Sensors and Materials, Vol. 25, No. 1 (2013) 79–97
MYU Tokyo

S & M 0904

Vibrotactile Representation of Three-Dimensional Shape and Implementation on a Vibrotactile Pad

Juan Wu*, Yu Ding, Dejing Ni, Guangming Song and Wei Liu

School of Instrument Science and Engineering, Southeast University,
Sipailou 2#, Nanjing, Jiangsu Province 210096, China

(Received September 28, 2012; accepted November 19, 2012)

Key words: contour lines, 3D shape, vibrotactile displa***y***

Representing visual information to visually disabled persons is a great challenge in the field of human computer interface (HCI). The technology based on a tactile vision substitution system (TVSS) device has become a promising solution to this problem. This kind of device allows visually disabled persons to perceive specific visual information through tactile stimuli. Research on such a device also has potential social and economic impacts on normal individuals. We propose to develop a system in which the three-dimensional (3D) shape of an object is acquired and transformed into a vibrotactile stimulus. The tactile represention procedure includes three steps: Firstly, the shape-from-shading (SFS) algorithm is adopted to extract the 3D shapes of an object. Then, specific contour lines are extracted for shape representation. Finally, the contour lines are mapped into a vibrotactile array (20×20), and the tactors corresponding to the lines are actuated using variable spatiotemporal vibration parameters. The preliminary experimental results suggest that subjects could successfully identify the difference in contour, and shape perception can be facilitated by adjusting the vibration parameters.

1. Introduction

Tactile displays for conveying spatial information such as shape and texture have been studied for over 40 years.[1] They can be used for fabricating sensory substitution devices for the visually impaired, as well as for people with normal vision engaged in specific work under conditions where there is restricted vision or high cognition load.[2]

Since the 1960s, researchers have developed many tactile display methods and prototypes to convey small-scale force to the users. The most common design approach for these displays uses arrays of pins that rise against the user's skin to approximate

*Corresponding author: e-mail: juanwuseu@seu.edu.cn

arbitrary shapes. Actuation approaches for pin-based tactile displays can be broadly classified into two categories: static and vibrotactile.[3] Static displays can control the pins to rise 5–57 mm with amplitude. These devices are often intended to stimulate the sensation of contact with 3D objects. Thomas P. Way and Kenneth E. Barner designed a tactile image creation system (TACTICS).[4,5] The system processes an image on microcapsule paper, which expands at various heights to stimulate the user's fingers. Jiang *et al*. used shape memory alloys (SMAs) to design a tactile display board that fits on the user's palm. The board can show the outlines of simple objects by changing the length of SMAs.[6] Wagner *et al*. used radio-controlled servomotors to actuate an array of mechanical pins. The prototype was used to convey shape information to the users.[7] The pins could deflect individually and be used to sculpture a 3D shape precisely. However, the required hardware usually has the shortcomings of a slow response and a complicated mechanical structure.

Vibrotactile displays typically consist of tactors that vibrate at frequencies of 200 –250 Hz. In this range, the frequency perceptual threshold is at the lowest, and low-cost and small actuators could be used. These TVSS devices are mostly used for tactile image and letter display. In 1969, Bach-y-Rita *et al.* introduced a tactile-visual conversion system.[8] The system converts brightness information taken from an object's image to a vibration stimulation signal. It fixes a 20×20 motor array on the back of a seat so that users can feel the vibration and perceive the outline of objects in the image when they sit on the seat. In 1970, Bliss *et al.* developed an optical-to-tactile image conversion system named Optacon.[9] The system converts brightness information of an image to vibrotactile stimulation. It drives the motor array that is fixed on a user's finger to display tactile information. In 1987, Kaczmarek *et al.* developed a tactile vision-substitution system (TVS). The system is based on image segmentation technology. It divides an image into 6×24 regions and displays the content in each region continuously and dynamically. Through user's memory splicing, it can show the entire image.[10] In 1997, Kurze proposed a rendering method for generating tangible drawings of spatial real-world objects based on a theory of haptic image perception and understanding. The method is based on an analysis of the process of drawing used by blind people and cognitive considerations.[11] For vibrating devices, usually the frequency, intensity, and time rhythm are adjustable, which correspond to the spatial information. Little work on 3D shape representation with a vibrotactile device has been carried out before.

The goal of this work is to extract the 3D contour characteristics from the image captured by the camera, and represent the 3D shape on a vibrotactile display. It is intended to exploit the common vibrating array taking into the consideration human factors (*e.g.*, tactile illusion) to recreate the virtual presence of a fluctuant surface on the 2D vibrotactile pad. The design focuses on extracting the 3D shape characteristics from an image and mapping into the tactor array using a spatial-temporal vibrating pattern. To convey the perception of touching 3D objects, the human's perception of a vibrating stimulus is primarily considered. The vibration parameters such as vibration time interval, duration, and intensity on the contour line are experimentally determined. The rest of the paper consists of three sections. In § 2, we introduce the principle of the system, in § 3, we describe the experimental procedure and evaluation experiments, in § 4, we present our results and discussion, and in § 5, we present our conclusion and discuss our future works.

2. Principle

The general scheme for a tactile shape display system comprises three parts. They are 3D shape extraction, tactile coding, and vibration actuation. Two problems must be primarily considered before the implementation of 3D shape display on the vibrotactile device. One is the human factor, which is the difference in sensitivity between vision and tactile perception. The other is the fact that the commonly used tactor array has limits in scale and output performance.

To solve the first problem, namely, the resolution of the tactile sensation is much less than that of vision, it is suggested to extract the characteristics of a 3D object and decrease the content that needs to be transformed. Previously, a compression algorithm of tactile-vision conversion was studied in relation to the tactile-vision substitution system. Researchers in relevant fields have obtained many important results. In 1988, Bourbakis and Klinger[12] developed the Tyflos system, which is a wearable navigation system designed for the blind. The system uses a motor array to express 3D information. They presented a hierarchical scheme for the coding of a reduced amount of image information. In 2006, Lin developed a tactile vision-substitution system. He proposed to collect images whose resolution is 64×64 pixels and use the wavelet algorithm to decrease the resolution into 32×32 pixels. Then, corresponding motors were driven to tactilely display the image.[13] In 2007, Dakopoulos *et al*. proposed an improved algorithm and emphasized the important parameters affecting the system.[14] Krufka *et al.* proposed a new method of visual-to-tactile conversion. Methods that automatically convert graphics into raised line images have recently been investigated. The proposed algorithm extracts object boundaries and employs a classification process that is based on a graphic's hierarchical structure to determine critical outlines. A single parameter is introduced into the classification process, enabling users to tailor graphics to their own preferences. The resulting outlines are printed using a Braille printer to produce the tactile output.[15]

It should be noted that almost all of the solutions mentioned above focus on the 2D outline of an object. In the real environment, without prior training, users cannot even distinguish a ball from a circle in the image. It has significant limitations to the blind and people engaged in dangerous and complex jobs (such as pilots, firefighters, and divers) who rely on tactile sensation to perceive the surroundings. In order to display the 3D shape of the object, this paper proposes to use a vibration motor array to build a tactile system and focus on a new method to represent 3D shape. The method is mainly based on contour lines of an image. It simplifies the 3D profile of an object with specific contour lines, and then maps the contour lines onto the vibration motor array. It can represent the 3D information of the object by changing the amplitude, rhythm, and order of the corresponding motors. Users can perceive the 3D shape of an object when exploring the motor array. An experiment is designed to assess the effect of different vibration modes.[16,17] The results confirm the effectiveness of the method.

There are two limitations of the tactile sense, which restrict the vivid representation of a 3D shape. One is the fact that, nowadays, the resolution of a motor array is much lower than the pixel resolution of an image,[18] resulting in loss of some information

during the mapping. The other is the limited ability to distinguish vibration intensity. The experiments show that vibration intensity is no more than 7–8 levels.[19] The insensitivity of vibration intensity restricts the levels that we can utilize when using different vibration intensities to represent a 3D shape.

To solve the first problem, it is proposed to keep basic information such as outlines and contour lines. It is much easier to map the basic information onto a motor array. For the latter problem, we propose to take the front projection of an object as the target image, and simplify height information to several contour lines. The number of contour lines is determined by the gradient of surface height. An algorithm is proposed to calculate the number and height of contour lines. By this method, we overcome the limitations of human tactile sense and find an effective method to complete the vibrotactile representation of a 3D shape. The method is mainly composed of three parts: (1) 3D contour extraction, (2) contour line reconstruction and mapping, and (3) tactor actuation.

2.1 *3D contour extractions*

The flow chart of 3D contour extractions is shown in Fig. 1. The original image is either taken from a digital camera or the standard image stored in the computer. Both the object and background in the original image must be monochromatic. Therefore it is easier to extract basic information such as outlines and contour lines. The original image is converted to grayscale after specific processing. The shape-from-shading (SFS) technique[20] is used to extract height information from the image. An algorithm is proposed to determine the number of contour lines. It mainly depends on the variance of height information to complete the determination.

In this section, we mainly focus on the SFS technique and determination of the number of contour lines. We will talk about them respectively in §§ 2.1.1 and 2.1.2.

2.1.1 *Height information extraction*

The commonly used method uses a binocular vision device to obtain several images, and extract height information through the parallax of the images.[21] However this method relies heavily on hardware and cannot be operated easily. SFS can overcome the disadvantage mentioned above and only one image is needed in each rebuilding process. Therefore, SFS is chosen to extract height information in the method.

In practice, the brightness of an image is influenced by many factors such as the model of image formation, light source direction, and surface normal. In order to simplify the algorithm, we make some assumptions: (1) both the object and background

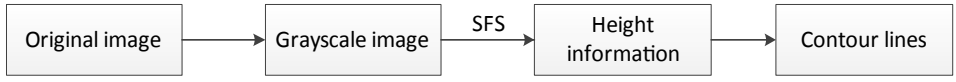


Fig. 1. Flow chart of 3D contour extractions.

in the image are monochromatic, (2) the model of image formation is the Lambertian model, (3) the light source is at infinity, (4) the light source direction is the same as the surface normal at each pixel in the image. However, the SFS algorithm based on the Lambertian model is morbid. To overcome this disadvantage, the Tsai and Shah algorithm[22] is used in the SFS algorithm. The algorithm retains the linear term of the Taylor expansion of the reflection function and makes the SFS algorithm linear.

Combined with assumptions and the simplification mentioned above, the SFS algorithm can be proposed. We assume that $E(x, y)$ represents the gray value of pixel in the image whose coordinate is (x, y); $R(p, q)$ represents the reflection intensity of point in the original object whose coordinate is (x, y, z). The point and pixel have the same 2D coordinate system (x-y coordinate system). $z = Z(x, y)$ represents the height at point (x, y). According to the conditions, we can obtain eq. (1)

$$E(x, y) = R(p, q), \tag{1}$$

where $p = \dfrac{\partial Z}{\partial x}$, $q = \dfrac{\partial Z}{\partial y}$.

The implementation of the Tsai and Shah algorithm is as follows: First, reflection function is discretized. Equations (2) and (3) are the discretized forms of reflection function,

$$p = \frac{\partial Z}{\partial x} = Z(x, y) - Z(x - 1, y), \tag{2}$$

$$q = \frac{\partial Z}{\partial y} = Z(x, y) - Z(x, y - 1). \tag{3}$$

Then, p and q are substituted into eq. (1). Thus, we can obtain the transformation of eq. (1), which is shown in eq. (4).

$$f(Z(x, y)) = E(x, y) - R(Z(x, y) - Z(x - 1, y), Z(x, y) - Z(x, y - 1)) = 0 \tag{4}$$

The Taylor expansion of eq. (4) is figured out and simplified to eq. (5).

$$0 = f(Z(x, y)) \approx f(Z^{n-1}(x, y)) + (Z(x, y) - Z^{n-1}(x, y)) \frac{\partial f(Z^{n-1}(x, y))}{\partial Z(x, y)}, \tag{5}$$

where $Z^n(x, y)$ represents the result of nth iteration.
Finally, $Z(x, y)$ is replaced by $Z^n(x, y)$. We can obtain an iterative equation as follows.

$$Z^n(x, y) = Z^{n-1}(x, y) + \frac{-f(Z^{n-1}(x, y))}{\frac{\partial f}{\partial Z(x, y)}(Z^{n-1}(x, y))} \tag{6}$$

The height of each pixel can be calculated after iteration. We use the matrix H to represent the heights of an image. The dimensions of H are the same as the resolution of the original image. For example, if the original image's resolution is $m \times n$ pixels, then a matrix H, which has m rows and n columns, can be figured out. $H(i, j)$ represents the height value of the pixel whose coordinate is (i, j), where $i = 1,2,3...m$, $j = 1,2,3...n$. In order to make the result more intuitive, a height map is used to represent the height value of each pixel. The height map is a 3D graph constructed by point sets $(i, j, H(i, j))$.

2.1.2 *Determination of number of contour lines*

We use SFS to extract the height information of a vase, a ball, and a cone from images and obtain the corresponding height map. As shown in Figs. 2(b) and 2(c),

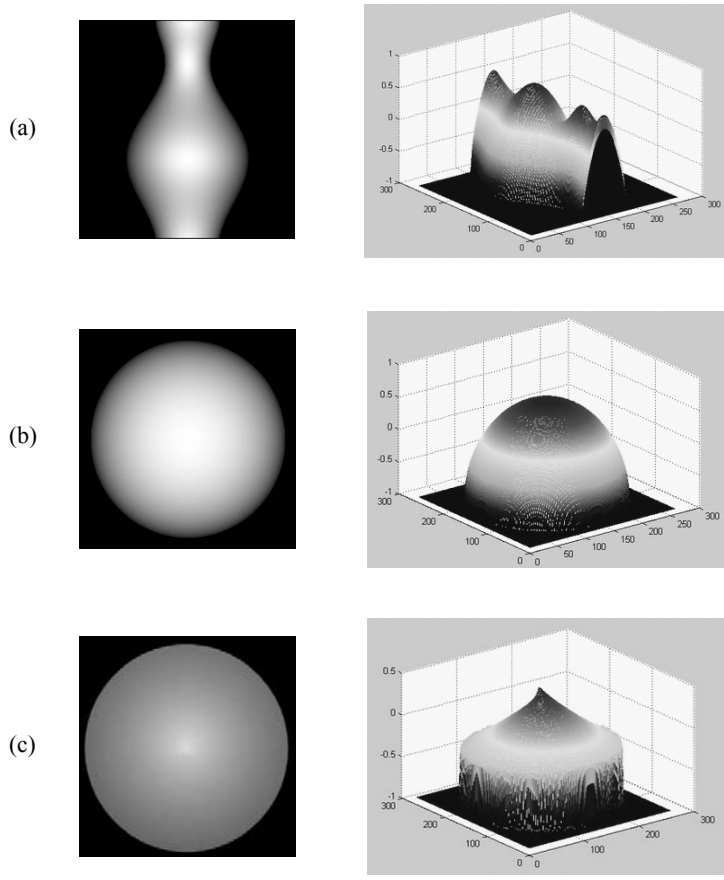


Fig. 2. Typical grayscale images and height maps: (a) grayscale image and height map of vase, (b) grayscale image and height map of ball, and (c) grayscale image and height map of cone.

although the 2D outlines of the ball and cone are the same, their 3D shapes are not the same. Therefore, if the tactile system only displays the 2D outline of the object, the user is unable to perceive the bump surface and fails to recognize objects.

In order to overcome this disadvantage, height information is used to distinguish the objects that have the same 2D outline. However, as mentioned above, there is a marked difference in sensitivity in vision and tactile perception. Therefore, the height information is simplified to several contour lines. The number of contour lines reflects the complexity of surface height. It should be carefully determined. We propose an algorithm for determining the number of contour lines. The algorithm depends on the variance of height value to set the number of contour lines. The specific steps are as follows:

(1) Heights are normalized using eq. (7), where $H_{\max}$ represents the maximum of heights, $H_{\min}$ the minimum height, and $h(i, j)$ the normalized data corresponding to $H(i, j)$.

$$h(i,j) = \frac{H_{\max} - H(i,j)}{H(i,j) - H_{\min}} \tag{7}$$

(2) The average of normalized data is calculated using eq. (8), where $\bar{h}$ represents the average.

$$\bar{h} = \frac{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} h(i,j)}{m \times n} \tag{8}$$

(3) The variance of normalized data is calculated using eq. (9), where V_{ar} represents the variance of normalized data.

$$V_{\text{ar}} = \frac{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (h(i,j) - \bar{h})^2}{m \times n} \tag{9}$$

(4) If $V_{\text{ar}} < 0.02$, the number is set to be 1. If $0.02 < V_{\text{ar}} < 0.06$, the number is set to be 3. If $V_{\text{ar}} \geq 0.06$, the number is set to be 5.

We assume that the number is N_l. The height of each contour line can be determined.

$$tp_k = h_{\min} + k \times \frac{h_{\max} - h_{\min}}{N_1}, \tag{10}$$

where tp_k is the height value of kth ($k = 1,2,3,...N_l$) contour line.

2.2 *Contour line reconstruction and mapping*

It usually occurs that the trace of pixels with the same height cannot be a continuous contour line or that contour lines are often too thick. Therefore, it is necessary to connect

and sharpen the contour lines. On the other hand, the resolution of a motor array is much lower than that of the height map. A compression algorithm is proposed to complete the mapping. The algorithm is efficient and easy to understand. It can keep the basic information of the height map during the mapping.

2.2.1 *Sharpening of contour lines*

Each contour line can be seen as the boundary line of two different regions. If there are pixels of different regions in a pixel's 8-neighborhood, then the pixel can be seen as the pixel at the contour line. Therefore, we can use this feature to sharpen the contour.

Take the kth contour line as an example, the pixel whose height is less than tp_k is set to be black and others are set to be white, where tp_k is the height of the kth contour line. Then, we connect the dots that have both black pixels and white pixels in their 8-neighborhood. By this method, we can obtain a continuous thin contour line. In order to record pixels accurately, the matrix S is introduced. If the resolution of the image is $m \times n$, S has m rows and n columns. $S(i, j)$ is the state of the pixel whose coordinate is (i, j). If a pixel is not at the contour line, its state is set to be 0. If a pixel is at the kth contour line, its state is set to be k. In order to make the result more intuitive, a contour map is used to show the contour lines of an object in an image. The pixel that is not at the contour line is set to be black. The pixel that is at the contour line is set to be a specific color. The color of pixels at different contour lines is different.

The contour maps of a vase, a ball, and a cone are shown in Fig. 3. Owing to their relatively simple shape, it is sufficient to select only three contour lines to constitute the contour map. As shown in Fig. 3, the blue, green, and red lines indicate three different contour lines; their corresponding heights are increasing. It is shown in Figs. 3(b) and 3(c) that the distances between contour lines of the ball and cone are, respectively, not equal and equal. It shows that the heights of the surface of the ball and cone change nonlinearly and linearly, respectively. This proves that it is reasonable to use the number of contour lines and the distance between adjacent contour lines to respectively reflect the complexity and surface height variation of the object.

2.2.2 *Mapping*

Space mapping is the process that maps the contour of an object to the motor array (as shown in Fig. 4). The resolution of the motor array used in the research is 20×20. In order to record each motor in the array unambiguously, a 2D array, $M(i, j)$, is used (i = 1,2,3...20 represents the row of the motor array, j = 1,2,3...20 represents column of the motor array). The origin of the motor array is in the upper left.

Note that all the images maintain the same aspect ratio during the mapping. This is because a change in aspect ratio will distort an object's shape. For example, the resolution of the original contour map is $m \times n$ and it has N_l contours. The resolution is compressed to $w \times h$ after mapping, where

$$w = \left[\frac{m}{\max(m, n)} \times 20 \right], \quad h = \left[\frac{n}{\max(m, n)} \times 20 \right], \tag{11}$$

where [ ] means the ceil function. The original contour map is divided into $w \times h$ regions; each region corresponds to a motor.

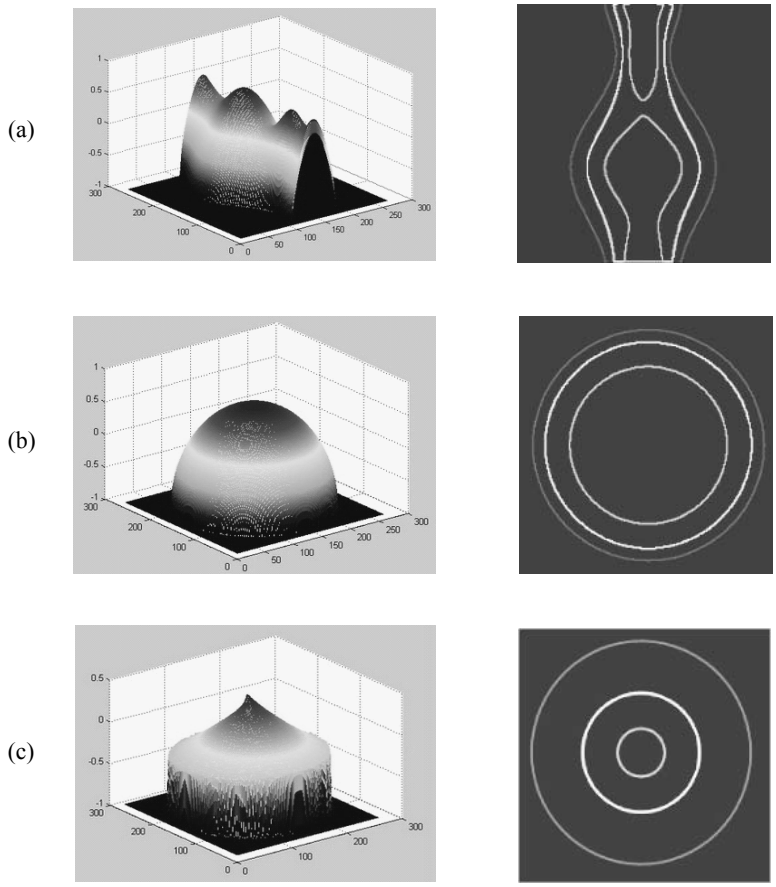


Fig. 3. Typical objects' height map and contour map: (a) vase, (b) ball, and (c) cone.

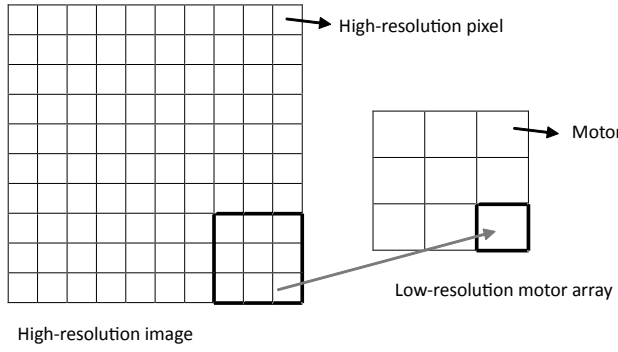


Fig. 4. Schematic diagram of mapping.

The rules of drawing a contour line in a lattice of low resolution are as follows:

(1) If the region does not contain any pixels of a contour line, then set $M(i, j) = 0$.
(2) If the region contains pixels of the kth (k = 1,2...N_l) contour line, then one can calculate the proportion of the pixels of a contour line in the region. If the proportion is higher than 0.01, then set $M(i, j) = k$; otherwise, set $M(i, j) = 0$.
(3) If the region contains pixels of different contour lines, find the contour lines (k_{max}th, k_{max} = 1,2...N_l) that have the most pixels in the region. Then, calculate the proportion of the pixels of the k_{max}th contour lines in the region. If the proportion is higher than 0.01, then set $M(i, j) = k_{max}$; otherwise, set $M(i, j) = 0$.
(4) Each dot is assigned after a series of calculations defined above, and then connect the dots with the same positive value.

In this way, space mapping from an image to a tactile array is finished. Figure 5 shows the contour map and the map mapped into a lattice of low resolution. The blue, green, and red dots represent the three different contour lines. The white dots represent the background of the map.

2.3 *Motor actuation*

The vibration motor is driven by a square wave whose frequency is 100 Hz. The duty cycle of the square wave varies. We define 11 (level 0–level 10) vibration intensity levels. Level 10 denotes the duty cycle of the square wave is 100%. Vibration at level 10 is the strongest. Level 0 denotes the duty cycle of the square wave is 0%. The square wave at level 0 cannot drive motors to vibrate. The levels have a linear relationship with the duty cycle. According to the linear relationship, the effective values of the square wave at different levels can be determined. Table 1 shows the relationship between intensity, duty cycle, and effective value.

Specifically, the relationship between vibration intensity and effective value is shown in Fig. 6. The figure shows that the intensity of motors could be linearly controlled in the range from 1 to 10.

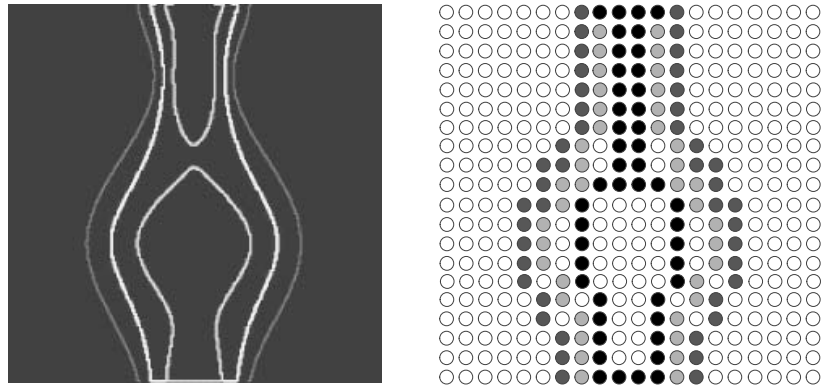


Fig. 5. Contour map and tactile display of vase.

Table 1
Relationship between intensity, duty cycle, and effective value.

Vibration intensity level	0	1	2	3	4	5	6	7	8	9	10
Duty cycle	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Effective value (V)	0	1	1.49	1.94	2.42	2.86	3.34	3.80	4.21	4.55	4.85

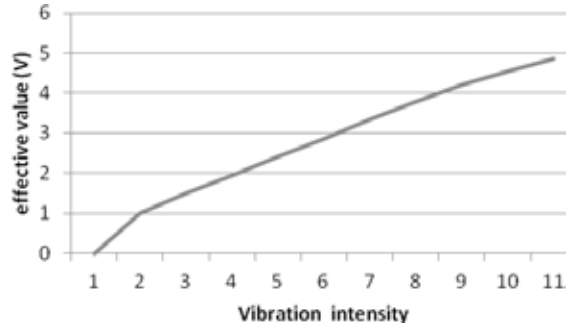


Fig. 6. Relationship between intensity and effective value.

The duration of vibration and vibration interval are both set manually. We give an example to explain two concepts. Suppose there are two motors A and B. A vibrates first and B vibrates after A. The duration of vibration is used to record how long A or B vibrates. The vibration interval is used to record the interval between the start-up of A and that of B (as shown in Fig. 7). In this paper, the specific form of number such as t_D/t_I is used to describe the time parameters of vibration, where t_D represents the duration of vibration and t_I represents vibration interval.

3. Experimental Procedure

This experiment is carried out to test the performance of the proposed tactile-vision substitution in 3D shape display.

3.1 *Experimental design*

The contour lines of a ball are shown in Fig. 8(a) ($L1$, $L2$, and $L3$). Their height values are H_{L1}, H_{L2}, H_{L3} ($H_{L1} < H_{L2} < H_{L3}$). Because the height variation of the ball is uniform, the distance between contour lines is the same. Figure 8(b) shows the compressed contour map in the lattice space.The blue, green, and red lines respectively represent $L1$, $L2$, and $L3$.

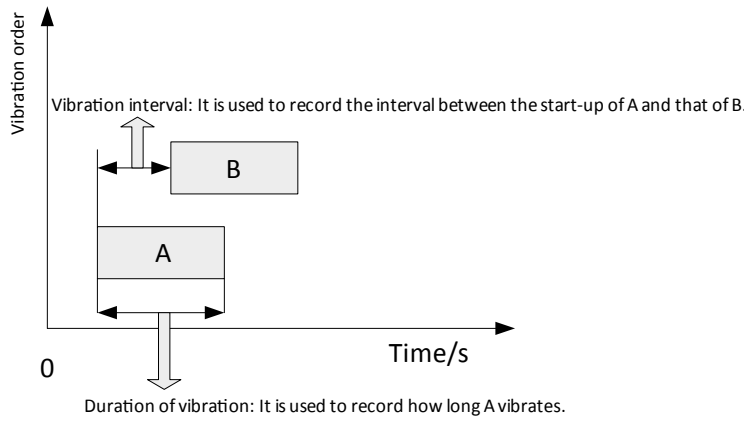
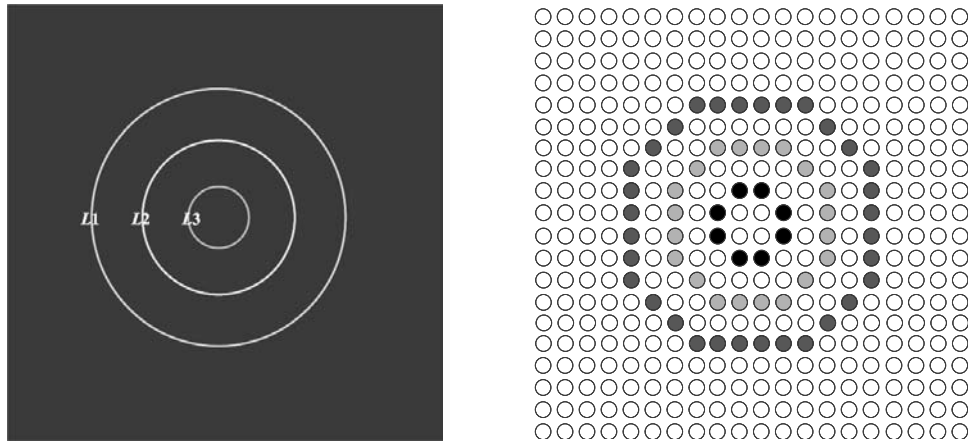


Fig. 7. Schematic diagram of vibration order.



(a) (b)

Fig. 8. Contour map and lattice space map: (a) contour map and (b) lattice space map.

There are six vibrotactile modes to express the 3D shape of the object using different display modes (static or dynamic), durations of vibration, vibration intervals (fixed or varied), and vibration intensities (linear or nonlinear).

(1) In the static display mode, all motors corresponding to contour lines vibrate at the same time. The vibration intensity of the motor is set by the height of the corresponding

contour line. A larger height means a higher intensity. When people perceive vibrotactile information, they touch some illusory dots. It is helpful to enhance the resolution of the lattice.[23] In this mode, the duration of vibration is a very important parameter. If it is too short, the user cannot fully perceive the shape. If it is too long, the user will become uncomfortable or lose concentration. In the experiment, the duration of vibration/vibration interval is set to be 4000/0 ms. As shown in Fig. 8(b), the color-coded dots represent the motors that need to be driven. The motors of different contour lines are distinguished by color. In the static display mode, these motors will vibrate at the same time and last for 4000 ms.

In the dynamic mode, motors corresponding to the same contour line vibrate at the same time. Motors of different contour lines vibrate at a specific order (as shown in Fig. 9).[24] The order is indicated by black arrows. It takes two cycles to complete the whole representation. It is also feasible to change the distance between contour lines, motor vibration duration, and number of contour lines to obtain different dynamic vibrotactile modes.

(2) The duration of vibration/vibration interval can be either fixed or varied. In this experiment, the fixed and varied modes are both selected to determine which mode can help increase the recognition rate. As shown in Fig. 8, in the fixed mode, the durations of vibration/vibration interval of all the corresponding motors are set to be 900/600 ms. In the varied mode, a longer duration of vibration/vibration interval represents a larger height. Therefore, the durations of vibration/vibration interval of $L1$, $L2$, and $L3$ are respectively set to be 1600/1400, 1200/1000, and 800/600 ms.

(3) As mentioned above, the vibration intensity of each contour line is different. A higher intensity means a larger height. However, the difference in the vibration intensity of the motors at the adjacent contour lines can either be the same or different. If the difference is the same, the situation is called the linear mode. If the difference is different, the situation is called the nonlinear mode. In the linear mode, the vibration intensities I_{L1}, I_{L2}, and I_{L3} (corresponding to $L1$, $L2$, and $L3$, respectively, $I_{L1} < I_{L2} < I_{L3}$) are respectively set to be 5, 3, and 1, and they are set to be 5, 4, and 1 in the nonlinear mode.

In summary, all the parameters of the different modes are shown in Table 2.

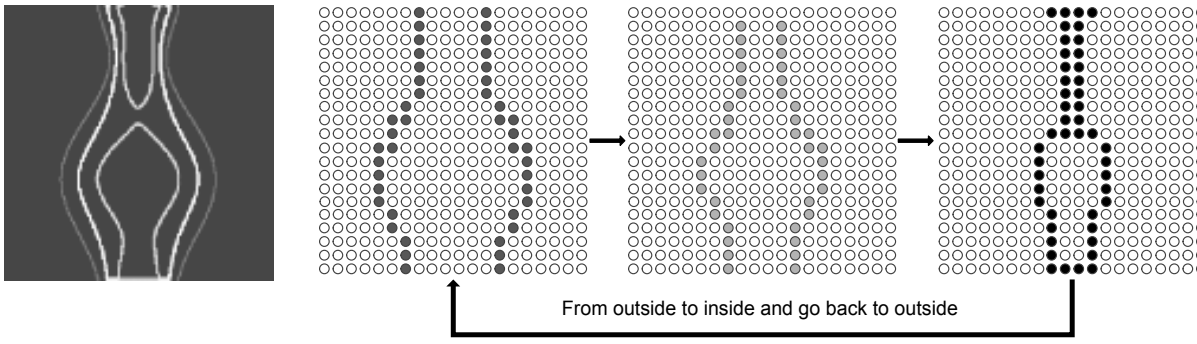


Fig. 9. Schematic diagram of dynamic mode: (a) contour map and (b) dynamic mode.

Table 2
Parameters of different modes.

		Duration of vibration (ms)			Vibration interval (ms)			Vibration intensity		
Mode		$L1$	$L2$	$L3$	$L1$	$L2$	$L3$	I_{L1}	I_{L2}	I_{L3}
Dynamic mode	Mode 1	600	600	600	900	900	900	5	3	1
	Mode 2	600	600	600	900	900	900	5	4	1
	Mode 3	1400	1000	600	1600	1200	800	5	3	1
	Mode 4	1400	1000	600	1600	1200	800	5	4	1
Static mode	Mode 5	0	0	0	4000	4000	4000	5	3	1
	Mode 6	0	0	0	4000	4000	4000	5	4	1

Modes 1, 2, 3, 4, 5, and 6 respectively represent dynamic fixed and linear mode, dynamic fixed and nonlinear mode, dynamic varied and linear mode, dynamic varied and nonlinear mode, static and linear mode, and static and nonlinear mode.

3.2 *Participants*

Fifteen subjects participated in the experiment (ten males, five females). They never participated in a similar experiment before. The average age of the participants was 24.

3.3 *Apparatus*

The device used in the experiment is shown in Fig. 10. It is mainly composed of three parts: image acquisition module, visual/tactile information conversion module, and vibrotactile display module. The image acquisition module uses a USB camera, which is hot-swappable and can acquire images in real time. The visual-to-tactile information conversion module uses a personal computer (PC) to process data. Its main work includes four aspects: (1) Control the image acquisition module for image acquisition. (2) Do image processing and extract image features. (3) Map the feature information into the tactile lattice and complete the vibrotactile coding. (4) Drive motors in vibrotactile display module to express tactile information. The vibrotactile display module is made up of 20×20 vibration motors.

3.4 *Procedure*

Before the experiment, the subjects are told that the greater vibration intensity represents a greater height value. During the experiment, the subjects respectively perceived six representation modes and determined which one matched with their perception the most. Each mode randomly appears 10 times, depending on the vibrotactile information they perceived. The participants chose the most likely shape from five options: A) plane circular, B) ball, C) uncertain, D) concave surface, and E) uneven surface. There are 6×10 trials in total for each participant.

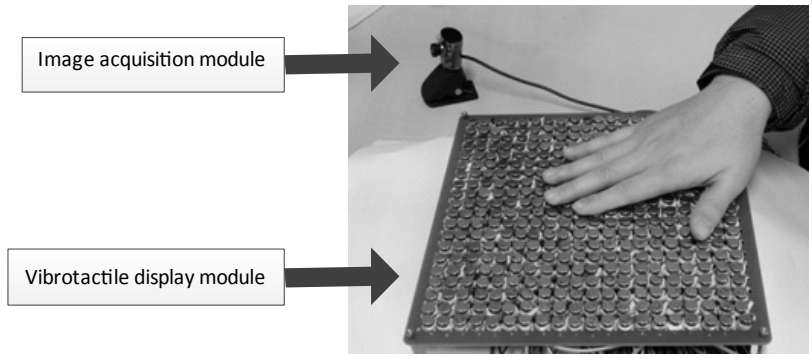


Fig. 10. Experimental setup.

4. Results and Discussion

The recognition rates are shown in Fig. 11. It shows that the recognition rates in Modes 1 and 2 were respectively 23.3 and 19.3%. The average recognition rate in the static mode is 21.3%. The rates of erroneously judging the ball as a plane circular in Modes 1 and 2 are respectively 66.7 and 69.3%. The recognition rates in Modes 3, 4, 5, and 6 are respectively 50.7, 47.3, 74.7, and 52.7%. The average recognition rate in the dynamic mode is 56.3%. ANOVA shows that the vibration mode (two levels: static, dynamic) has a significant impact on recognition rate [$F(1,14) = 28.775$, $p < 0.0001$]. The accuracy of recognition is significantly higher in the dynamic mode than in the static mode. Therefore, it proves that the static display mode cannot express height information effectively, whereas the dynamic mode can express the height information of objects better.

From ANOVA, we can determine that the duration of vibration/vibration interval (factor A1 represents the fixed mode and A2 represents the varied mode) has a significant impact on recognition rate [$F(1,14) = 9.787$, $p < 0.007$]. The recognition rate of the varied mode is significantly higher than that of the fixed mode. The intensity increasing mode (factor B; B1 represents the linear mode, B2 represents the nonlinear mode) has a significant impact on recognition rate [$F(1,14) = 5.530$, $p < 0.034$]. The recognition rate is significantly higher in the linear mode than in the nonlinear mode. The interactive impacts of factors A and B on recognition rate are significant [$F(1,14) = 4.822$, $p < 0.045$]. Therefore, it is neccssary to perform a simple effect analysis. From the data above, we determined that the recognition rate is more sensitive to factor A($p_A < p_B$)[25] because the difference between p_A and p_B is relatively large. The recognition rate is much less sensitive to factor B than to factor A. Therefore, a simple effect analysis of factor B is performed at levels A1 and A2. The results are shown in Table 3.

From Table 3, we know that the impact of factor B is not significant at the A1 level [fixed mode; $F(1,14) = 0.21$, $p < 0.657$]. However, the impact of factor B is significant at the A2 level [varied mode; $F(1,14) = 12.04$, $p < 0.004$]. The results of the simple effect

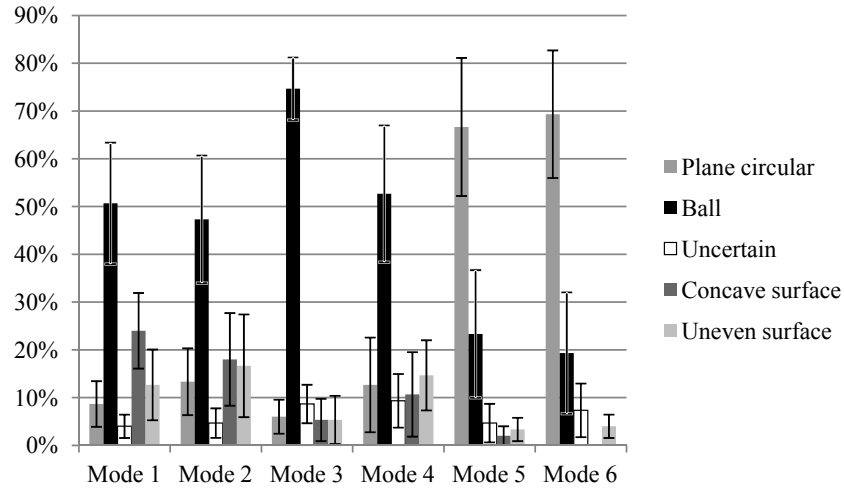


Fig. 11. Results of perception.

Table 3
Results of simple effect analysis.

Level	df of factor	df of error	*F*	*p*
A1	1	14	0.21	0.657
A2	1	14	12.04	0.004

analysis indicate that factor B has little influence on recognition rate in the fixed mode. Therefore, there is almost no difference between the recognition rates in the two modes. However, factor B has a significant influence on the recognition rate in the varied mode. Therefore, it is recommened to utilize this influence to improve the rate. Figure 12 shows the recognition rates of all the subjects in the fixed and linear mode, and the fixed and nonlinear modes. Figure 13 shows the recognition rates of all the subjects in both varied and linear mode, and varied and nonlinear mode.

The average recognition rates in the four modes are shown in Fig. 11. From Fig. 11, we determined that the average recognition rate in the varied and linear mode is the highest. From the discussion of the results of simple effect analysis, factor B has no significant impact on the recognition rate in the static mode. However, factor B has a significant impact on the recognition rate in the varied mode. Therefore, the dynamic varied and linear mode is the best choice for the expression of the 3D shape.

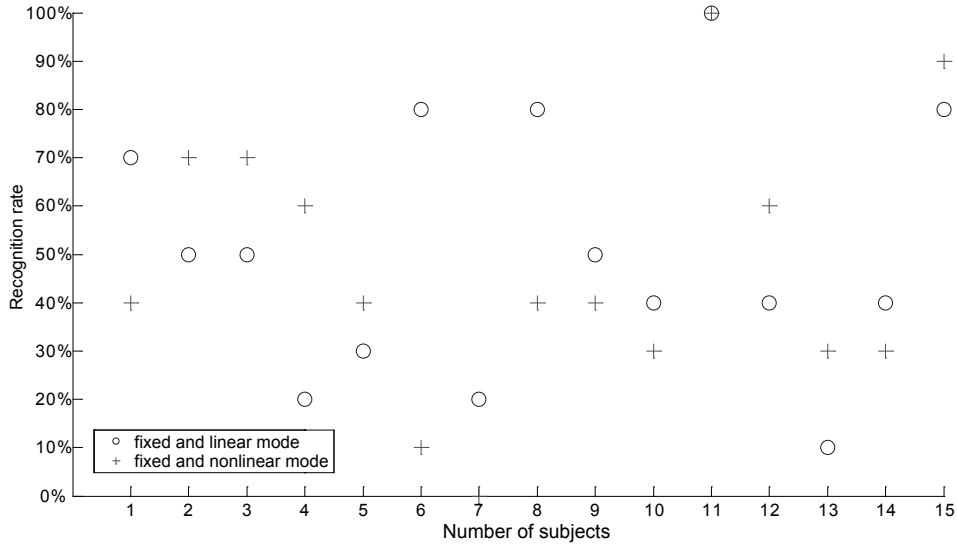


Fig. 12. Recognition rates of all subjects in both fixed and linear mode, and fixed and nonlinear mode.

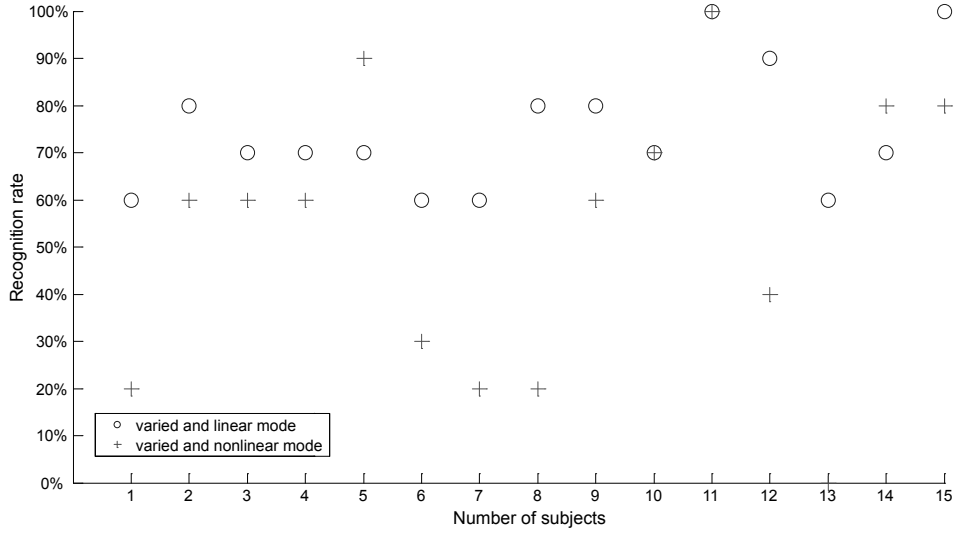


Fig. 13. Recognition rates of all subjects in varied and linear mode, and varied and nonlinear mode.

5. Conclusions and Future Work

In this paper, we proposed a vibrotactile method to express the 3D shape of an image. The front projected image of an object was processed by the SFS method. The 3D shape of the object was extracted. Then, the 3D contour map was simplified with specific height contours. Finally, the height contours were linearly mapped onto the motor arrays, and motors in the position of contours were actuated with the corresponding intensity and sequence. The marked differences in the vibration parameters make it easy to identify the difference in the height map with the vibrating stimulus. The experiments were conducted to prove the effectiveness of the method.

In this system, common vibration motors that are frequently used in mobile phones are adopted to build the vibrotactile array. The motors have big dimensions. To avoid the strong interference between motors, space must be maintained. All these reasons limit the density of the motor array; thus, it is difficult to use the system to display the 3D shapes of complex objects. In order to express more complex shapes and improve the perception using this method, our future work will focus on two aspects: (1) expansion of the size of the motor array, reduce the size of the motor, and improve the control accuracy, and (2) selection of more sets of parameters and conduct more experiments to obtain the optimal vibrotactile mode.

Acknowledgements

The research reported in this paper was carried out at the Robotic Sensor and Control Laboratory, School of Instrument Science and Engineering, Southeast University, Nanjing, Jiangsu, China.

This research was supported in part by the Natural Science Foundation of China under Grant Nos. 60905045 and 61105075, Natural Science Foundation of Jiangsu Province under Grant No. BK2011254, and Program for New Century Excellent Talents in University under No. Grant NCET-10-0330.

References

1 Y. Eriksson: How to Make Tactile Pictures Understandable to the Blind Reader, http://archive.ifla.org/IV/ifla65/65ye-e.htm (accessed on 2010-07-15).
2 T. K. Ferris and N. Sarter: IEEE Trans. Haptic **3** (2010) 199.
3 I. Yasushi, W. Kazufumi and F. Shuichi: IEEE Comput. Graphics Appl. **17** (1997) 53.
4 T. P. Way and K. E. Barner: IEEE Trans. Rehabil. Eng. **5** (1997) 81.
5 T. P. Way and K. E. Barner: IEEE Trans. Rehabil. Eng. **5** (1997) 95.
6 C.-G. Jiang, K. Uchida and H. Sawada: Proceedings of the 2011 International Conference on Advanced Mechatronic Systems, eds. M. Deng, D. Y. Wang and H. Wang (IEEE, Piscataway, NJ, 2011) p. 570.
7 C. R. Wagner, S. J. Lederman and R. D. Howe: 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2002, eds. H. Z. Tan and A. M. Okamura (IEEE, Piscataway, NJ, 2002) p. 354.
8 P. Bach-y-Rita, C. C. Collins, F. A. Saunders, B. White and L. Scadden: Nature **221** (1969) 963.

9 J. C. Bliss, M. H. Katcher, C. H. Rogers and R. P. Shepard: IEEE Trans. Man Mach. Syst. **11** (1970) 58.
10 K. Kaczmarek, P. Bach-y-Rita, W. J. Thompkins and J. G. Webster: IEEE Trans. Biomed. Eng. **32** (1985) 602.
11 M. Kurze: Conference on Human Factors in Computing Systems, ed. S. Pemberton (ACM, Atlanta, 1997) p. 22.
12 N. G. Bourbakis and A. Klinger: Pattern Recognit. **22** (1988) 317.
13 C. N. Lin: The 28th IEEE EMBS Annual International Conference, eds. A. Hielscher and J. A. Leehan (IEEE, Piscataway, NJ, 2006) p. 5261.
14 D. Dakopoulogs, S. K. Boddhu and N. Bourbakis: Proceedings of the 7th IEEE International Conference on Bioinformatics and Bioengineering, ed. J. Y. Yang (IEEE, Piscataway, NJ, 2007) p. 930.
15 S. E. Krufka, K. E. Barner and T. C. Aysal: IEEE Trans. Neural Syst. Rehabil. Eng. **15** (2007) 310.
16 H. Shu and Y. X. Zhang: Methods of Psychology Research — Design of Experiments and Data Analysis, ed. D. L. Peng (People's Education Press, Beijing, 2008) Chap. 6.
17 F. J. Gravetter and L. B. Wallnau: Statistics for the Behavioral Sciences, eds. A. M. Wang and Y. Li (China Light Industry Press, Beijing, 2008) Chap. 13.
18 E. R. Kandel, J. H. Schwartz and T. M. Jessell: Principles of Neural Science, 4th ed., eds. S. Mack and J. Dodd (McGraw-Hill, New York, 2000) p. 598.
19 L. A. Jones and N. B. Sarter: Hum. Factors **50** (2008) 90.
20 A. M. Bruckstein: Comput. Vision Graphics Image Process. **44** (1988) 139.
21 V. Pradeep, G. Medioni and J. Weiland: 32nd Annual International Conference of the IEEE EMBS, ed. R. Armentano (IEEE, Piscataway, 2010) p. 6233.
22 P. S. Tsai and M. Shah: Image Vision Comput. **12** (1994) 487.
23 D. S. Alles: IEEE Trans. Man Mach. Syst. **11** (1970) 85.
24 G. v. Békésy: J. Acoust. Soc. Am. **29** (1957) 489.
25 Z. Sheng, S. Q. Xie and C. Y. Pan: Probability Theory and Mathematical Statistics, 4th ed., eds. L. Wang and Y. C. Li (Higher Education Press, Beijing, 2008) Chap. 8.