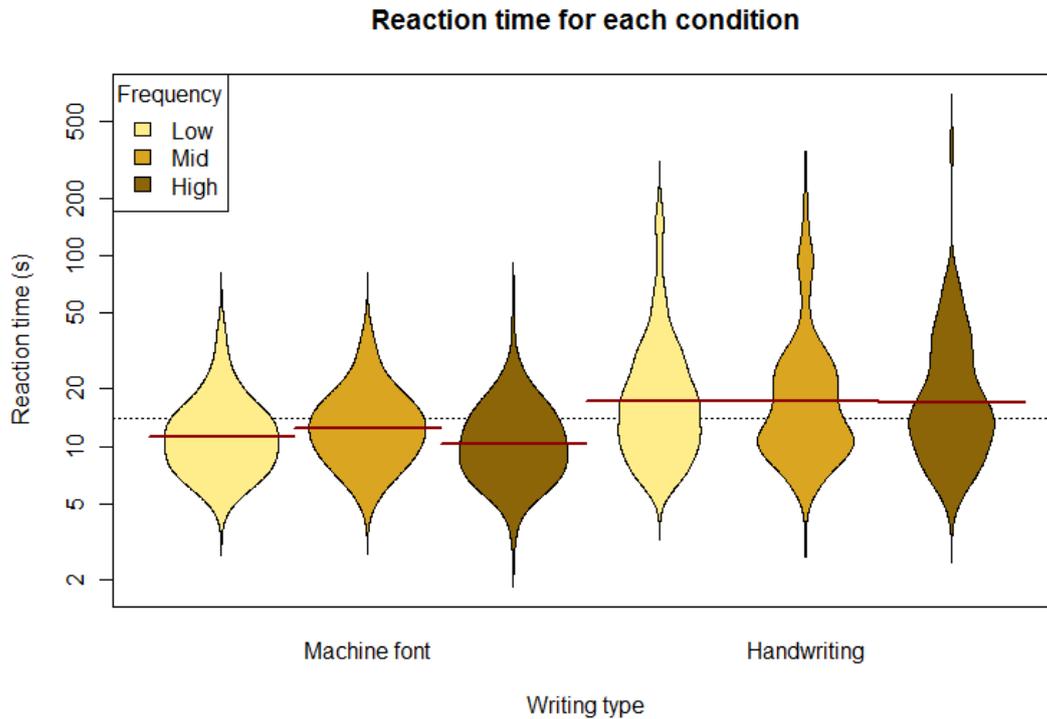


Figure 10: Beanplot showing reaction time in seconds for different character pair frequencies in the machine font and handwriting condition. Shape shows density. The red lines show average reaction times per condition, the dotted line shows total average reaction time.

Figure 10 shows a bean plot with reaction times for the writing type and frequency conditions. A response with a response time of 2277.83 seconds (over half an hour) was treated as outlier and omitted from the results. The mean time of all machine font stimuli (12.60 ± 6.56) and handwriting stimuli (24.11 ± 31.37) is significantly different (twosample t-test, $p < 3.47 \times 10^{-10}$, $DF = 351.328$). There was a significant effect at $p < 0.05$ on response time for frequency in the machine font condition (ANOVA, $p = 0.03247$, $F(2) = 3.4645$). Post hoc analysis using the Tukey's HSD test indicated that the mean response time between medium frequency

(13.84 ± 6.79) and high frequency (11.51 ± 6.13) was significantly different. No significant effect on response time for frequency was found in the handwriting condition (ANOVA, $p = 0.9752$, $F(2) = 0.0252$). These results indicate that there is a difference in response time between machine font and handwritten stimuli. Within the machine font condition the reaction time for the high frequency character pairs is lower than those for the medium frequency character pairs. No such effects were noticeable when comparing response times for handwritten stimuli.

4 Discussion

Our study shows that the features present in the background between character pairs can successfully be used by humans to identify the character pairs and shows that background can play a role in the recognition of characters. In the case of machine font characters the accuracy rate far exceeds the guessing chances. In the case of handwriting it was successful with character pairs that appear in medium and high frequency in Dutch written text. There was no effect found on the low frequency handwritten pairs. Responses showed that in many cases participants gave incorrect replies because they picked the more frequent appearing character pair from the answers in the response screen. Perhaps because they thought character pairs such as 'n x' or 'm w' couldn't exist in Dutch written text. In future research, response screens could be made with character pairs made entirely of character pairs that occur at the the same frequency as the stimuli to see if this is the case.

Accuracy rates and response times indicate that recognizing machine font character pairs is easier for humans than handwritten character pairs. To further understand this difference it would be interesting to study the effects of connecting strokes, irregular spacing and ascenders/descenders on background and recognizing accuracy.

In this experiment a small sample of possible characters pairs was researched. Future work could increase the number of samples by automating the creation of backgrounds images. On successful automating, large quantities of samples can be generated and then it becomes interesting to use the background features of character pairs in machine learning algorithms

to classify character and use the classifier as an extension in handwriting recognition applications.

5 Acknowledgment

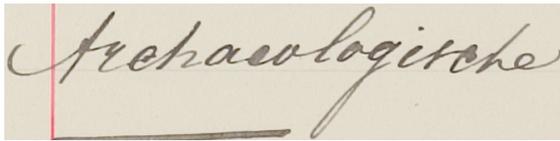
I thank Jean-Paul van Oosten and Lambert Schomaker for the useful remarks and guidance on a preliminary version of this paper.

References

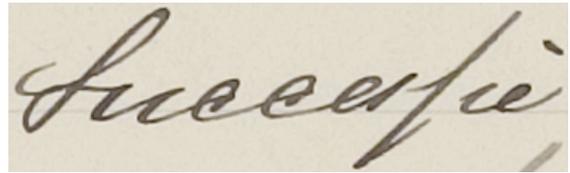
- Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007, May). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(21), 9087-9092.
- Casey, R. G., & Lecolinet, E. (1996, Jul). A survey of methods and strategies in character segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *18*(7), 690-706.
- Chen, Y.-K., & Wang, J.-F. (2000). Segmentation of single-or multiple-touching handwritten numeral string using background and foreground analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *22*(11), 1304-1317.
- Chung, S. T. L. (2002, Apr). The effect of letter spacing on reading speed in central and peripheral vision. *Investigative ophthalmology & visual science*, *43*(4), 1270-1276.
- Kabinet der Koningin. (1903). *Archief van het kabinet der koningin*. Den Haag (Netherlands).
- Kanizsa, G. (1979). *Organization in vision: essays on gestalt perception*. Praeger.
- Nath, R. K., & Rastogi, M. (2012). Improving various off-line techniques used for handwritten character recognition: a review. *International Journal of Computer Applications*, *49*(18).
- Pal, U., Belaïd, A., & Choisy, C. (2003, Jan). Touching numeral segmentation using water reservoir concept. *Pattern Recognition Letters*, *24*(1-3), 261-272.
- Peterson, M., & Rhodes, G. (2006). *Perception of faces, objects, and scenes: Analytic and holistic processes*. Oxford University Press, USA.
- Plamondon, R., & Srihari, S. (2000, Jan). Online and off-line handwriting recognition: a comprehensive survey. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *22*(1), 63-84.
- Ringach, D. L., & Shapley, R. (1996, Oct). Spatial and temporal properties of illusory contours and amodal boundary completion. *Vision research*, *36*(19), 3037-3050.
- Roy, P. P., Bhowmick, S., Pal, U., & Ramel, J. Y. (2012). Signature based document retrieval using ght of background information. In *Frontiers in handwriting recognition (icfhr), 2012 international conference on* (p. 225-230).
- Schomaker, L., Franke, K., & Bulacu, M. (2007). Using codebooks of fragmented connected-component contours in forensic and historic writer identification. *Pattern Recognition Letters*, *28*(6), 719 - 727.
- Schomaker, L., & Segers, E. (1999). Finding features used in the human reading of cursive handwriting. *International Journal on Document Analysis and Recognition (IJ DAR)*, *2*(1), 13-18.
- Srihari, S. N., & Kuebert, E. J. (1997). *Integration of hand-written address interpretation technology into the united states postal service remote computer reader system* (Vol. 2).
- Tinker, M. A., & Paterson, D. G. (1931). Studies of typographical factors influencing speed of reading. VII. variations in color of print and back-

- ground. *Journal of Applied Psychology*, 15(5), 471-479.
- Uit den Boogaard, P. (1975). *Woordfrequenties in geschreven en gesproken nederlands*. Oosterhoek, Scheltema & Holkema, Utrecht. (Werkgroep Frequentieonderzoek van het Nederlands)
- van der Zant, T., Schomaker, L., & Valentijn, E. (2008). Large scale parallel document image processing. In *Proceedings of document recognition and retrieval xv, IS&T/SPIE international symposium on electronic imaging*. (p. 68150S-68150S).
- van Oosten, J.-P., & Schomaker, L. (2014). Separability versus prototypicality in handwritten word-image retrieval. *Pattern Recognition*, 47(3), 1031 - 1038.

Appendix



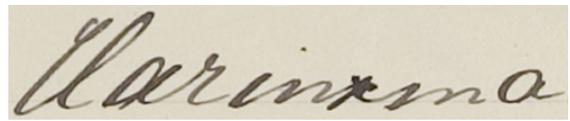
(a) 'a e' contained in Archaeologische.



(b) 'c c' contained in succesje.



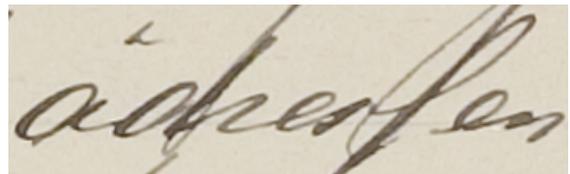
(c) 'm w' contained in stoomwezen.



(d) 'n x' contained in Harinxma.



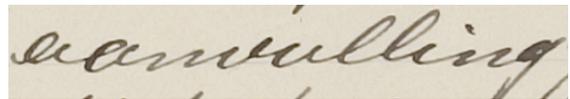
(e) 'o o' contained in voor.



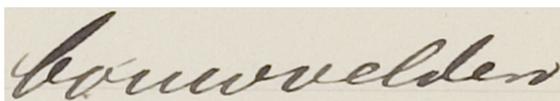
(f) 'r e' contained in adressen.



(g) 'v a' contained in van.



(h) 'v u' contained in aanvulling.



(i) 'w v' contained in bouwvelden.

Figure 11: Words from Kabinet der Koningin (1903) used to obtain the character pairs in the experiment.