Optimizing Chinese Character Displays Improves Recognition and Reading Performance of Simulated Irregular Phosphene Maps

Yanyu Lu,[1] Han Kan,[1] Jie Liu,[1] Jing Wang,[1] Chen Tao,[1] Yao Chen,[1] Qiushi Ren,[2] Jie Hu,[3] and Xinyu Chai[1]

[1]School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China
[2]Department of Biomedical Engineering, College of Engineering, Peking University, Beijing, China
[3]Department of Knowledge-Based Engineering, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China

PURPOSE. A visual prosthesis may elicit an irregular phosphene map relative to a regular electrode array. This study used simulated irregular phosphene maps as a way of optimizing the display methods of Chinese characters (CCs) to improve recognition and reading performance.

METHODS. Twenty subjects with normal or corrected sight participated in two experiments (9 females, 11 males, 20–30 years of age). Experiment 1: two character display methods were proposed: selecting phosphenes covered by character strokes on a simulated phosphene array (projection method) and finding the phosphene closest to the expected location in some range of an irregular phosphene array as a substitute (nearest neighbor search [NNS] method). The recognition accuracy of CCs was investigated using six levels for the coverage ratio of stroke and phosphene area and for search range, respectively, for two methods, for several irregularity levels. Experiment 2: reading accuracy (RA) and reading efficiency (RE) were measured using the regular array correspondence and NNS methods.

RESULTS. Experiment 1: projection and NNS methods were significantly affected by coverage ratio or search range. NNS significantly improved CC recognition accuracy to the highest at $81.3 \pm 2.7\%$ and $59.1 \pm 5.2\%$, respectively, for different irregularity levels, compared with the projection method. Experiment 2: RA and RE significantly decreased as the distortion level increased; NNS significantly improved RA (from approximately 40% to >80%) and RE (from approximately 13 char/min to >40 char/min) when reading more irregular paragraphs.

CONCLUSIONS. The performance of CC recognition and paragraph reading when using an irregular phosphene array can be improved through optimizing the display method.

Keywords: retinal prosthesis, simulated prosthetic vision, Chinese character displays, character recognition, paragraph reading

Age-related macular degeneration (AMD) and retinitis pigmentosa (RP) are the two most common retinal degenerative diseases that cause a loss of vision.[1] These two diseases result in the loss of photoreceptors or the retinal pigment epithelium, but nerve cells of middle and inner retina remain viable. Visual prostheses are being developed to restore functional vision by electrically stimulating different locations along the remaining visual pathway that elicit phosphenes and create a discrete visual percept (cortex,[2] optic nerve,[3] or retina[4]; Zrenner E, et al. *IOVS* 2007;48:ARVO E-Abstract 659). A retinal prosthesis can be categorized as epiretinal, subretinal, or suprachoroidal based on the location of the implant. An epiretinal device is implanted on the inner surface of retina and stimulates retinal ganglion cells and axons,[4] whereas a subretinal device is implanted below the outer surface of the retina (Zrenner E, et al. *IOVS* 2007;48:ARVO E-Abstract 659) and a suprachoroidal device is placed in a sclera pocket.[5]

Retinal prostheses are neurostimulators that generally use an external camera and an image processing chip to drive the implanted electrodes; an exception to this is a subretinal microphotodiode array.[6] The stimulating electrode array is often placed near the fovea to better utilize what remains of the high-resolution visual pathway and is expected to produce a regular phosphene map. However, the distribution of ganglion cells is distinctly nonuniform with regions around the fovea 5 to 7 cells deep,[7] and in nonfoveal regions, axons from more peripheral sites form a nerve fiber that lies above the retinal ganglion cell bodies. This distribution of cells and axons can distort the creation of a phosphene map that matches a regular electrode array. Humayun et al.[4] indicated that if the axons from other areas of the retina were stimulated in addition to bipolar and ganglion cell somata, the reported percepts did not always match the shape of the electrode stimulation pattern. Rizzo et al.,[8] on the basis of short-term surgical trials, reported that patients had percepts that matched the stimulation pattern only 48% and 32% of the time for single- and multiple-electrode trials. Similarly, Fujikado et al.[5] reported that the topographical correspondence between the gravitational center of the perceived phosphene and each electrode was not always consistent when using a suprachoroidal prosthesis.

Irregular phosphene maps will distort the geometry of the presented objects or characters. To study the effect of this phosphene map distortion on the performance of visual tasks, some studies modeled irregular or stochastic phosphene maps based on descriptions from human trials and then conducted psychophysical simulation studies, which have been used to understand the requirements of a visual prosthesis and assess the performance on daily tasks.[9–14] Hallum et al.[15] investigated stochastic fields of phosphenes and the performance of inaccurate phosphene mapping by using techniques based on Fourier analysis. Cai et al.[16] adopted a two-dimensional normal distribution to devise a model of the phosphenes' positional uncertainty. Visual acuity was tested on normally sighted subjects and the results showed a decrease of visual acuity as the irregularity index increased. Our previous study[17] investigated the effect of distortion on the recognition of pixelized Chinese characters (CCs), which were processed according to a regular array correspondence method (the greater the offset of a phosphene's location elicited by an electrode, the greater the character distortion formed by the phosphene map). The results showed that recognition accuracy significantly decreased with an increase in the distortion level; for the recognition of CCs at higher distortion levels, it decreased to <60%.

Building on the previous methods and results, the present study focused on optimizing the display of CCs to improve character recognition and paragraph reading performance. We used simulated phosphene maps with a higher distortion level. In the first experiment, two display methods (the projection and nearest neighbor search methods) were proposed and several levels of parameters (coverage ratio and search range, respectively), which determined the effect of two methods were examined. Next, we examined the effect of distortion on paragraph reading, comparing the regular array correspondence method with the best method identified in Experiment 1. The experimental setups and results were presented separately for ease of interpretation.

EXPERIMENT 1: THE EFFECT OF DIFFERENT DISPLAY METHODS ON RECOGNITION OF CHINESE CHARACTERS

Subjects

Ten volunteers with normal or corrected visual acuity (20/20; five females, five males, 20–30 years of age) were recruited from Shanghai Jiao Tong University. They were all native Chinese speakers and passed a Standard Chinese Proficiency Test. All subjects were informed of the purpose and procedures of the experiments and signed an informed consent form prior to participation. The research adhered to the tenets of the Declaration of Helsinki and ARVO.

Experimental Setup

The test platform consisted of a personal computer (2.90 GHz, Intel Core i5-2310 CPU, 4 GB DDR3 RAM; Lenovo, Inc., Beijing, China), a 17-inch CRT screen (1024 × 768, 85 Hz/110 MHz; EMC; Proview Technology Co., WuHan, China), an eye tracker (ViewPoint EyeTracker; Arrington Research, Inc., Scottsdale, AZ) used to monitor eye movements, a light-shielding device (40 × 40 × 50 cm) to avoid the influence of extraneous light around the screen, a headrest to maintain head position and viewing distance, and a microphone.

The computer ran in-house software Visual C++ (Microsoft, Redmond, WA) to control the experimental procedure, transform the CCs to the simulated irregular phosphnene map, display the phosphenes, and record the recognition time

and verbal responses of the subjects. The CRT screen was connected to the computer via a Video Graphics Adapter distribution board, which presented the visual stimuli (pixelized CCs) to the subjects.

Presentation Materials

A CC library was chosen based on frequency statistics for commonly used modern characters issued by the State Language Work Committee and the State Education Commission.[18] The first 500 characters from the statistical table were selected to minimize any effects due to unfamiliarity. CCs in this set of 500 have 1 to 16 strokes and frequency of use ranging from 0.042% to 3.593%.

Visual Stimuli

All CC recognition tasks were performed using a 12 × 12 dot array where the recognition accuracy of the pixelized CCs reached 100%.[19] The distortion of the dot array was simulated by a 2D Gaussian distribution with 0 mean and mutual independence (Formulas 1 and 2), as was adopted in our previous study.[17]

$$P(\Delta x, \Delta y) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{\Delta x^2 + \Delta y^2}{2\sigma^2}\right) \qquad (1)$$

$$\sigma = kS \qquad (2)$$

where Δx and Δy are the deviations from the regular matrix in the horizontal and vertical directions, and σ is the SD deciding the degree of positional uncertainty of phosphenes. The irregularity index k measures positional uncertainty and S is the center-to-center distance between two neighboring dots in the array. The previous study indicated that recognition accuracy of CCs decreased to <60% at $k = 0.3$ and 0.4; thus we adopted different display methods of CCs to improve the recognition accuracy at these two distortion levels.

Since the gray level did not significantly affect the recognition of CCs,[17] the visual stimuli were represented only as Gaussian spots[20] with 256 of the center gray value on a black background to simulate phosphenes perceived by a prosthesis wearer. The diameter of each Gaussian spot measured at 95% intensity was 33 arcmin, with a 23 arcmin gap between neighboring spots. The visual field (or the size of pixelized CC) was set to 10 to 11°, which is the optimal angle for CC recognition.[21]

CC Display Method 1: The Projection Method

The *projection method* superimposes a CC image onto a simulated irregular phosphene array of the same size and then selects those phosphenes that are covered by the character strokes, thus forming a pixelized CC (Fig. 1).

$$P = \begin{cases} 1 & \frac{A(P \cap W)}{A(P)} \geq TH \\ 0 & otherwise \end{cases} \qquad (3)$$

where P indicates the phosphenes forming a character for the simulated prosthetic vision, $A(P)$ represents the phosphene area, $A(P \cap W)$ represents the phosphene area covered by the character stroke line W, and coverage ratio $A(P \cap W)/A(P)$ represents the ratio between these two areas. Because the stroke line width of a character used the Hei typeface, the value of TH in the formula was a key point in the projection method. A TH value that was too low produced a number of phosphenes that seriously deviated from their regular location
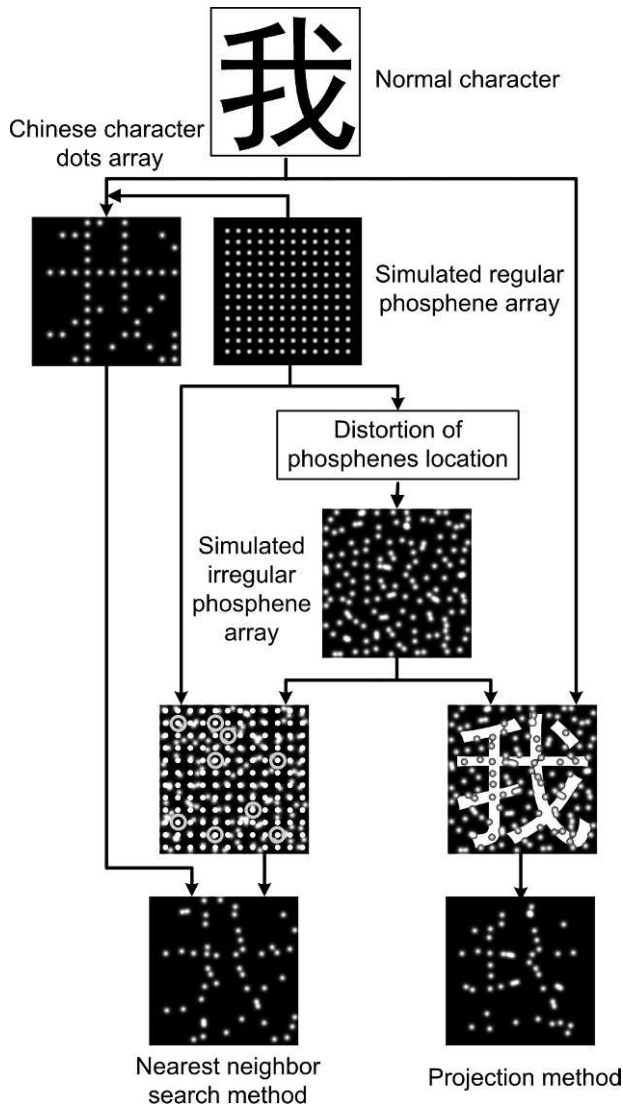
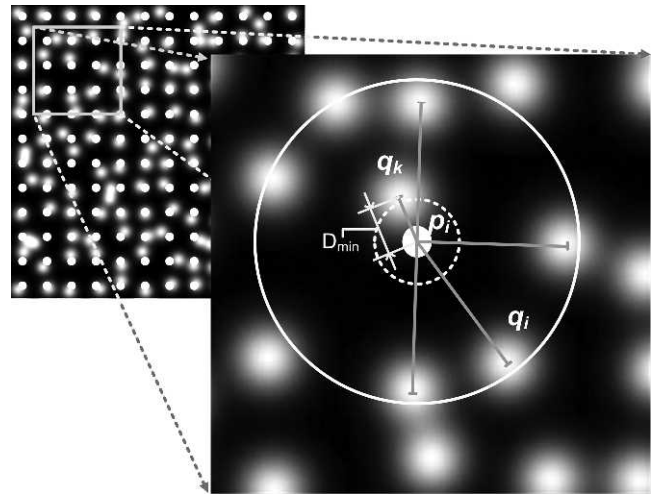FIGURE 1. Flow chart showing two display methods of Chinese characters using simulated irregular phosphene maps.

FIGURE 2. Nearest neighbor search method: p_i is the ideal phosphene location elicited by the electrode, whereas the electrode (e_i) actually elicits phosphenes at q_i with an offset from the regular grid. Therefore, in the appropriate search range, the dot (q_k) elicited by another electrode (e_k) is closer to p_i and thus replaces q_i to express the visual information. If there are no dots within the search range (*dashed circle*), the information at that location is missing.

The CC display method using NNS is a modification of a regular array correspondence method.[17] For each electrode (e_i) of the electrode array (E), $e_i \in E$ and p_i ($p_i \in P$, where P is an ideal regular phosphene array) is the ideal phosphene at a regular location elicited by e_i, whereas e_i actually elicits the phosphene q_i ($q_i \in Q$, where Q is the actual irregular phosphene array) with an offset from the regular grid. Therefore, in the appropriate search range (a circle with center at p_i), the dot q_k elicited by the electrode (e_k) is closest to p_i and thus replaces p_i to express the visual information at this location (Fig. 2). If there were no dots within the search range, the information at that location is missing. Thus, we can define a transformation table between the (regular) electrode array and (irregular) phosphene array in which no electrode is represented by more than one phosphene, and some electrode may be not presented, due to distance or unavailability of a matching phosphene.

When displaying a CC, the character is pixelized according to the ideal phosphene array (regular array) using the GB2312-80 code (Chinese National Standard) to get an expected dot set and then the lookup table is consulted to get the actual phosphene set used to display the character in the irregular phosphene map.

Similar to the projection method, the key point of the NNS was to choose an appropriate search range that was neither too small, which caused a large number of dot dropouts, nor too large, which retained too many distorted dots and destroyed the structure of the CC. Six different radii were chosen for the search range, starting from $0.5S$, where dot dropout was >23%, to $1.0S$ (neighboring site) in $0.1S$ increments. Twelve character libraries (two distortion levels × six search ranges) with 500 characters per library were formed; 100 CCs were randomly selected from each library and used in each trial (Fig. 3).

Procedure

A training test period was performed to help familiarize the subject with the experimental environment and procedure. The training materials consisted of 200 CCs randomly selected from the CC libraries of different simulated phosphene maps. The subject's left eye was covered with an eye patch to simulate a monocular prosthesis and the subjects were asked

and formed a highly distorted character that consumed more energy when produced by the prosthetic device. A high *TH* value (such as *TH* = 1) meant that fewer phosphenes formed a character and some key strokes were missing. In the present study, six *TH* values (1/6 to 6/6 in 1/6 increments) were tested to find the optimal value during the presentation of simulated phosphene maps with two different distortion levels ($k = 0.3$ and 0.4). Twelve CC libraries (two distortion levels × 6 *TH* value) with 500 CCs per library were formed; 100 CCs were randomly selected from each library and used in each trial (Fig. 1; also see Fig. 3 later in the text).

CC Display Method 2: NNS Method

The nearest neighbor search (NNS) method was an optimization problem for finding the nearest points in metric space.[22] It was defined as follows:

Given a data set of points (S) in a metric space (M) and a query point ($q \in M$) find the closest point, p', in S to q. M is d-dimensional Euclidean space.

$$NN(q) = \{p' \in S | \forall p \in S : ||p' - q|| \leq ||p - q||\} \quad (4)$$
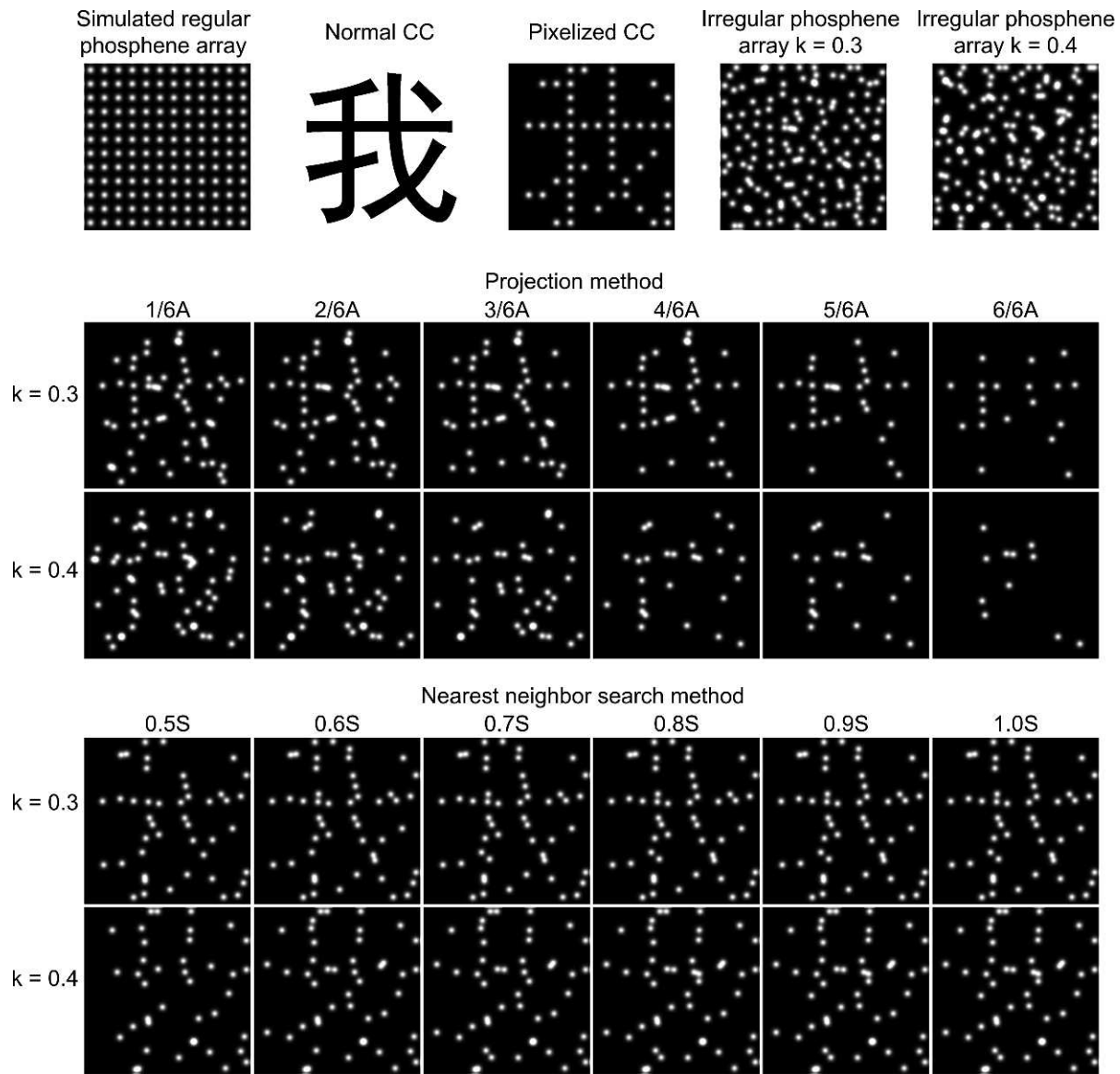
FIGURE 3. Examples of simulated phosphene maps of CCs formed from the projection and nearest neighbor search methods using 0.3 and 0.4 distortion levels (k). Large coverage and small search radius yield the sparsest image.

to read aloud a recognized CC as soon as possible or respond with "pass" for nonrecognition. Each CC was presented for 4 seconds, but subjects could move on to the next CC by pressing the keyboard if they completed the task in less time.

The verbal responses of each subject were recorded and analyzed for recognition accuracy. The data points represented the mean number of correct responses (% ± SD) for combined data from all subjects. Data were analyzed with two-way ANOVA to examine the effect of the parameter (coverage ratio or search range) and distortion level on the CC recognition (a Bonferroni correction was applied to multiple comparisons; SPSS 16.0 for Windows; IBM SPSS Inc., Armonk, NY) and paired t-test (two-tailed) to compare the two methods.

RESULTS

Projection Method

CC recognition accuracy significantly changed with an increase in the coverage ratio for both distortion levels when using the projection method (Fig. 4A) and reached a peak at a ratio of 3/6 ($k = 0.3$: 62.4 ± 9.8% and $k = 0.4$: 47.0 ± 9.7%). There was no significant interaction effect between distortion and coverage ratio. The distortion level had a significant impact on recognition accuracy, which at a lower distortion level was higher for each coverage ratio ($P < 0.05$). Recognition accuracy obtained for CCs at the optimal coverage ratio (3/6) was compared with the recognition of CCs at other coverage ratios. According to the multiple comparisons, the accuracy at a coverage ratio 3/6 was significantly higher than that at other coverage ratios (5/6 and 6/6; $P < 0.05$ after Bonferroni correction) for both distortion levels.

Nearest Neighbor Search Method

Recognition accuracy using the NNS method reached a peak when using a search range of $0.6S$ (81.3 ± 2.7%) and a $k = 0.3$ distortion level (Fig. 4B). With the exception of the $0.5S$ search range (57.5 ± 5.7%), recognition accuracy at other search ranges ($0.6S$–$1.0S$) was significantly higher compared with the
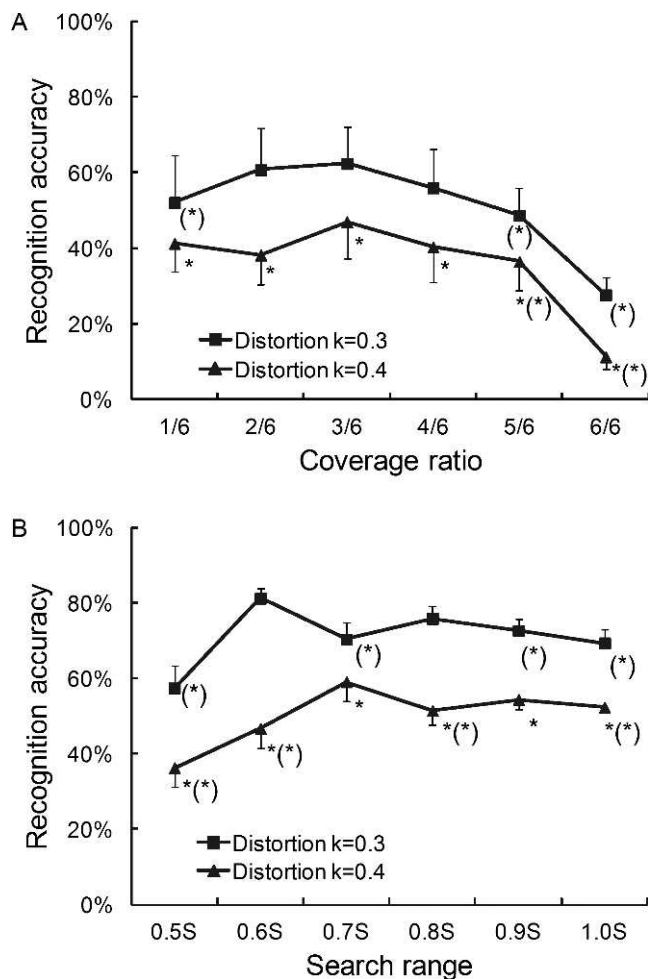
FIGURE 4. Recognition accuracy of CCs as a function of (A) the coverage ratio using the projection method and (B) different search ranges using the nearest neighbor search method at different distortion levels (k). *Error bar* represents the variability of the means of the 10 subjects. Accuracy (mean ± SD) obtained for CCs with the optimal level of the parameters in the different methods (coverage ratio: 3/6, search range: 0.6S and 0.7S at $k = 0.3$ and 0.4 distortion levels, respectively) at each distortion level was compared with the accuracy at other levels (significance shown in parentheses, *$P < 0.05$). Significant comparisons between recognition accuracy at different k levels with the same parameter level are indicated without parentheses.

recognition at other search ranges. When $k = 0.3$, recognition at 0.6S was significantly higher than that at other search ranges, with the exception of 0.8S, whereas at the higher distortion level, except for 0.9S, there were significant differences between other search ranges and 0.7S ($P < 0.05$ after Bonferroni correction). At each search range, the higher distortion level resulted in lower recognition accuracy.

EXPERIMENT 2: EFFECT OF DISPLAY METHOD ON PARAGRAPH READING

Experimental Setup

Ten volunteers with normal or corrected visual acuity of 20/20 (four females, six males, 20–30 years of age), who did not participate in Experiment 1, were recruited for the second experiment. The test platform consisted of two personal computers (2.90 GHz, Intel Core i5-2310 CPU, 4 GB DDR3 RAM; Lenovo, Inc.), a 17-inch CRT screen (1024 × 768, 85 Hz/ 110 MHz; EMC; Proview Technology Co., WuHan, China), a head-mounted display (HMD; Z800 3Dvsior, 800 × 600 resolution, 40° diagonal field of vision; eMagine, Rochester Hills, MI), a microcamera (640 × 480, 30 Hz; Philips Inc., Best, The Netherlands), an eye tracker (View Point EyeTracker; Arrington Research, Inc.), headrest, and a microphone.

One computer was used to show a normal paragraph containing CCs. The other computer ran in-house software written in Visual C++ (Microsoft) to process the images captured by the microcamera (640 × 480, 30 Hz; Philips Inc.) and display the pixelized paragraphs in the HMD.

Materials

In all, 200 paragraphs were chosen from primary school Chinese language textbooks (grades 4–6; *People's Education Press*, 5th ed., 2008, PR China) to form a paragraph library and 95% of the CCs in the paragraphs were listed among the top 1000 CCs in the frequency of use table[18]; these CCs can provide nearly 92% of daily reading and writing information. All paragraphs had a similar distribution of CC frequency of use and number of strokes, such that, for each paragraph, (1) the number of CCs was between 40 and 44, (2) the average number of strokes per CC was between 6.6 and 7.6, and (3) the average CC frequency of use was between 0.31% and 0.47%. Each paragraph on the computer screen was formed by five lines with nine characters per line.

Processing Paragraph Images

Image Preprocessing. Because the optimal size necessary for accurate pixelized reading was 5° × 5°[23,24] and the resolution of a single CC was set to 12 × 12, 3 × 3 CCs with a 36 × 36 dot array were presented in the 15° × 15° visual field. To achieve consistency between the visual field taken from the camera (640 × 480, 30 Hz; Philips Inc.) and that presented to subjects, pictures from the camera were cropped to a 15° × 15° visual angle and denoised (Fig. 5).

Unlike the CC processing method according to the GB 2312-80 Code in Experiment 1, which required real-time and accurate character recognition of computers, another CC processing method was reduced sampling of the CC image, which was a real-time process, but in some cases may divide a thick stroke into two parallel lines, thus causing incorrect display and recognition of the CC. Therefore, we selected a parallel thinning algorithm[25] to thin the stroke thickness before the reduced sampling used to optimize the display.

Paragraph Display Method 1: Regular Array Corre-

maximum achieved using the projection method at the same distortion level (81.3 ± 2.7%, 70.4 ± 4.3%, 75.9 ± 3.0%, 72.7 ± 2.9%, 69.4 ± 3.4% vs. 62.4 ± 9.8%; paired two-tailed *t*-test, $P < 0.05$, respectively). Unlike the condition for the $k = 0.3$ distortion level, recognition accuracy with the NNS method reached a maximum at a search range of 0.7S for the $k = 0.4$ distortion level, meaning that the optimal search range changed as the distortion level increased. Statistical analysis indicated that this maximum was significantly higher than the maximum using the projection method (59.1 ± 5.2% vs. 47.0 ± 9.7%, paired two-tailed *t*-test, $P < 0.05$); the accuracies at search ranges between 0.8S and 1.0S were all higher than the maximum when using the projection method at the same distortion, albeit not significantly.

The search range of the NNS method had a significant impact on CCs recognition ($P < 0.05$) for both distortion levels. Recognition at an optimal search range ($k = 0.3$: 0.6S and $k = 0.4$: 0.7S) at each distortion level was compared with
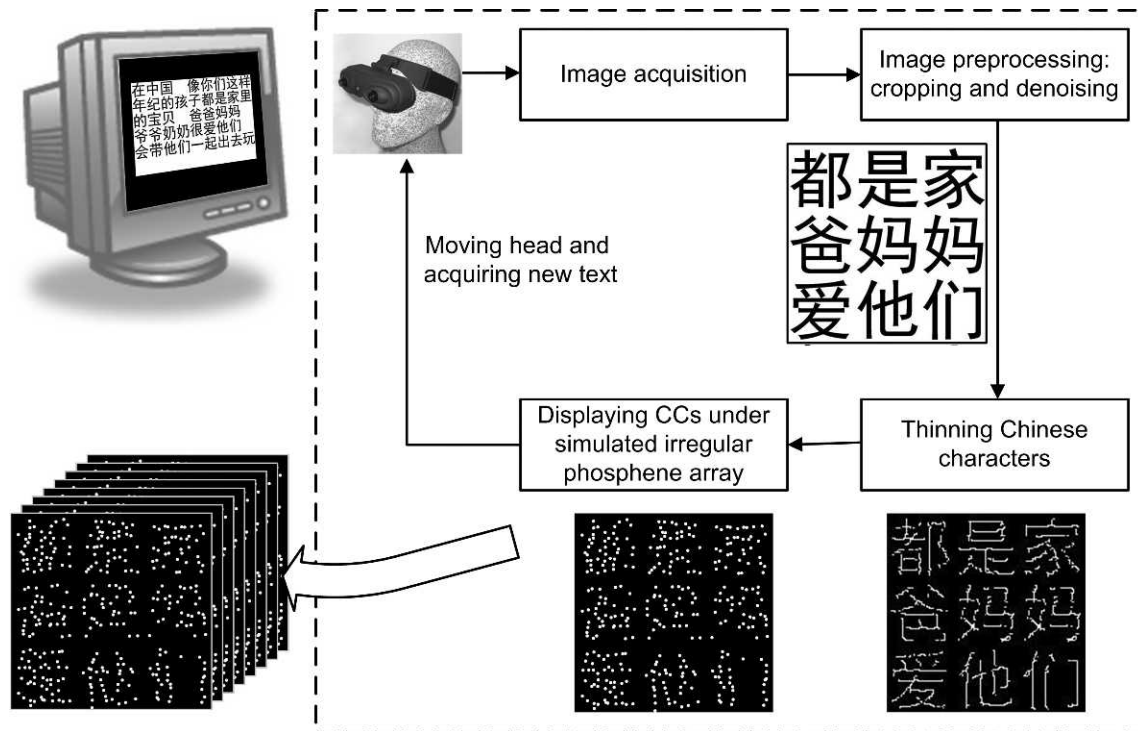
FIGURE 5. Flow chart of processing steps for paragraph reading.

spondence Method. The regular array correspondence method was adopted in the previous study. Normal CCs images were pixelized according to a regular 36×36 dot array corresponding to the regular array and the position of each dot was distorted according to Formulas 1 and 2. The more the locations of the phosphenes elicited by the electrodes deviated from the expected locations in the visual map, the more distorted were the CCs formed by the phosphenes. Because reading accuracies at $k = 0.1$ and 0.2 distortion levels were near 100% in the pretest, four blocks in the formal experiment were used with distortion indexes of $k = 0.3$, 0.4, 0.5, and 0.6 to avoid a ceiling effect. Twenty different paragraphs were randomly chosen and presented in four blocks (five paragraphs for each block).

Paragraph Display Method 2: Nearest Neighbor Search Method. According to the results of the regular array correspondence method, the NNS method was used with distortion levels of 0.5 and 0.6, where the recognition accuracy was <80%. In the CC recognition experiment, the recognition accuracy was not optimal for a search range larger than $0.8S$, where the dots were more suitable for the substitution of a neighboring ideal phosphene. Therefore, the search range was set to $0.5S$ to $0.8S$ in $0.1S$ incremental steps. Forty different paragraphs were randomly chosen (five paragraphs for each search range $\times$ four search ranges $\times$ two distortion levels).

Procedure

Prior training was performed to help subjects become familiar with the experimental environment and procedure. The training materials were 10 paragraphs randomly selected from the 200 paragraphs representing different simulated phosphene maps. The training and formal test used different paragraphs, but followed the same format: each subject wore the HMD, and moved his or her head to capture different parts of the normal text presented on the CRT screen and read aloud what was seen on the HMD. After the subject finished one

paragraph, the experimenter clicked the mouse to record the time and the next paragraph was presented. The simulated phosphene map remained stable when reading one paragraph, whereas between different paragraphs, the phosphene map changed. All paragraphs were presented in a pseudorandom, counterbalanced order to evenly distribute the learning effect. Subjects were allowed a 5-minute rest after finishing five paragraphs.

The verbal responses of each subject were recorded and analyzed for reading accuracy (RA $= CC_{Correct}/CC_{Total}$) and reading efficiency (RE $= CC_{Correct}/T_{Total}$, where T_{Total} denotes the total time to read one paragraph). Data were analyzed with one-way ANOVA (a Bonferroni correction was applied to multiple comparisons) to examine the effect of distortion levels using regular array correspondence method or search range using the NNS method and paired t-test (two-tailed) to compare two methods at the same distortion level.

RESULTS

Regular Array Correspondence Method

The RA of paragraphs containing CCs significantly decreased with an increase in the distortion level ($P < 0.05$). The RA was $98.4 \pm 1.2\%$ at a $k = 0.3$ distortion level, nearly 100% and still >85% even at $k = 0.4$. The largest RA decrease appeared to be between distortion levels $k = 0.5$ ($68.0 \pm 7.7\%$) and 0.6 ($38.5 \pm 6.7\%$). Figure 6A also showed that there were significant RA differences between different distortion levels ($P < 0.05$ after Bonferroni correction).

RE was also significantly affected by the distortion level of the simulated phosphene maps ($P < 0.05$), especially at a level where $k = 0.3$ to 0.4 for which RE decreased from 55.7 ± 13.2 char/min to 31.7 ± 13.8 char/min (Fig. 6A). As the distortion level increased, RE decreased monotonically to <15 char/min at a distortion level of 0.6. There were no significant RE
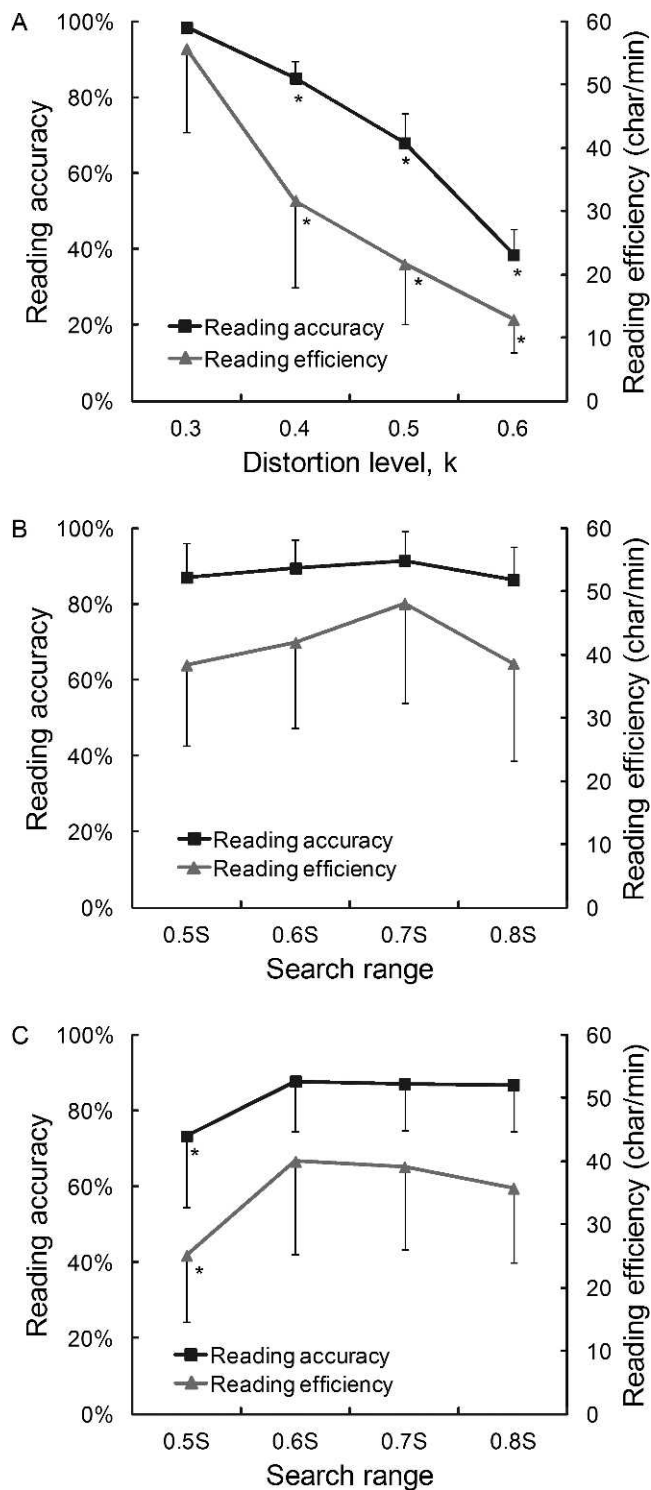
FIGURE 6. Reading accuracy and efficiency of paragraphs containing CCs were plotted as a function of (**A**) distortion level using the regular array correspondence method and (**B**, **C**) search range using the nearest neighbor search method ([**B**]: $k = 0.5$ distortion level; [**C**]: $k = 0.6$). *Error bar* represents the variability of the means of the 10 subjects. Accuracy or efficiency values obtained for paragraphs at the lowest k level (**A**) or the optimal search range ([**B**]: 0.7S at $k = 0.5$; [**C**]: 0.6S at $k = 0.6$) were compared with values obtained at other levels (*$P < 0.05$).

differences between level $k = 0.5$ and its neighboring ranges (0.4 and 0.6).

Nearest Neighbor Search

The paragraph RA using the NNS method exceeded 80% for all search ranges when using a $k = 0.5$ distortion level (Fig. 6B) and was significantly higher (paired two-tailed t-test, $P < 0.05$) compared with the regular array correspondence method (68.0 $\pm$ 7.7%). As the search range increased from 0.5S to 0.7S, RA increased from 87.1 $\pm$ 8.9% to 91.5 $\pm$ 7.6%, and then decreased to 86.4 $\pm$ 8.7% as the search range increased to 0.8S. However, statistical analysis indicated that there were no significant differences between different search ranges ($P >$ 0.05). RE trends following an increase in the search range were similar to that seen for RA, and reached a peak for a 0.7S search range (48.1 $\pm$ 15.5 char/min); similarly, there were no significant differences between different search ranges ($P >$ 0.05). REs when using a search range between 0.5S and 0.8S and the NNS method were all >38 char/min and significantly higher (paired two-tailed t-test, $P < 0.05$) compared with the regular array correspondence method (21.7 $\pm$ 13.7 char/min).

For a $k = 0.6$ distortion level of the simulated phosphene array, with the exception of the 0.5S search range (73.3 $\pm$ 18.6%), RAs using the NNS method exceeded 80%, and were significantly higher (paired two-tailed t-test, $P < 0.05$; Fig. 6C) relative to RA when reading paragraphs manipulated by the regular array correspondence method (38.5 $\pm$ 6.7%). RA reached a peak (87.8 $\pm$ 13.2%) with a search range of 0.6S, unlike a $k = 0.5$ distortion level, and then slightly decreased to 86.8 $\pm$ 12.2% as the search range increased to 0.8S. RA was only significantly lowered with a search range of 0.5S ($P <$ 0.05). A similar trend was seen in RE values, which reached a peak (40.0 $\pm$ 14.6 char/min) with a search range of 0.6S and showed significant differences only between 0.5S and other search ranges ($P < 0.05$). The RE values for four search ranges using the NNS method were all significantly higher than that using the regular array correspondence method (paired two-tailed t-test, $P < 0.05$).

DISCUSSION

Effect of Different Displaying Methods

Chinese characters were displayed according to either a projection or a nearest neighbor search method, and differed from the display method corresponding to the regular electrode array in our previous study.[17] With a less distorted phosphene array ($k = 0.1$ or 0.2 for recognition of characters and $k = 0.3$ or 0.4 for reading paragraphs), the characters produced by the regular array correspondence method of the previous study resulted in better performance (>85% recognition accuracy); with a more distorted phosphene array, the recognition accuracy using the projection or nearest neighbor search methods was higher compared with the regular array correspondence method. Under these conditions the nearest neighbor search method significantly improved the performance of CC recognition compared with the performance following the projection method. Therefore, we can choose the optimal display method in terms of the distortion level of the phosphene map to benefit the perception of the prosthesis wearer.

After the implantation, researchers can test the phosphene array perceived by a prosthesis recipient and then map the ideal phosphene map according to the electrode array. Comparison of the projection and NNS methods in this study showed that NNS was to find a closest substitute for the

expected dot and the dot grid using NNS was a better approximation of the regular position and could present more regular lines or contours, in spite of some dropout. Since CCs are formed by regular lines and recognized by their contour information,[26] the NNS method may be a suitable strategy to make CCs under irregular phopshene array more recognizable. It is noted that the distortion model in this study was a 2D Gaussian function and the distribution of simulated phosphenes in the visual field had no obvious orientation bias. If the phosphene distribution elicited by the visual prosthesis has an orientation bias (i.e., most phosphenes are compressed in one orientation, with more dispersion in the perpendicular orientation), there would be a more serious effect of the projection method on the representation of images.

Effects of Parameters for the Two Methods

The results revealed that the coverage ratio, when using the projection method, or the search range, when using the NNS method, had a significant impact on the recognition of pixelized CCs. Changing the coverage ratio or search range meant changing the number of the dots that formed the strokes of pixelized CCs. When using a large coverage area or smaller search range, the strokes of the CCs appeared to be incomplete or missing, resulting in an unrecognizable CC, which was consistent with the results of our previous study, which showed that the phosphene dropout rate in the array significantly affected the recognition of pixelized CCs[17]; however, in spite of more information, the use of a smaller coverage area or a large search range resulted in distorted strokes, thereby disrupting the whole character structure. Under these conditions the addition of more dots interfered with the recognition of the "*character's radical*," which was an important process in CC recognition.[27,28] Therefore, a parameter value for which recognition accuracy reached a peak is the best compromise between the dropout and distortion of the phosphene array.

Interestingly, the results with the NNS method showed that CC recognition accuracy reached a peak and, although slightly lower, accuracy tended to plateau as a function of the search range. Expanding the search range added more distorted dot positions to the array. In the study, the search range linearly increased, but resulted in a nonlinear increase in information and the distortion level of the array. Clearly, a search range between $0.7S$ and $0.8S$ (e.g., $k = 0.3$ distortion level) increased the number of dots (average approximately 4) and had a greater impact on the increase in information than the distortion level. This was in agreement with the subjective feedback from the subjects who reported that CCs with a $0.8S$ had clearer strokes compared with a search range of $0.7S$.

Contrary to expectations, the search ranges of the NNS method had only a slight impact on paragraph RA and indicated that phosphene dropout had less influence. Unlike recognizing static CC images processed off-line, when reading paragraphs in real time, subjects could increase spatial frequency through temporal integration and get more information through head movement to reduce the effect of dropout. Fornos et al.[20] compared the effect of off-line versus real-time pixelization on reading with simulated prosthetic vision, for words written in the Latin alphabet, and found that real-time stimulus pixelization favored reading performance. Clinical studies have also demonstrated the effect of head movement by which patients with a retinal prosthesis device (60 electrodes; Argus II Retinal Prosthesis System; Second Sight Medical Products, Inc., Sylmar, CA) can read large single letters (daCruz L, et al. *IOVS* 2010;51:ARVO E-Abstract 2023). Nevertheless, head movement requires more time for recognition. In the present study, some subjects had more difficulty

reading through the use of head movement and needed more time to become familiar with the experimental procedure and this caused a large between-subject variability in RE.

The Effect of the Context

Although 90% of the characters in the paragraphs were among the 1000 most commonly used characters, in contrast to Experiment 1 and our previous study, which used the most commonly used 500 CCs, paragraph reading had a much better performance relative to single character recognition with an equivalent or lower distorted phosphene map. This result can be attributed to the linguistic context for reading, which was consistent with the findings of Biemiller in a study of reading English[29] and Zhao et al.[23] in a study of reading Chinese paragraphs by subjects using simulated prosthetic vision. The subjects responded that some distorted characters could not be recognized, but they were able to guess the meaning due to the context of the text, for example, the association of commonly used words. This finding is important for prosthetic vision users in that it suggests that prosthesis wearers can increase reading capacity through linguistic context, even though there is a limited number of phosphenes and a distorted map.

The Effect of Distortion

Clinical trials of visual prostheses reported a lack of retinotopic correspondence between the stimulating site and the perceived location in the visual field (i.e., the phosphene map appeared distorted with respect to the corresponding electrode array).[4,8] Our results indicate that regardless of the CC display method, distortion has a significant impact on the performance of Chinese character recognition and paragraph reading. A similar trend has been demonstrated in CC recognition[17] and object recognition (Zhou C, et al. *IOVS* 2010;51:ARVO E-Abstract 3030). Although paragraph RE at higher distortion levels was still poor and unacceptable for reading, RA and the accuracy of character recognition were markedly improved using the NNS method compared with the regular array correspondence method. Our results demonstrate the importance of optimizing CC display methods and suggest that, in addition to more effective image processing and encoding strategies, we should consider optimizing the design of the prosthetic device (e.g., the distribution of electrodes[30,31]) to keep the impact of the distortion to a minimum and optimize the perceptions for maximum benefit.

Limitation of the Study

The simulation in our experiments depended on a highly idealized map. The phosphene was mimicked by a round spot with a Gaussian distribution, whereas other forms of phosphenes elicited by visual prostheses have also been reported in clinical trials.[33] The geometric irregularity of the phosphene map was simulated by a spatial probability distribution function, which focused on the local randomness of the phosphene location. The asymmetry of the phosphene distribution or the global geometric deformation should receive more attention in future studies.

In conclusion, due to the lack of retinotopic correspondence between the stimulation site of a visual prosthesis and the perceived location in the visual field, we proposed several display methods and estimated their effect on CC recognition and paragraph reading. The results indicated that a nearest neighbor search method significantly improved the performance of character recognition and paragraph reading under conditions where the simulated phosphene maps were more

distorted; the optimal search range achieving maximum accuracy of recognition of CCs or Chinese reading changed according to different distortion levels. Paragraph reading performance decreased with an increase in the level of distortion, similar to that observed with single character recognition, but was less sensitive to the distortion of the phosphene map. We expect the methods in the study can be used or adapted for a visual prosthesis to enhance Chinese character reading ability of blind patients.

References

1. Weiland JD, Humayun MS. Visual prosthesis. *Proc IEEE*. 2008;96:1076–1084.

2. Brindley GS, Lewin WS. The sensations produced by electrical stimulation of the visual cortex. *J Physiol*. 1968;196:479–493.

3. Veraart C, Raftopoulos C, Mortimer JT, et al. Visual sensations produced by optic nerve stimulation using an implanted self-sizing spiral cuff electrode. *Brain Res*. 1998;813:181–186.

4. Humayun MS, Weiland JD, Fujii GY, et al. Visual perception in a blind subject with a chronic microelectronic retinal prosthesis. *Vision Res*. 2003;43:2573–2581.

5. Fujikado T, Kamei M, Sakaguchi H, et al. Testing of semichronically implanted retinal prosthesis by suprachoroidal-transretinal stimulation in patients with retinitis pigmentosa. *Invest Ophthalmol Vis Sci*. 2011;52:4726–4733.

6. Palanker D, Vankov A, Huie P, Baccus S. Design of a high-resolution optoelectronic retinal prosthesis. *J Neural Eng*. 2005;2:S105–S120.

7. Curcio CA, Allen KA. Topography of ganglion cells in human retina. *J Comp Neurol*. 1990;300:5–25.

8. Rizzo JF III, Wyatt J, Loewenstein J, Kelly S, Shire D. Perceptual efficacy of electrical stimulation of human retina with a microelectrode array during short-term surgical trials. *Invest Ophthalmol Vis Sci*. 2003;44:5362–5369.

9. Dagnelie G. Psychophysical evaluation for visual prosthesis. *Annu Rev Biomed Eng*. 2008;10:339–368.

10. Dagnelie G, Barnett D, Humayun MS, Thompson RW Jr. Paragraph text reading using a pixelized prosthetic vision simulator: parameter dependence and task learning in free-viewing conditions. *Invest Ophthalmol Vis Sci*. 2006;47:1241–1250.

11. Dagnelie G, Keane P, Narla V, Yang L, Weiland J, Humayun M. Real and virtual mobility performance in simulated prosthetic vision. *J Neural Eng*. 2007;4:S92–S101.

12. Dagnelie G, Walter M, Yang LC. Playing checkers: detection and eye-hand coordination in simulated prosthetic vision. *J Modern Optics*. 2006;53:1325–1342.

13. Hayes JS, Yin VT, Piyathaisere D, Weiland JD, Humayun MS, Dagnelie G. Visually guided performance of simple tasks using simulated prosthetic vision. *Artif Organs*. 2003;27:1016–1028.

14. Thompson RW Jr, Barnett GD, Humayun MS, Dagnelie G. Facial recognition using simulated prosthetic pixelized vision. *Invest Ophthalmol Vis Sci*. 2003;44:5035–5042.

15. Hallum LE, Cloherty SL, Taubman DS, Suaning GJ, Lovell NH. Psychophysics of prosthetic vision: III. stochastic rendering, the phosphene image, and perception. *Conf Proc IEEE Eng Med Biol Soc*. 2006;1:1169–1172.

16. Cai S, Fu L, Zhang H, Hu G, Liang Z. Prosthetic visual acuity in irregular phosphene arrays under two down-sampling schemes: a simulation study. *Conf Proc IEEE Eng Med Biol Soc*. 2005;5:5223–5226.

17. Zhao Y, Lu Y, Zhou C, Chen Y, Ren Q, Chai X. Chinese character recognition using simulated phosphene maps. *Invest Ophthalmol Vis Sci*. 2011;52:3404–3412.

18. *Modern Chinese Characters Commonly Used in the Frequency Statistics*. Beijing: The State Language Work Committee; 1989.

19. Chai X, Yu W, Wang J, Zhao Y, Cai C, Ren Q. Recognition of pixelized Chinese characters using simulated prosthetic vision. *Artif Organs*. 2007;31:175–182.

20. Fornos AP, Sommerhalder J, Rappaz B, Safran AB, Pelizzone M. Simulation of artificial vision, III: do the spatial or temporal characteristics of stimulus pixelization really matter? *Invest Ophthalmol Vis Sci*. 2005;46:3906–3912.

21. Huang X, Cai Z, Chen L. The effect of visual angle on the recognition of Chinese characters. *Psychol Sci*. 2004;27:770–773.

22. Knuth DE. *The Art of Computer Programming*. Reading, MA: Addison-Wesley; 2011.

23. Zhao Y, Lu Y, Zhao J, et al. Reading pixelized paragraphs of Chinese characters using simulated prosthetic vision. *Invest Ophthalmol Vis Sci*. 2011;52:5987–5994.

24. Feng H, Lu Q, Liu Y. Proper visual angle, dimension and number of figure. *Chin J Ergonom*. 2000;6:25–28.

25. Zhang TY, Suen CY. A fast parallel algorithm for thinning digital patterns. *Commun ACM*. 1984;27:236–239.

26. Tseng S-C, Chang L-H, Wang C-C. An informational analysis of the Chinese language: I. The reconstruction of the removed strokes of the ideograms in printed sentence-texts. *Acta Psychol Sinica*. 1965;4:281–290.

27. Ding G, Peng D, Taft M. The nature of the mental representation of radicals in Chinese: a priming study. *J Exp Psychol Learn Mem Cogn*. 2004;30:530–539.

28. Yeh SL, Li JL. Role of structure and component in judgments of visual similarity of Chinese characters. *J Exp Psychol Hum Percept Perform*. 2002;28:933–947.

29. Biemiller A. Relationships between oral reading rates for letters, words, and simple text in the development of reading achievement. *Reading Res Q*. 1977-1978;13:223–253.

30. Chen SC, Hallum LE, Lovell NH, Suaning GJ. Visual acuity measurement of prosthetic vision: a virtual-reality simulation study. *J Neural Eng*. 2005;2:S135–S145.

31. Rodger DC, Fong AJ, Li W, et al. Flexible parylene-based multielectrode array technology for high-density neural stimulation and recording. *Sensors Actuators B Chemical*. 2008;132:449–460.

32. Chen SC, Suaning GJ, Morley JW, Lovell NH. Simulating prosthetic vision: I. Visual models of phosphenes. *Vision Res*. 2009;49:1493–1506.