Character Image Transformation for Low Vision Aids Based on Learning of Conspicuous Structures

Hideaki Orii ^{a,*}, Masatoshi Nagata ^b, Hideaki Kawano ^b, Hiroshi Maeda ^b, Takaharu Kouda ^a

^a Fukuoka University, 8-19-1 Nanakuma, Jonan-ku, Fukuoka 814-0180, JAPAN

^b Kyushu Institute of Technology, 1-1 Sensui-cho, Tobata-ku, Kitakyushu 804-8550, JAPAN

*Corresponding Author: oriih @ fukuoka-u.ac.jp

Abstract

Many people who have visual disturbance, called as “low vision”, such as old sight and cataract feel inconvenient in daily life. For example, climbing stairs, walking the street, reading information on a newspaper or a book, etc.. Especially, they have a difficulty to read characters without using optical aids. In this paper, we propose a novel algorithm to transform a character appearance for compensating the differences of visual perception such as blurring. In the proposed method, characters in an image are transformed into the characters which can be recognized with ease according to pre-learning rules. In the experiments, we apply the proposed method to various characters and fonts. The experimental results show the effectiveness of proposed method.

Keywords: image transformation, ridge regression, low vision aids.

1. Introduction

In recent years, the number of weak eyesight people including mild low-vision, such as presbyopia and cataracts, has been increased. There are various definitions of weak eyesight. In general, “low vision” means (1) corrected both eyesight is 0.05 to 0.3, (2) though there are restrictions on daily life and learning, the people with visual impairment other than visual acuity can do visual activities. The definition (1) means that using braces, such as glasses or contact lenses, vision doesn't improve enough. Blindness means corrected both eyesight is below 0.05.

The weak eyesight has to live with inconvenience, in a lot of situations e.g. climbing stairs, walking the street,

Fig.1 Comparison of legibility with respect to character size and font.

reading information including route maps and signs at public places, such as stations and airports. Figure 1 shows the comparison of the legibility with respect to character size and font. The difficulty of recognition varies according to the structure and thickness, the size of characters.

To solve these problems, the daily life support systems by wearable computing is studied and developed. Various research and development has been proceeding because of minimization of component and performance improvement in recent years. For example, Tanaka et al. [1] have proposed assistive technologies for transmitting voice to the user recognizes the characters. However, unlike such as the paper, and to recognize the character of any information that is present in daily life there is a technical challenge. Sakamaki et al. [2] have proposed a system to be used for visual support image enlargement technology. We have proposed wearable vision aid system employing eye expression recognition [3]. This system recognizes user's act of squinting by using a wearable camera, and displays the scaled view image to the user. Those are the approaches of increasing the visibility of the user by presenting enlarged visual information. However, consideration of

Fig.2 Overview of proposed method and system.

"visibility of the user" that cannot be realized with simple enlargement is not made, since the previous image enlargement technologies have been developed for the purpose of maximum reproducibility.

In this paper, we propose a novel algorithm to transform a character appearance for compensating the differences of visual perception, e.g. blurring, visual field limitation, etc.. In the proposed method, characters in an image are transformed into the characters which can be recognized with ease according to pre-learning rules.

2. Proposed Method

There are individual differences in appearance of weak eyesight depending on the symptoms. We focus attention on the defocus, such as myopia (shortsightedness), hyperopia (farsightedness) or presbyopia. Figure 1 is obtained by simulating the characteristics. This example shows that even if the sighted people can read characters, those could not be able to read in low vision. Even when presented with expansion, there may be difficult to read by the font type. In this way, it is insufficient to only make the presentation of a simple extension image for partially sighted. Those expansions in consideration for "visibility" and structure transformation of character are required.

In this study, we assume proposed method will be implemented to the vision support system using a wearable camera. A transformed character image by proposed method is presented for user using AR (Augmented Reality) technology. Figure 2 shows the overview of proposed method and system. In the proposed method, a skeleton of the character is extracted by applying the thinning process on the character image. Next, skeleton images are transformed based on pre-learning rules. Finally,

(a) Hiragana "あ". (b) Hiragana "え". (c) Hiragana "ほ".

Fig.3 Skeleton of the characters.

transformed character images are presented for a user by AR system. The detail of these processes is described in Sec. 2.1 and 2.1.

2.1 Character Skeleton Extraction

In this section, we describe skeleton extraction of input character image as a pretreatment for structure transformation. The skeleton of character is obtained as its thinning image. Generally, the characters have a wide variety of characteristics, such as size and font, thickness, in the living environment. In order to deal with diversity of these characters, the thinning process is used in this study. The otiose information of character, such as the thickness or the decoration, is lost by the thinning process. By learning the structure transformation processing for this skeleton image, the proposed method may have applicability to the characters of the type and font thickness different.

First, the input character image is binarized for value "0" or value "255". Then, the binarized image is thinned. Figure 3 shows the skeleton of the characters as the result of thinning.

2.2 Structure Transformation Based on Pre-Learning

In this section, the structure Transformation process for the character skeleton image obtained by thinning. In order to reflect the characteristics of the thickness or font type that the user feels to be "easy to read", let the user select the character from same characters on various fonts. The parameter of structure transformation is learned using it. The detail of these processes is described in Sec. 2.2.1 and 2.2.2.

2.2.1 Pre-Learning

In this section, we describe the learning of structure transformation parameter based on the sensitivity of the user. In pre-learning, first, the user choose the ideal appearance character image about recognition on each characters. Thus, bad appearance character images about recognition are obtained as a result this process. Then, the bad appearance character's skeleton could be obtained by

Fig.4 Build process of learning image data set.

Fig.5 Build process of learning vector data set.

thinning of those. The learning image set is made from those image pairs. Figure 4 shows the build process of the learning image data set.

Next, the vector data set for learning is made from this image data set. Figure 5 shows the build process of the learning vector data set. A pair of vector data, y_i and x_i , is obtained by raster scanning on skeleton image and ideal appearance image. Thus, skeleton character image vector data y_i and ideal character image vector data x_i are $D(=S_w \times S_w)$ dimensional vector data, when the patch size is $S_w \times S_w$.

In this paper, the relationship between y_i and x_i is learned with ridge regression. In this ridge regression, a projection parameter matrix representing the relationship between the vector data pairs is obtained as a matrix to minimize the diagonal elements of the following equation:

$$\sum_i \{ (y_i - Ax_i)(y_i - Ax_i)^T + \alpha A^T A \} \quad (1)$$

This objective function is obtained by adding a regularization term to the least squares method in the regression analysis. The purpose of this restriction is to be constant below a certain sum of squares parameter. Therefore, an estimated matrix of a projection parameter matrix can be obtained as following equation:

Fig.6 Result of proposed method. Hiragana “れ”.

$$\hat{A} = [X^T X + \alpha I]^{-1} X^T Y \quad (2)$$

Here, let $X = [x_1, x_2, \dots, x_N]$ be the matrix consist of ideal appearance vector data, let $Y = [y_1, y_2, \dots, y_N]$ be the matrix consist of skeleton character image vector data, let I be an identity matrix of size D , let $\alpha \geq 0$ be the complexity parameter.

2.2.2 Structure Transformation

In this section, the details of the structure of the character conversion process is described. In this process, the input character image is skeletonized in advance. First, input vector data p_i ($i = 1, \dots, N$) are obtained from an input skeleton image by raster scanning along with learning process. Then, the ideal character structures $\hat{q}_i$ are estimated from each input vector data p_i as following equation:

$$\hat{q}_i = \hat{A} p_i \quad (3)$$

Here, $\hat{A}$ is a projection parameter matrix obtained in Sec. 2.2.1. Then, an output image is reconstructed from the estimated vectors $\hat{q}_i$ ($i = 1, \dots, N$). The estimated value substitutes the corresponding point on an output image. Finally, the Transformed character image is obtained as binarized image of the output image.

3. Experimental Results and Discussions

In this section, we apply the proposed method to the various character images, to verify its effectiveness. Figure 6, 7, and 8 are the result of applying the proposed method to Hiragana “れ”, “ほ”, and “を”. As shown those figures, thickness of the characters in transformation results are almost the same despite the input characters have various thickness. However, there are some hollowing places in

Fig.7 result of proposed method. Hiragana “ほ”.

Fig.8 Result of proposed method. Hiragana “を”.

each character. This is presumably because it does not perform learning good. For patch learning, it occurs that some vector data include almost no character skeleton. In this paper, we introduce "filling" process to before final output. Thus, those hollowing place could be filled.

Figure 9 shows the result of a document image, and Fig. 10 shows the comparison of simple enlarged images for a blur of the "visibility". Figure 10 is obtained by simulating the appearance of partially sighted by the Gaussian filter. As can be seen from Fig. 9, the filling process is functioning effectively. Further, as shown in Fig. 10 (b), it can be seen that clear, easily recognizable and clearly character who has been subjected to structural transformations.

4. Conclusions

We propose a novel algorithm to transform a character appearance for compensating the differences of visual perception. In the proposed method, characters in an image are transformed into the characters which can be recognized with ease according to pre-learning rules. We apply the proposed method to various characters and fonts. The results show the effectiveness of proposed method. In the future, we would like to study the presentation method of

はFナンバーと呼ばれる。また、 $\cos \alpha$ の値で暗くなること（周辺減光）を表す。実際は、 $\cos \alpha$ が含まれており、レンズが筒状になり、 $\cos \alpha$ は、対象表面の明るさである放射輝度 L は、 L が光が対象表面で反射したものである。それと、 $\cos \alpha$ が与えられていると、あとは対象表面の放射輝度 L の対象表面での分布は、反映していることになる。そこで、画像上、画像上の陰影から対象表面の空間での向きを推定する可能性がある。これが、陰影からの形状

(a) input image.

(b) enlarged image.

Fig.9 Result of a document image.

(a) unprocessed image.

(b) proposed method.

Fig.10 Comparison of "visibility" of blurred images.

the processing result to respond to various symptoms, such as "field of view is narrow", and, the various parameter automatic setting methods such as patch size on learning.

References

- (1) M. Tanaka, H. Goto : “Wearable Reading Assistant Device for the Blind,” IPSJ SIG Notes. CVIM 2008(115), pp.125-130, (2008)
- (2) T. Sakamaki, T. Yamaguchi, K. Shimamura : “A Magnification Control for the Vision Support Function Based on the Magnifying Glass Metaphor,” Int. Conf. Next Era Information NETworking (NEINE'06) , pp.99-102, (2006)
- (3) H. Orii, M. Aoki, H. Kawano, H. Maeda : “Eye Expression Recognition of Wearable View Magnification Interface for Low Vision,” The 1st International Conference on Industrial Application Engineering 2013 (ICIAE2013), pp.194-197, Mar. 27-28, (2013)