Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Towards a 2D Tactile Vocabulary for Navigation of Blind and Visually Impaired

Dimitrios Dakopoulos, Nikolaos Bourbakis
Assistive Technologies Research Center (ATRC)
Wright State University, Dayton, OH 45435-0001
dakopoulos@gmail.com, nikolaos.bourbakis@wright.edu

***Abstract*—In this paper we present research work towards a 2D tactile vocabulary for training visually impaired for their independent mobility. The vocabulary is associated to a 2D tactile array (vibration array), which is a part of a wearable navigation prototype (Electronic Travel Aid) called Tyflos. The vibration array is currently consisting of 16 vibrating elements arranged in a 4×4 manner. Each motor can be independently driven with square pulses of varying frequencies. The vibration array can represent dynamically the 3D space of the user's field of view: the 2D arrangement represent the x-y coordinates while the frequencies represent the z coordinate (which is the distance of an obstacle from the user). Different navigation and human factor criteria have been used to create the 2D tactile vocabulary. Finally, using the continuous feedback from the users, the goal is to balance between a minimal vocabulary for easy learning and a rich vocabulary that will still be able to represent efficiently the 3D navigation space.**

Keywords*— *Tactile vocabulary, tactile display, vibration array, blind's navigation, visually impaired, electronic travel aid, ETA.

I. INTRODUCTION

The American Foundation for the Blind (AFB) estimates that "there are approximately 10 million blind and visually impaired people in the United States" [1], [2]. The need for assistive devices for visually impaired individuals is unquestionable. The last decades a variety of portable or wearable navigation systems have been developed to assist visually impaired people during navigation in known or unknown, indoor or outdoor environments [3], [5]. There are three categories of navigation systems [4]: i) vision enhancement, ii) vision replacement, and iii) vision substitution. Vision enhancement involves input from a camera, processing of the information, and output on a visual display. In its simplest form it may be a miniature head-mounted camera with the output on a head-mounted visual display (as used in some virtual reality systems). Vision replacement involves displaying the information directly to the visual cortex of the human brain or via the optic nerve. Vision substitution is similar to vision enhancement but with the output being non-visual - typically tactual or auditory or some combination of the two. ETAs (Electronic Travel Aids) are the most popular visual substitution devices that transform information about the environment that would normally be relayed through vision into a form that can be conveyed through another sensory modality.

Our navigation prototype is called Tyflos and belongs to the category of vision substitution (ETA). One of its most important components is the 2D vibration array, which offers to the blind-user a sensation of the 3D surrounding space. While, haptic interfaces [6] have been troubling scientists from different fields since the 1960s, to our knowledge, the exploitation of such a 2D tactile interface for representing the 3D environment for mobility purposes has not been performed.

The 16 vibrating elements are capable of producing 2^{16}=65,536 vibration patterns; words, in terms of the Vibration Array Language (VAL) [4]. Experimental results in haptic devices show that the users need to some extent to learn those patterns, thus the large number of pattern would correspond to an undesirable heavy cognitive load.

This paper presents the work towards the selection of a set of patterns that will constitute the 2D tactile vocabulary, a set that the user will be able to learn and distinguish between them with. In a more formal way, we try to create the dictionary of the VAL; what are the combinations of symbols that create "valid" words. This set will be selected using criteria based on different possible navigation scenarios, safe navigation and simplicity and using two rule-generation approaches: the vertical and the horizontal. A pattern vibrotactile recognition work has been done also by Jones et. Al [7] but the patterns were simple and they had the form of directional or instructional cues. Finally, in order to match the generated patterns with one from the tactile vocabulary, we propose a pattern matching methodology using a modified Euclidean dissimilarity measure.

This paper is organized as follows. First a quick overview of the Tyflos navigation system with is presented. The next section describes the experimental set-up followed by details on the different approaches followed towards the creation of the tactile vocabulary along with the experimental results. The next section describes the proposed pattern matching methodology followed by the presentation of the real-life navigation scenarios. Finally we conclude with an overall discussion and future work.

II. OVERVIEW OF THE TYFLOS NAVIGATOR

Tyflos navigation system was conceived in the mid 90s and various prototypes have been developed. The Tyflos navigation system is consisted of 2 basic modules: the Reader and the Navigator which is an Electronic Travel Aid (ETA). The main goal for the Tyflos system is to integrate different navigation assistive technologies such as: a wireless handheld computer, cameras, range sensors, GPS sensors, microphone, natural

978-1-4244-2794-9/09/$25.00

language processor, text-to-speech device, a digital audio recorder etc and methodologies such as region based segmentation, range data conversion, fusion etc. in order to offer to the blind more independence during navigation and reading. The audio-visual input devices and the audio-tactile output devices can be worn (or carried) by the user. Data collected by the sensors is processed by the Tyflos' modules each specialized in one or more tasks. In particular, it interfaces with external sensors (such as GPS, range sensors, etc.) as well as the user, facilitating focused and personalized content delivery. The user communicates the task of interest to the mobility assistant using a multimodal interaction scheme [8].

The role of the Navigator is to capture environmental data from various sensors and give the extracted and processed content to the user with the most appropriate manner. Previous Tyflos prototypes are designed using many of the technologies mentioned above and tested yielding promising results. The latest Tyflos Navigator system prototype developed is shown in Figure 1. It consists of two cameras, an ear speaker, a microphone, a 2D vibration array vest controlled by a microprocessor and a portable computer and it integrates various software and hardware components (Fig. 2).

The stereo cameras create a depth map of the environment (which can be verified by the range sensor's output). A high-to-low resolution algorithm drops the resolution of the depth map into a low resolution keeping necessary information for navigation such as safe navigation paths and objects of interest (moving objects and people; using motion detection and face-detection methodologies). This final "image" is a representation of the 3D space and it is converted into vibration sensing on a 2D vibration array vest that is attached on the user's abdomen or chest. The element of the array that vibrates represents the direction where an object is detected and the different vibration levels represent the distance of the object. Optional audio feedback can inform the user for objects of interest or for hazardous situations.

The main advantages of the Tyflos are that is free-ears and that the use of the 2D vibration array with the variable vibration frequencies offers the user a more accurate representation of the 3D environment (including ground and head height obstacles) giving also information for distances

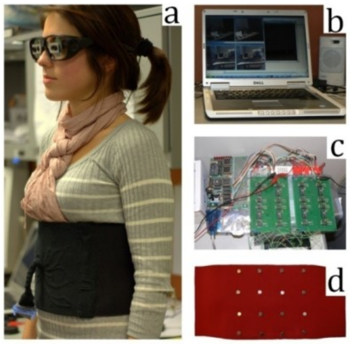


Figure 1. User wearing the 2nd Tyflos' prototype. a) stereo cameras attached on conventional eyeglasses and vibration array vest placed on the user's abdomen, b) portable computer, c) microcontroller and PCBs, d) arrangement of the 4×4 vibrating elements inside the vibration array vest.

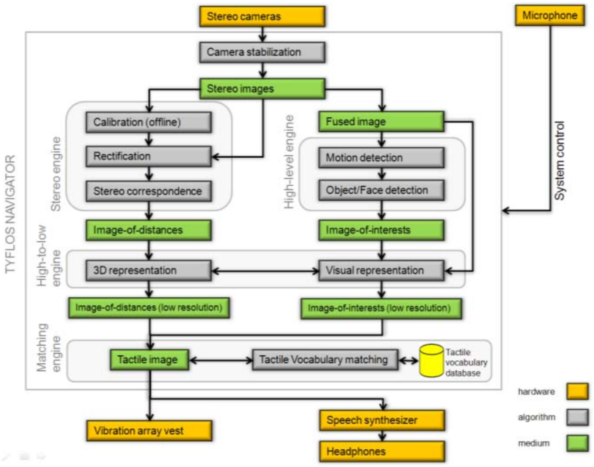


Figure 2. The hardware and software architecture of the 2nd Tyflos prototype.

III. VIBRATION ARRAY LANGUAGE

The modeling of the information sent to the user the vibration array is performed using a formal language called VAL (Vibration Array Language) [4]. The characteristic of the VAL is that it can represent any possible obstacle (or combination of obstacles) in various distances. From a C++ OO perspective the vibration array is an object that in every given moment can contain a set of symbols forming a word. Every symbol is of the form $A = (x, y, length, V[\,])$ where x, y are the coordinates of the first vibrating element on the array, *length* its length and *V[]* is an array that holds the vibration levels for every vibrating element of the symbol.

TABLE I. VARIOUS REPRESENTATION OF THE 4 VIBRATION LEVELS.

Vibration			**Distance range [m]**	**RGB representation**		**Grayscale representation**	
Level	**Freq [Hz]**	**Strength**					
0	0	None	[3,∞)	🟦	Cyan	⬛	Black
1	~2	Low	[2,3)	🟨	Yellow	▩	Dark gray
2	~4	Medium	[1,2)	🟥	Red	▨	Light gray
3	~8	High	(0,1)	🟥	Burgundy	⬜	White

Three simulated cases are provided (Fig. 3) covering different possible scenarios during navigation to demonstrate the flexibility of the VAL to describe every 3D formation of the 3D environment. Here, for better demonstrating the use of VAL, we use a 32×32 vibration array. The vibration frequencies are represented in the z-axis and correspond to 4 vibration levels that represent how far the obstacle is from the user. Also, for visualizing the vibration levels an RGB and a grayscale color is assigned for each of them (Table I).

The first case (Fig. 3 left) is a vertical obstacle which can be a standing/walking person. The VAL representation will be:

$W_1 = A_1\#A_2\#A_3$, where $A_1 = (5, 12, 25, V[\,])$, $A_2 = (5, 13, 25, V[\,])$, $A_1 = (5, 14, 25, V[\,])$ and $V = (\underbrace{3\;3\ldots3}_{\#25})$

The second case (Fig. 3 middle) includes a side and vertical obstacles which can be a person in a corridor. They are represented as:

$W_4 = W_3 \# A_5 \# A_6$, where $A_5 = (10, 25, 17, V[\,])$, $A_6 = (10, 26, 17, V[\,])$ and $V = (\underbrace{1\ 1\ \dots\ 1}_{\#17})$

The third case (Fig. 3 right) is a complex obstacle such as a workstation in an office. The VAL representation is:

$W_7 = A_1 \# A_2 \# \dots \# A_{13}$, where $A_i = (6, 5 + i, 20, V[\,])$, where $i = 1,2, \dots 13$ and $V = (\underbrace{3\ 3\ \dots\ 3}_{\#6}\ \underbrace{1\ 1\ \dots\ 1}_{\#6}\ \underbrace{2\ 2\ \dots\ 2}_{\#8})$

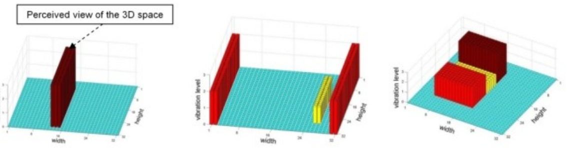


Figure 3. The 3 different navigation cases and 3D representations of the navigation space using the VAL language.

IV. EXPERIMENTAL SET-UP

A series of experiments were performed and described in the next sections for testing the user's ability to recognize symbols and patterns on the 2n prototype's 4×4 vibration array. 10 subjects were selected with their ages varying from 14 to 60. They were all normal-sighted and they have never been trained with or used the 2D vibration array or other tactile feedback devices. They were asked to wear the vest and we made sure that all the vibrating motors are in contact with their body.

Before every experiment we explain to the user the purpose of it and run a demo version of the experiment so that they will get familiar with the different patterns. During the experiment the subjects were standing so that the vest is making proper contact with their body and we tried to minimize our interaction with the subject for not distracting him/her.

V. FREQUENCIES

The vibration array has the ability to represent the 3D navigational space. The 3rd dimension, which is the distance of the subjects from the user, is represented with the different vibration levels of the motors. The correspondence of the vibration levels to vibration frequencies and distance is shown in Table I.

An experiment was performed to validate the user's capability to recognize the different vibration frequencies. Ten subjects were selected as previously (6 male, 4 female) with their ages varying from 24 to 40). Three elements were randomly selected and random vibration levels were sent to them for every trial. The subjects were asked to identify the vibration level for each element (vibration level 0 was excluded due to its simplicity).

Fig. 4 presents the identification accuracy for the 3 vibration levels. We notice that for levels 1 and 3, most of the subjects responded in a good manner showing identification accuracy over 60% and many times over 75-80%. The subjects had difficulties identifying level 2. There are two possible explanations for that. The first is that the frequencies 1 and 2 are relatively close (1Hz and 2Hz correspondingly), compared to frequency 3 (10Hz) and this was mentioned by the subjects during the experiment.

The second explanation, which also concludes our discussion, is that that none of the users were trained with the vibration array. Studies show that training of the users is a very crucial part during the development of an assistive device such as Tyflos and we strongly believe that the users trained with the system will dramatically increase their performance.

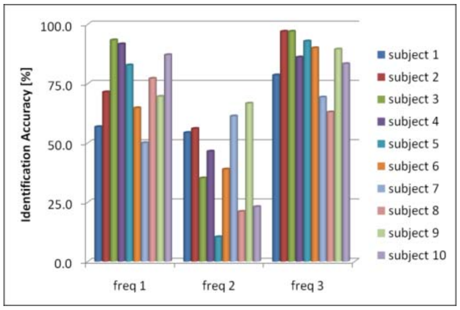


Figure 4. Identification accuracy for the different vibration levels/frequencies.

VI. VERTICAL RULES

During navigation the most important information is whether there is an open path to navigate and this corresponds to a specific direction (x-coordinates) but the major advantage of the two-dimensional array, compared to the one-dimensional, is that it can inform the user about how high or low an object is (y-coordinate). Thus, if the path is not open then the user can be informed for the position of the obstacle in the y-coordinates. We argue here that this information doesn't have to be very detailed. Thus, some rules can be set to reduce the number of patterns that can appear in the vertical dimension (y):

We select 6 types of VAL column-type symbols shown in Table II. They can appear in the four different positions (i.e. the four columns of the 4×4 array) thus, the possible navigation patterns are 6^4=1296.

TABLE II. THE 6 VIBRATION SYMBOLS (ACTIVATED VIBRATING MOTORS ARE FILLED WITH BLACK).

S0 : A	S1: B	S2 : C	S3 : D	S4 : E	S5 : F
○	○	○	●	●	●
○	○	○	●	○	●
○	○	●	○	○	●
○	●	●	○	○	●

Experiments (vertical rules)

We experiment with the users' ability to recognize the different vertical symbols in different positions.

A. Position recognition experiments

For the fist set of experiments (#1 to #5) one of the symbols S1 to S5 is sent to the vibration array in random position (a, b, c or d). The user was asked to identify the correct position by naming it (a, b, c or d) and as soon as he/she identified it, a new pattern was sent in a random position. One hundred measurements/trials were performed for each experiment. The results are shown in Table III.

B. *Symbol and symbol/position recognition experiments*

In experiment #6 (Table III), a random symbol (S0 to S5) is sent to position "a" and the user was asked to identify the correct symbol. In the final experiment #7, the user was sent a random symbol in a random position and the user was asked to identify the correct symbol and the correct position. 300 measurements/trials were taken for each experiment.

TABLE III. EXPERIMENT #1 TO #5 FOR POSITION IDENTIFICATION ACCURACY FOR THE DIFFERENT SYMBOLS (B, C, D, E, F) AND EXPERIEMENT #6: SYMBOLS IDENTIFICATION ACCUARCY

	Experiments #1 to #5					**Experiment #6**
	Position accuracy [%] for different patterns					**Symbols identification accuracy [%]**
subject	**B**	**C**	**D**	**E**	**F**	
#1	99.0	100.0	98.0	100.0	100.0	95.0
#2	99.0	100.0	100.0	99.0	100.0	92.0
#3	99.0	99.0	97.0	98.0	98.0	79.3
#4	97.0	100.0	98.0	85.0	98.0	77.3
#5	100.0	100.0	98.0	100.0	99.0	83.3
#6	96.0	100.0	98.0	98.0	100.0	88.0
#7	99.0	99.0	97.0	97.0	99.0	89.3
#8	96.0	100.0	99.0	92.0	95.0	66.7
#9	99.0	97.0	96.0	100.0	96.0	82.7
#10	100.0	100.0	98.0	94.0	99.0	90.7

C. *Discussion on vertical rules*

Table III shows that the users had at least **92% accuracy in identifying the position/direction of a predefined symbol**. On the contrary in experiment #7, were the symbol was random, those percentages are considerably smaller (Fig. 5). A possible explanation for this difference is

- The users didn't only have to identify the patterns and/or positions of symbols but they also had to say corresponding letter for each one of them. This requires more thinking and so possible mistakes. Indeed, many subjects said that that many times they were confused and used they were using the wrong letter to describe a pattern. Thus in experiment 7, that symbols and positions had to be identified, the possibility for saying the incorrect letter was higher.
- *Fatigue*; many subjects complained that they got tired, especially during the final experiment which is also the longest. This can probably result to more incorrect letter selection (as discussed before).

On the contrary, Table III and Fig. 6 show that they had some problems identifying the correct symbol. The confusion matrices (Table IV and Table V) show that the major problem was distinguishing if there is one or two consecutive vibrators active; difficulties identifying the symbol B and D where most of the times they were incorrectly identified as C and E accordingly (28.3% and 30.2% misidentification).

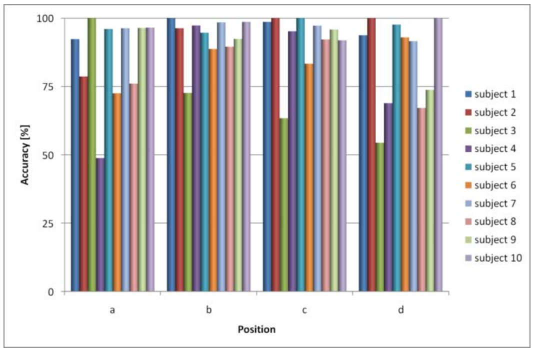


Figure 5. Experiment #7: User's accuracy in identifying the correct position of a random symbol.

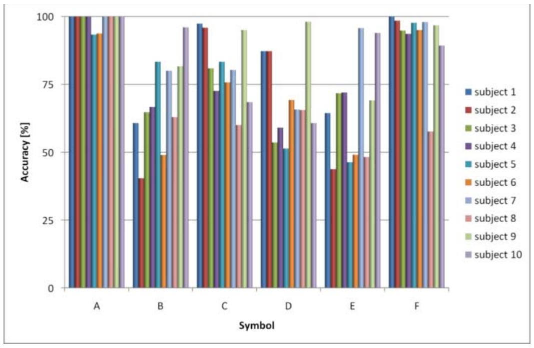


Figure 6. Experiment #7: Users' accuracy in identifying the different symbols in random positions.

TABLE IV. EXPERIMENT #6: SYMBOLS' CONFUSION MATRIX IN PERCENTAGES (AVERAGE FROM THE 10 SUBJECTS).

	Guessed symbol					
Actual symbol	**A**	**B**	**C**	**D**	**E**	**F**
A	100.0	0.0	0.0	0.0	0.0	0.0
B	3.3	85.2	**10.1**	0.5	0.9	0.0
C	0.0	9.9	83.2	3.7	0.2	3.1
D	0.0	2.2	4.4	76.3	**11.6**	5.5
E	5.1	0.3	1.0	**12.4**	80.8	0.4
F	0.4	0.2	9.5	**10.4**	0.4	79.1

TABLE V. EXPERIMENT #7: SYMBOLS' CONFUSION MATRIX IN PERCENTAGES (AVERAGE FROM THE 10 SUBJECTS).

	Guessed pattern					
Actual pattern	**A**	**B**	**C**	**D**	**E**	**F**
A	100.0	0.0	0.0	0.0	0.0	0.0
B	0.7	68.6	**28.3**	0.9	0.9	0.6
C	0.0	**10.4**	81.2	2.6	0.4	5.5
D	0.2	0.3	2.4	69.8	9.1	**18.3**
E	0.2	1.3	1.8	**30.2**	65.7	0.7
F	0.0	0.3	3.4	3.8	0.2	92.3

From bibliography [9]-[13] we know that the spatial acuity on the torso is higher than the one that our array has. As far as navigation purposes, this confusion between symbols B-C and D-E is not of major concern since the subject can still identify the low-obstacle (for example if there was a confusion between B and F then it will be more important since the symbols represent different situations i.e. low-obstacle and tall obstacle. A possible explanation for that misidentification is that the vibrators were not making proper contact with the user; the abdomen area is not uniform so not all the vibrators were placed as firmly. Despite the problem discussed above, overall the accuracy percentages are still high enough, considering that all subjects did not receive any training.

VII. HORIZONTAL RULES

In this approach, the rules are set by comparing a symbol with the symbols next to it. The idea is that the information that a symbol carries can be correlated with the neighboring symbols by emphasizing or de-emphasizing it. For example, if a large ground object is between two tall obstacles, it can be de-emphasized and be considered as a small obstacle, keeping its nature as a ground obstacle but emphasizing the nature of the tall ones.

The six horizontal rules:

1. If S1 has S2 on one side, AND the other side is S2/3/5/B then S1 will be transformed to S2 (Note: SB is the pseudo-symbol of the border of the array).
2. If S4 has S3 on one side, AND the other side is S2/3/5/B, then S4 will be transformed to S3.
3. If S2/3 has S5 on one side and S5/B on the other, then S2 will be transformed to S1 and S3 to S4.
4. If S2 has S3/4 on one side then S2 will be transformed to S1 and the side S3/4 to S4
5. If S3 has S1/2 on one side then S3 will transformed to S4 and the side S2 to S1.
6. If S1/4 is next to S0 then transform S1/4 to S2/3.

Now, by applying the above vertical and horizontal rules we reach a tactile vocabulary of 298 words which is 0.45% of the initial 65,536 patterns.

VIII. PATTERN MATCHING

As shown in the software and architecture (Fig. 2) the final step before sending the pattern to the vibration array is a pattern matching. This is because the output of the high-to-low methodologies doesn't necessarily match with one of the patterns from the VAL vocabulary, so we need a methodology that maps the low resolution images to words from the VAL vocabulary.

A. A modified dissimilarity measure

Various similarity and dissimilarity measures have been used for pattern recognition/matching applications [14]. The Euclidean distance is a fundamental dissimilarity measure. In the case of grayscale images, the Euclidean distance of two pixels is $d = \sqrt{p_1^2 - p_2^2}$ where p_1 is the grayscale value of pixels 1 and p_2 the grayscale value of pixel two. Larger distance means larger difference of intensity/color which means high dissimilarity of the pixels.

We can define the basic Euclidean dissimilarity measure between two images A and B of the same dimension m×n:

$$D = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \sqrt{\left(\mathrm{A_{ij}} - \mathrm{B_{ij}}\right)^2} = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \left|\mathrm{A_{ij}} - \mathrm{B_{ij}}\right| \quad (1)$$

where A_{ij} and B_{ij} are the values of pixels (i,j) of images A and B correspondingly.

Two identical images will results in D=0 which is zero dissimilarity. The larger the D, the more different the images are but this measure doesn't incorporate any pixel's spatial distribution and in our case, this is crucial so a modification is necessary.

The Vibration Array Language is based on vertical patterns, which correspond to the different vibration symbols and in terms of navigation to the directions of the different navigation paths. Thus, a solution for a better pattern matching would be first to match the columns of the image with the available VAL language symbols: a parameter λ is introduced to modify the Euclidean distance dissimilarity measure of Eq.1 where now the new measure will be called D_t where t stands for "tactile":

$$D_t = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \left|\mathrm{I_{ij}} - \mathrm{P_{ij}}\right| (1 + \lambda) \quad (2)$$

where I is the initial image, P the pattern with which I is compared and

$$\lambda = \begin{cases} \frac{1}{\sqrt{m \cdot n}}, & \text{if } \mathrm{I_{ij}} > 0 \wedge \mathrm{P_{ij}} = 0 \\ 0, & \text{in any other case} \end{cases} \quad (3)$$

where m and n are the dimensions of the images.

λ reflects if the "active" pixels are matched correctly. Active pixels we call the pixels that carry distance information, excluding the pixels that correspond to greater than the maximum distance (i.e. no-vibration pixels). If an active pixel is not represented in the pattern then the dissimilarity is increased. For larger images the dissimilarity is increased less because the pixels carry less spatial information since the pixels correspond to smaller environmental space.

Table VI shows an example of how the modified dissimilarity measure applies in comparison with the standard Euclidian dissimilarity brings two possible matches: symbol S3 and S5 with dissimilarity 1. For a 4×1 image λ=0.5 and the modified dissimilarity matches only with S5. The incorporation of spatial characteristics on the new measure is evident because the new dissimilarities are more dispersed e.g. S2 has smaller dissimilarity with the image than S0 because it has less active pixel incorrect matches. From a navigation point of view symbol S5 represents better the image because represents better the active pixels (3 out of 3). Finally symbol S2 matches better than S0 because it represents at least one of the three active pixels.

Fig. 7 presents the outline of the proposed pattern matching methodology: The initial image after the high-to-low methodologies is first updated using the modified dissimilarity measure so that its columns correspond to valid language symbols. The new image is mapped to a pattern with the smaller dissimilarity value. Finally the selected pattern is updated so that the distance information corresponds to the initial image; active pixels in the matched pattern inherit the distance information from the corresponding pixels in the initial image. If the initial image pixel is not active then it inherits it from the closest active pixel of the same column. If there are more than one pixels then the pixel that is in the same half (bottom or up) is selected.

TABLE VI. EXAMPLE OF SYMBOL MATCHING USING THE STANDARD AND MODIFIED DISSIMILARITIES. THE BLACK CIRCLES CORRESPOND TO ACTIVE PIXELS I.E. CORRESPONDING MOTORS VIBRATE.

	Possible matching symbols					
Image column	**S0**	**S1**	**S2**	**S3**	**S4**	**S5**
● ● ● ○	○ ○ ○ ○	○ ○ ○ ●	○ ○ ● ●	● ● ○ ○	● ○ ○ ○	● ● ● ●
Standard dissimilarity	3	4	3	**1**	2	**1**
Modified dissimilarity	4.5	5.5	4	**1.5**	3	**1**

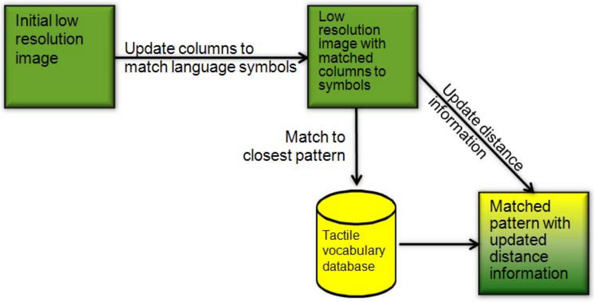


Figure 7. Pattern matching methodology flow.

IX. NAVIGATION SCENARIOS

The last experimental part involves testing with real navigation scenarios. 11 videos were recorded from inside the school of engineering. The videos include different possible navigation scenarios, including moving or static people, over-hanging obstacles, low height obstacles, doors etc. The captures from the left and right cameras were processed off-line (stereo, high-to-low, pattern matching) and a selection of the final vocabulary frames (179 in total) from each scenario was presented to the users. Some characteristic frames including the left/right camera, the high resolution disparity map and the mapped tactile word are shown in Fig. 9.

The users were asked to do a quick sketch of the vibration pattern sent to them. This includes position of open paths and obstacles with distance/frequency information.

Discussion on navigation scenarios

The experimental results from the navigation scenarios are presented in Table X and Table XI and they give us some important feedback for the evaluation and further development of the Tyflos Navigation prototype but before further discussion we want to emphasize on two points

- The users did not receive any training with our prototype or any other system with tactile interface.
- The navigation scenarios are captured from real-life indoor environments; not synthetic or simulated.

The most important information that a visually impaired user needs to know while navigating, is whether there are open safe navigation paths and in which direction. For our current prototype a safe navigation path is determined by a column that is free of vibrations. Table X shows that in **97.6% of the frames the users were able to detect a fully open frame (scene without obstacles)** which is a relatively easy task but important to notice. We also see that **in 91.5% of the cases the users were able to identify correctly the open paths in images containing obstacles.**

TABLE VII. CORRECT IDENTIFICATION OF: PATTERNS WITHOUT VIBRATIONS; OPEN PATHS WITHIN PATTERNS; NON-ZERO FREQUENCIES.

subject	empty images [%]	open paths [%]	non-zero freqs [%]
0	97.0	87.4	26.6
1	100.0	96.4	34.4
2	97.0	93.2	32.5
3	97.0	86.5	27.6
4	97.0	93.7	35.2
5	100.0	89.6	18.5
6	90.9	95.0	26.7
7	97.0	91.4	30.9
8	100.0	86.9	42.6
9	100.0	94.6	30.5
Average	***97.6***	***91.5***	***30.6***

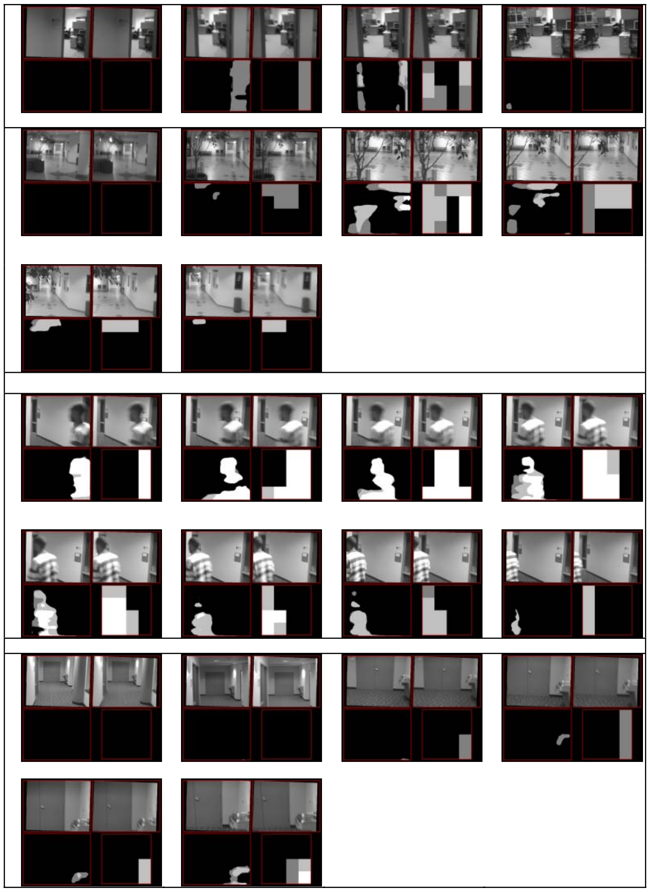
TABLE VIII. AVERAGE CONFUSION MATRIX FOR THE DIFFERENT FREQUENCIES.

	guess			
actual	**0**	**1**	**2**	**3**
0	***91.6***	4.0	3.2	1.3
1	36.8	***27.6***	31.8	3.8
2	34.1	15.0	***35.8***	15.1
3	30.4	4.6	29.3	***35.7***

When it comes to obstacle identification the users did not perform as good. Their average identification accuracy for vibration level different that zero was 30.6% (Table VII). The confusion matrix (Table VIII) shows that the users misidentified vibration levels with close frequencies. For example when the level was 3, 35.7% of the users guessed correctly level 3 but 29.3% guessed level 2, while a small 4.6% guessed level 1.

We can also notice a misidentification of non-zero levels as zeros with 36.8%, 34.1% and 30.4% correspondingly. The most possible explanation is that the users misidentified the symbol: as seen in the previous experiments, the confusion matrix of symbols (Table V) shows that the users misidentified the symbols with similar percentages of up to 30.2%.

Figure 8. Examples of selected frames. Top left and top right are left and right high resolution camera frames. Bottom left is the high resolution disparity map. Bottom left is the (4×4; low resolution) tactile vocabulary pattern sent to the vibration array. Black pixels correspond to no vibrations; dark gray to vibration frequency 1; light gray to frequency 3 and white to frequency 3 (the highest).



X. DISCUSSION

In this paper we present work towards the creation of a 2D tactile vocabulary that will help blind and visually impaired for independent mobility as part of the Tyflos navigation prototype. A vocabulary of 298 was created using different approaches and the experimental results are promising. The subjects, although they have not received any training with the prototype tactile array, were possible to detect safe navigation paths on videos from real-life indoor scenarios with an average accuracy of 92.5%. The feedback received during the conduction of the experiments will be the key towards the refinement of the vocabulary. For example

- More tests can be performed on the vertical and horizontal rules.
- Change the frequency levels to maximize perception (e.g. many users complained that level 1 and 2 are very similar to distinguish)
- Changes of the hardware (e.g. experiment with synchronized vibration frequencies)
- Changes in the software methodologies (high-to-low mapping, pattern matching etc)

Tyflos, as an assistive prototype, is an evolving system and it has to adapt to the needs of the users made possible only through continuation of the experiments with emphasis on experiments including blind and visually impaired users which are currently being performed.

ACKNOWLEDGEMENT

This work is partially supported by an NSF and AIIS Inc grant. Also, the authors wish to thank CHI, PDC, CSCW volunteers, all publications support and staff who wrote and provided helpful comments on previous versions of this document.

REFERENCES

[1] National Federation of the Blind: http://www.nfb.org/

[2] American Foundation for the Blind: http://www.afb.org/

[3] Blasch, B.B., Wiener, W.R. and Welsh R.L., *Foundations of orientation and mobility,* 2nd edition, AFB Press, 1997.

[4] Dakopoulos, D., Boddhu, S. K., Bourbakis, N. A 2D Vibration Array as an Assistive Device for Visually Impaired, *7th IEEE International Conference on Bioinformatics and Bioengineering*, *Vol. 1, 14-17* October 2007, Boston, MA, USA, pp.930-937.

[5] Dakopoulos, D., N. Bourbakis, N. Wearable Obstacle Avoidance Electronic Travel Aids: a Survey, *IEEE Trans. Systems, Man and Cybernetics,* to appear 2009.

[6] Chouvardas, V.G., Miliou, A.N., Hatalis, M.K. Tactile Displays: Overview and recent advances, *Displays, Vol. 29, 2008, pp. 185-194.*

[7] Jones, L.A., Lockyer, B., Piateski, E. Tactile display and vibrotactile pattern recognition, *Advanced Robotics,* Vol. 20, No.12*,* 2006, pp.1359-1374. N. Bourbakis, R. Keefer, D. Dakopoulos, A. Esposito, "A Multimodal Interaction Scheme Between a Blind User and the Tyflos Assistive Prototype", 20th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2008), Dayton, OH, USA, 3-5 November 2008.

[8] N. Bourbakis, R. Keefer, D. Dakopoulos, A. Esposito, "A Multimodal Interaction Scheme Between a Blind User and the Tyflos Assistive Prototype", 20th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2008), Dayton, OH, USA, 3-5 November 2008.

[9] Van Erp, J.B.F. Guidelines for the Use of Vibro-Tactile displays in Human Computer Interaction, *Eurohaptics 2002*, Edinburgh, Ed. by S. A. Wall et al., pp. 18-22.

[10] Van Erp, J.B.F.. Vibrotactile Spatial Acuity on the Torso: Effects of Location and Timing Parameters", 1st Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2005, pp. 80-85.

[11] Cholewiak, W.R., Collins, A.A., Brill, J.C. Spatial Factors in Vibrotactile Pattern Perception, *Eurohaptics 2001*, 1-4 July 2001, Birmingham, England.

[12] Weinstein, S., Intensive and extensive aspects of tactile sensitivity as a function of body-part, sex and laterality, in: The Skin Senses, edited by D. R. Kenshalo, Springfield, C. Thomas, pp. 195-218.

[13] Johnson, K.O., Philips, J.R. Tactile spatial resolution. I. Two point discrimination, gap detection, grating resolution and letter recognition, *Journal of Neurophysiology, Vol. 46, No. 6*, 1981, pp. 1177-1191.

[14] Theodoridis, S., Koutroumbas, K., *Pattern Recognition*, 3rd Edition, Academic Press, 2006.