

# Image processing strategy used for simulated phosphene map of artificial vision

Longmin Chen, Fanghua Yan, Zezhi Gong, Yayu Zheng<sup>1\*</sup>

<sup>1</sup>College of Information Engineering, Zhejiang University of Technology, No.288 Liuhe Road, Hangzhou, China

Received 24 November 2014, www.cmnt.lv

## Abstract

Phosphene generating mechanism is the most important theoretical foundation and key technology for artificial vision. Simulating the mapping relationship between visual image information and limited phosphene maps remains a difficult problem. To satisfy the practical requirement of clinical trials, we present an image processing strategy for simulated phosphene map. Based on block segmentation, a reducing pixelated image processing method is proposed. Meanwhile, an electrode intensity control strategy based on brightness grading is carried out as well. Finally spatial response experiment is performed on the artificial vision platform based on DSP to prove the algorithm availability. The experiment indicates that under the premise of ensuring the electrode stimulation accuracy, this system can precisely extract the brightness information of the pixel block and the contour information and transform them into electrode stimulus. The speed of the image processing strategy based on DSP is up to 30 frames per second after being optimized, which meets the real time needs of visual centers system completely.

*Keywords:* artificial vision, phosphene map, reducing pixelated, DSP

## 1 Introduction

Artificial vision is also named visual prostheses. The basic principle of artificial vision is converting visual information into electrical signals by artificial means, hereby imposing electrical stimulation on microelectrodes which are implanted in the visual pathway (retina, visual cortex or optic nerve) to induce phosphene, then restore the visual perception of the blind. The system structure of information processing of artificial visual image is represented in Figure 1. Although some of study results of artificial vision are recognized by numerous research institutes [1-3], current artificial vision researches still face a difficulty. That is, due to the limited number of implanted electrodes, there is no general image processing algorithm to get the minimum valid information from original complex visual image. However, investigations confirm that it is available to use simulated phosphene map to represent image information based on phosphene generating mechanism. This process is called pixelated vision [4, 5].

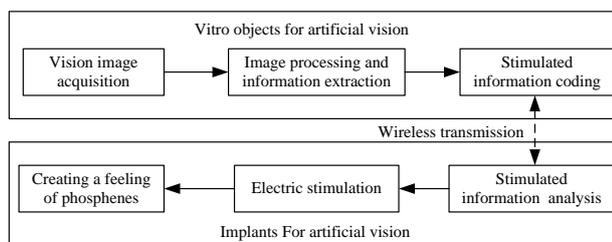


FIGURE 1 System structure of artificial vision

Natural graphics contains immense information, but the information that electrode can express is restrained. How to use limited electrode array to express the mass of natural scene is an important research. The blind need not to recognize the color, motion and stereo vision. Basic aims of visual repair are to restore the visual perception of the blind and further to identify the silhouette. Therefore, image processing for artificial vision focuses on getting information of brightness, contour information and simulated phosphene map to induce the microelectrode stimulation. For decades, scientists have intense interest in artificial vision research. With the development of science and technology, the artificial vision researchers have made enormous progress in recent years. United States Argus II artificial vision system has worked out 30 cases human clinical implanted [6], and the implanted electrode number has reached 60. As a result, patients can recognize simple characters. The natural graph is mapped directly into regular or discrete phosphene maps after single character recognizing arose by the later electric stimulus training constantly. Van Rheede et al [7] have built a real-time system for visual prosthesis and carried out image processing strategy based on ROI and fish eye zooming-in. Different experiments on scenario demonstrates that different visual tasks require different algorithms. ROI algorithm can refine more scene details, but low pixelated algorithm can provide more macro details. Zhao et al [8] have proposed image processing strategy based on DSP. According to image complexity, the image can be divided into 3 parts after being pre-processed. They are simple image, moderate image and complex images. Finally

\* Corresponding author's e-mail: yayuzheng@zjut.edu.cn

288×288 original images can reduce to 32×32 discrete pixel matrixes.

In summary, current visual image processing methods generally remain in simulation and experiment stage. There are few practical applications in medicine domain. For the requirements of specific clinical experiments and the limitation of only 4×6 micro electrode lattice being implanted in ophthalmic hospitals, based on block segmentation and gray grading, a method called reducing pixelated is presented to simulate the mapping relationship between brightness of pixel block information and phosphene map in this paper. Subsequently, an artificial vision system is proposed based on DSP for hardware platform.

## 2 Visual image processing strategy

Given that many image processing methods for artificial vision are still being explored in the laboratory, and the limit of 4×6 number of implant electrodes, we mainly discuss the following image processing strategy: on non-character scenarios, reducing pixelated image processing method based on block segmentation in this paper. The purpose is to build the brightness information of simulated phosphene map and characterize contour information of the object. What's more, to help ophthalmic hospitals further explore imaging mechanism of phosphine is also a significant aim, which can contribute to improving the design of implanted electrode array and the establishment of stimulation strategy.

### 2.1 REDUCING PIXELATED IMAGE PROCESSING METHOD BASED ON BLOCK SEGMENTATION

While only 4×6 number of electrodes can be implanted, a reducing pixelated image processing method based on block segmentation is designed as described in Figure 2.

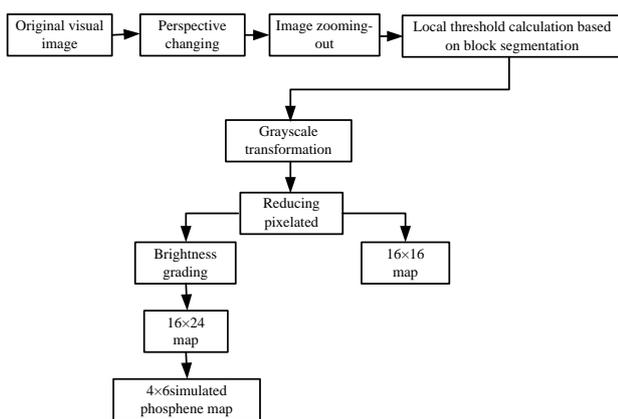


FIGURE 2 Flowchart of the reducing pixelated image processing method based on block segmentation.

#### 1) Perspective changing and image zooming-out.

Ordinarily, the micro-electrode array in retina implants is in the macular area of the eye, corresponding to the central angle of about 20°×20° [9]. We suppose the camera's view angle is about 45°×45° in this paper. The

input image resolution is 640×480. Its central region is tailored to deal with and the original image resolution is set as 240×360 because of the proportion of 4:6. In addition, perspective changing can effectively suppress camera noise introduced by few lights.

The step of image zooming-out is to reduce the 240×360 image to the 96×144 image and ensure the image is not distorted as far as possible. Typical image scaling algorithms include bilinear interpolation, tri-linear interpolation, wavelet interpolation, etc. In our proposed architecture, a specific hardware algorithm called poly-phase filter algorithm is applied. The scaling rate of the algorithm ranges from 0.25× to 4×. Through experiment, processing efficiency of hardware algorithm is found at least 10 times higher than the simplest bi-linear interpolation method.

#### 2) Local threshold calculation based on block segmentation.

Calculating threshold is the key of grayscale transformation. Common methods like fixed threshold and dynamic threshold are employed in many domains. The adaptive dynamic threshold algorithm known as otsu algorithm (OTSU) is an advanced algorithm [10]. This procedure mainly includes the following sections. In one image, the proportion of the number of pixels of the target image in the whole image is denoted by  $\omega_0$ , and average grayscale is denoted by  $\mu_0$ . Similarly, the proportion of the number of pixels of the background and its average grayscale are denoted by  $\omega_1$  and  $\mu_1$ . OTSU algorithm formula is defined as in Equation (1). A traversal method is applied to make interclass variance  $\sigma$  as a max threshold. The  $\sigma$  is the optimal threshold value, denoted by  $T$ .

$$\begin{cases} \omega_0 + \omega_1 = 1 \\ \sigma = \omega_0 \omega_1 (\mu_0 - \mu_1)^2 \end{cases} \quad (1)$$

When in practical use, due to uneven brightness and local highlights, the global Otsu algorithm often mistake part of target image for background. Consequently, an improved Otsu algorithm is proposed. An image processing method based on block segmentation is used to calculate the local threshold, the key of this method is dividing the number of blocks. Dividing too much will affect the efficiency, and too little is not good to subsequent grayscale transformation. After block segmentation, 3 parameters of small rectangular blocks should be calculated respectively in the Otsu algorithm. They are optimal segmenting threshold  $T_n$ , grayscale average of target object and background  $avgln$  and  $avghn$ ,  $n=1,2,\dots,6$ . If grayscale average values satisfy the expression  $(avghn-avgln) \geq \sigma$  ( $\sigma=30$ ,  $n=1,2,\dots,6$ ), this rectangular block is considered as non-flat area, which means there are significant differences between target object and background. Its area is marked as active block. Global threshold  $T$  will eventually be placed by average of all the effective threshold  $T_n$ .

#### 3) Grayscale transformation.

After acquiring global threshold  $T$ , usually binaryzation processing is taken on image. But approach

in this paper is different. Contrast stretching method and grayscale transformation is added in this method as in Equations (2), (3), (4), (5).

$$fA = 144, \tag{2}$$

$$fB = 128 - fA \times T / 16, \tag{3}$$

$$gray_i = fA \times i / 16 + fB, \tag{4}$$

$$gray[i] = \begin{cases} 255 & \text{if } (gray \geq 255) \\ 0 & \text{else if } (gray < 0) \\ gray_i & \text{else} \end{cases} \tag{5}$$

In the above equation  $i=0,1,\dots,255$ . Traversed 8-bit grayscales are converted to grayi by contrast stretching, then table  $gray[255]$  is updated. If one pixel in the image grayscale value is p, then the original grayscale value is replaced by  $gray[P]$ , so that the grayscale conversion will be carried out. For YUV422 format image, the chrominance components U and V are set to 0x80, which lead to a valid segmentation of target object and background. The image looks smoother and more noises are removed.

4) Generating 4x6 simulated phosphene map.

At first, an average grayscale value denoted by  $P_{avg}$  for each block is added up after dividing converted grayscale images into several small rectangular blocks of  $M \times N$ . Binaryzation processing to rectangular blocks is utilized subsequently. Here target blocks are marked as 1, whereas background as 0. Equations are expressed as (6), (7).

$$P_{avg} = \sum_{i=0}^{M \times N - 1} P_i / (M \times N), \tag{6}$$

$$P_{avg} = \sum_{i=0}^{M \times N - 1} P_i / (M \times N). \tag{7}$$

The  $P_i$  mentioned above denotes all pixel values of points in the  $M \times N$  rectangle.  $P_{M \times N}$  denotes pixel tag values of the rectangular block. The size of visual image after being transformed is  $96 \times 144$ . When  $M \times N$  is set as  $6 \times 6$ , a  $16 \times 24$  map can be generated. When  $M \times N$  is set as  $6 \times 9$ , a  $16 \times 16$  map can be generated.

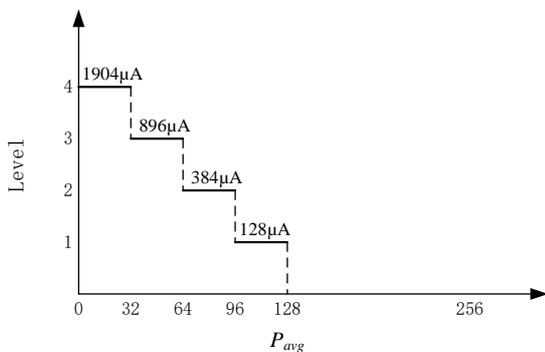


FIGURE 3 The brightness grading of simulated phosphene map

Brightness grading is unique to image processing for artificial vision. Brightness grade is used to control the

microelectrodes to stimulate current intensity. The higher the stimulation current, the clearer the phosphene generated in the visual cortex [11-12]. Average of  $M \times N$  rectangular block  $P_{avg}$  grayscale is regarded as a criterion for classification and be divided into 5 brightness levels which is revealed in Figure 3. This grading pattern indicates that the closer the target block tending to black target, the higher the brightness level of the simulated phosphene map. The maximum rating corresponding to stimulation current reaches  $1904 \mu A$ . The minimum rating means there is no electrical stimulation.

A  $16 \times 24$  bitmap will be divided into 24 small rectangular blocks with size of  $4 \times 4$  after acquiring. According to Equation (8), pixel analysis is put to use for each block separately. The needed  $4 \times 6$  simulated phosphene map can be generated ultimately. The  $P_{4 \times 4}$  denotes tag value of each corresponding block, the  $P_i$  denotes tag value of each pixel in a  $4 \times 4$  block.

$$P_{4 \times 4} = \begin{cases} 1 & \text{if } (\sum_{i=0}^{15} P_i \geq 6) \\ 0 & \text{else} \end{cases}, \tag{8}$$

After visual image processing mentioned above, the process of reducing the original  $240 \times 360$  image resolution to the  $4 \times 6$  simulated phosphene map which limited by the number of implanted electrodes is basically realized.

3 Experimental results and analysis

3.1 ARTIFICIAL VISION SYSTEM PLATFORMS

As TI's high performance multimedia processors, DM6437 is chose as the hardware platform for experiments. Its processing logic is clocked at 594MHz. DM6437 integrates video capture interface VPFE and output interface VPBE, with a 64-channel enhanced EDMA. Its chip storage architecture includes 80KB L1D RAM/Cache and 128KB L2 RAM/Cache. In this paper, an artificial vision hardware system is designed as shown in Figure 4a and Figure 4c. It includes the core board, interface board, and camera module. In Figure 4b, hardware is assembled in Vision Glass.



a)



b)



c)

FIGURE 4 Artificial vision hardware system

3.2 SPATIAL DOMAIN RESPONSE EXPERIMENT

So far, since few products are turn out for artificial vision in practical application, there is no actual accepted standard for experiments. In our experiment, a spatial domain response test standard applicable to artificial vision systems is developed. In Figure 5, pure video source matrixes (such as black and white displays) are with size of 4×6. Each electrode channel has a corresponding unique point in the space. The sequences as shown in Figure 5 separately correspond to electrode numbered as 1, 6, 11 and 16 which sends electrical stimulation signals.

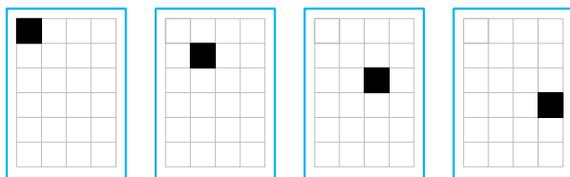


FIGURE 5 Standard sequences diagram of spatial domain response test

The illumination in this spatial domain response experimental environment shall be not less than 30Lu. Test steps are as follows. Firstly, the 24-inch display plays standard test sequences in turn. Secondly, the Vision Glass is moved to a certain distance so that the converted 240×360 image edge can overlap to the blue border edge of test sequence. Finally, a significant step is to test whether only one corresponding micro-electrode can generate stimulation signals through the oscilloscope.

Figure 6 performs a visual image processing of simulated phosphene map, which includes perspective transformation, image zooming, grayscale conversion and reducing pixelated image processing method. It can be seen that due to the influence of illumination and the shortness of camera, the black and white contrast difference of the collected 240×360 image is not obvious, but the pixel information of the test sequence is largely retained. Distorted areas are not appeared when image zoom out to a resolution of 96×144. This approach deals well with the binaryzation processing in the process of gray conversion and also effectively remove noises. In this frame processing, the optimal global threshold value T is 102.

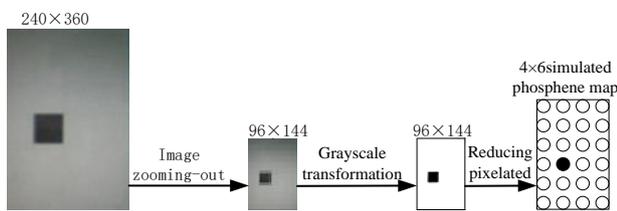


FIGURE 6 An example: Good quality with clear lettering

To further validate that system has certain anti-interference capability when generating 4×6 simulated phosphene map, the following test has been done.

1) For the situation that computer screen appears obvious dark side in the top edge occasionally, anti-interference capability of this system is firstly discussed

under the condition of uneven brightness. During the experiment, illumination is deliberately increased at the bottom to highlight the uneven brightness of top and bottom edges. The image processing results of fixed threshold algorithm, global OTSU dynamic threshold algorithm and local OTSU dynamic threshold algorithm are conducted separately.

Figure 7a and Figure 7b respectively are using the fixed threshold algorithm and global dynamic threshold algorithm to generate the 4×6 simulated phosphene map. Both of them do not meet the spatial response test standard. For the fixed threshold algorithm, local darker pixels in the upper edge are converted to local target objects, mistaking to simulate the mapping relationship between the original non-target area pixels and phosphene map with brightness rating of 1 after reducing pixelated. For global OTSU dynamic threshold algorithm, because darker pixels in the upper edge of the global image are in a larger proportion compared to object, the calculated global threshold T is larger, which contributes to a lot of the wrong object residual. Consequently, its actual effect is worse than the fixed threshold algorithm. The effect of local OTSU threshold algorithm used in our approach is given in Figure 7c. Here the 96×144 original images are divided into 48×48 rectangular blocks with 2×3 pixels. Figure 7c demonstrates that the object and background differences are most obvious in the 4th rectangle block. Other blocks are more flat. Setting local threshold value T4 as global threshold value T for gray conversion makes the converted effect better than the first two and has no effect to simulated phosphene map after reducing pixelated. Apparently, these processing results meet spatial response test standard as well.

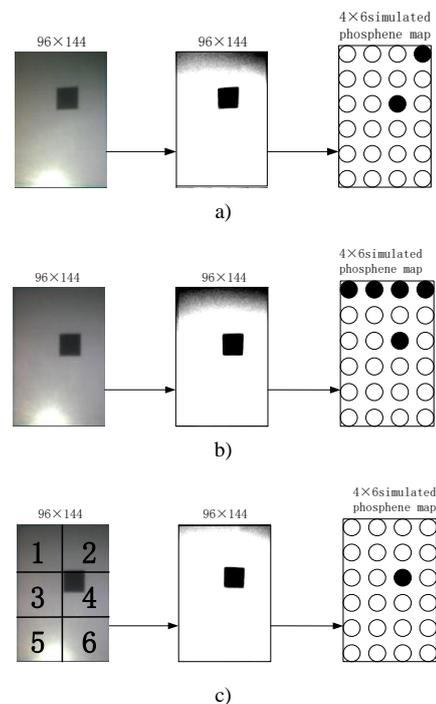


FIGURE 7 Spatial response experiment results

2) Standard test sequence with interference of discrete particle noise. Figure 8 displays the tailored visual images after resizing. After gray conversion, the noise significantly reduced. Furthermore no loss of black target block pixels is in this large area. The test perfectly accomplishes binaryzation steps. Ultimately, the generated 4×6 simulated phosphene map meets the spatial response test standard after the reducing pixelated.

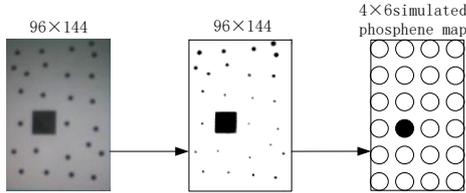


FIGURE 8 Process effects with interference of discrete particle noise

3) Standard test sequence with interference of black object with local loss in Figure 9. Obviously, a phenomenon of black object with local loss exists after resizing. After grayscale conversion, the test retains the basic pixel information of target block. Therefore, it cannot generate an error 4×6 simulated phosphene map after reducing pixelated, which avoids causing interference to clinical trials.

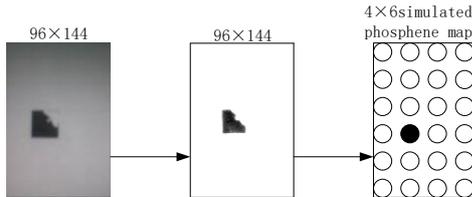
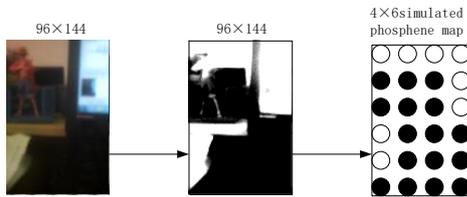


FIGURE 9 Process effects with interference of black target with local loss



a)



b)

TABLE 1 System memory consumption

| Key module                   | Resize | Gray Conversion | Simulated Phosphene | Electrode Stimulate | All   |
|------------------------------|--------|-----------------|---------------------|---------------------|-------|
| Before optimization (Cycles) | 6.0M   | 5.9M            | 3.0M                | 30.4M               | 45.3M |
| After optimization (Cycles)  | 0.4M   | 1.9M            | 1.0M                | 11.4M               | 14.7M |

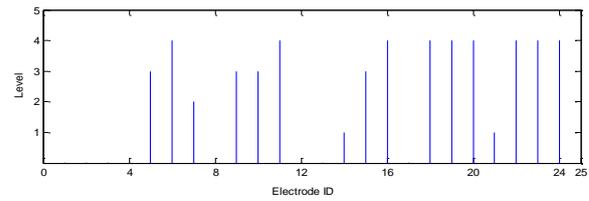


FIGURE 10 Process effects under the ubiquitous scenes

4) The final test is to verify the visual image processing effects under a ubiquitous scene. Figure 10a illustrates the tailored visual image information under a ubiquitous scene. The experiment basically retains pixel information of the target block after gray conversion. What's more, the corresponding 4×6 simulated phosphene map's distribution is consistent with the spatial response. Process effects can basically meet the hospital clinical trials' standard. Figure 10b is a real-time image information output. The display includes real-time visual images after perspective changing as well as a 4×6 or 16×16 simulated phosphene map to give doctors the most intuitive visual representation. Figure 10c gets analysis result on the brightness level of the 4×6 simulated phosphene map in Figure 10a. Experiment proves that brightness level of the valid simulated phosphene map is consistent with the grayscale information of actual object block.

In addition, we evaluate system storage resource cost for the implementation of simulated phosphene map's image processing strategy in DM6437. On-chip and off-chip storage resource consumption statistics is given in Table 1.

Moreover, in order to compare processing performance between before and after the optimization processing, we used the CCS5.4 Clock tool to do statistics for the consumed clock cycles by system core module. Since system hardware has certain requirements for low-power, the work main frequency of the selected DM6437 processors is 432MHz.

The data in Table 2 indicates that the consuming clock cycles of visual image processing algorithms significantly decline after the system was optimized. Then covering the time consumed by simulated phosphene map data coding and electric stimulation, a complete link restoration's speed reaches 30 frames per second, which fully meets the real time needs of visual centers system.

TABLE 2 The computational performance of key modules

| Memory | Size  |
|--------|-------|
| L1D    | 7.5KB |
| L2D    | 2KB   |
| DDR2   | 9.5MB |

#### 4 Conclusion

Artificial vision is a hot topic in many fields, such as clinical medicine, biomedical engineering and computer science. For solving the problem of simulated phosphene

map generation wanted by medical experiment, based on block segmentation, a reducing pixelated image processing method is proposed, which effectively works out the contradiction between the 4×6 limited electrode map and the big visual information. Meanwhile, electrode stimulus strategy is built based on the association between brightness grade and stimulus current intensity. Afterwards, this strategy of artificial visual system is verified by the DSP platform, the visual image processing speed reaches 30fps after optimization, which basically meets the processing needs of the animal or human visual cortex system in real time.

#### References

- [1] Luo Y H L, Da Cruz L 2014 A review and update on the current status of retinal prostheses (bionic eye) *British Medical Bulletin* **109**(1) 31-44
- [2] Hu J, Xia P, Gu 2014 Recognition of Similar Objects Using Simulated Prosthetic Vision *Artificial organs Pergamon:Oxford* **38**(2) 159-67
- [3] Xia P, Hu J, Chai X Y 2014 Visual prosthetic system modeling and visual information processing *Journal of Shanghai Jiaotong University* **48**(2) 164-7
- [4] Greenwald S H, Horsager A, Humayun M S 2009 Brightness as a function of current amplitude in human retinal electrical stimulation *Investigative Ophthalmology & Visual Science* **50**(11) 5017-25
- [5] Shepherd R K, Shivdasani M N, Nayagam D A X 2013 Visual prostheses for the blind *Trends in Biotechnology* **31**(10) 562-71
- [6] Eiber C D, Lovell N H, Suaning G J 2013 Attaining higher resolution visual prosthetics: a review of the factors and limitations *Journal of Neural Engineering* **10**(1) 1-17
- [7] Dorn J D, Ahuja A K, Caspi A 2013 The detection of motion by blind subjects with the epiretinal 60-electrode (Argus II) retinal prosthesis *JAMA Ophthalmology* **131**(2) 183-9
- [8] Van Rheede J J, Kennard C, Hicks S L 2010 Simulating prosthetic vision: Optimizing the information content of a limited visual display *Journal of Vision* **10**(14) 32
- [9] Zhao Y, Lu Y, Tian Y 2010 Image processing based recognition of images with a limited number of pixels using simulated prosthetic vision *Information Sciences* **180**(16) 2915-24
- [10] Wang J, Lu Y, Gu L 2014 Moving object recognition under simulated prosthetic vision using background-subtraction-based image processing strategies *Information Sciences* **277** 512-24
- [11] Lu Y, Wang J, Wu H 2014 Recognition of Objects in Simulated Irregular Phosphene Maps for an Epiretinal Prosthesis *Artificial Organs* **38**(2) 10-20
- [12] Tapia E, Mazzi C, Savazzi S 2014 Phosphene-guided transcranial magnetic stimulation of occipital but not parietal cortex suppresses stimulus visibility *Experimental Brain Research* **232**(6) 1989-97

| Authors   |  |
|---|--|
|  | <p><b>Longmin Chen, 1989.4, Zhejiang, China.</b></p> <p><b>Current position, grades:</b> master, grade 3.<br/> <b>University studies:</b> control science and engineering.<br/> <b>Scientific interest:</b> video image processing.</p>  |
|  | <p><b>Fanghua Yan, 1991.1, Zhejiang, China.</b></p> <p><b>Current position, grades:</b> master, grade 2.<br/> <b>University studies:</b> control science and engineering.<br/> <b>Scientific interest:</b> video image processing.<br/> <b>Publications:</b> Wang Z L, Yan F H and Zheng Y Y 2013 An adaptive uniform distribution surf for image stitching <i>Image and Signal Processing (CISP)</i> <b>2</b>, 888-92</p>   |
|  | <p><b>ZeZhi Gong, 1989.9, Zhejiang, China.</b></p> <p><b>Current position, grades:</b> Master, Grade 3.<br/> <b>University studies:</b> Control science and Engineering.<br/> <b>Scientific interest:</b> image processing about artificial visual prostheses.<br/> <b>Publications:</b> Gong Z Z, Cheng L M, Yang H S and Zheng Y Y Artificial Vision System for Animal Experiments <i>Computer engineering and design</i>.</p>   |
|  | <p><b>Yayu Zheng, 1978.4, Zhejiang, China.</b></p> <p><b>Current position, grades:</b> associate professor of Zhejiang University of Technology.<br/> <b>University studies:</b> BSc and PhD degrees from Zhejiang University, Hangzhou, China, in 2002 and 2008, respectively.<br/> <b>Scientific interest:</b> video analysis.<br/> <b>Publications:</b> Yayu Zheng, Fan Zhou, Xiang Tian, Yaowu Chen 2008 Lightweight content-adaptive coding in joint analyzing-encoding framework <i>IEEE Transactions on Consumer Electronics</i> <b>54</b>(2),614-622</p> |