

A Self Organized Map Method to Classify Auditory-Color Synesthesia from Frontal Lobe Brain Blood Volume

Takashi Kaburagi, Takamasa Komura, Yosuke Kurihara

Abstract—Absolute pitch is the ability to identify a musical note without a reference tone. Training for absolute pitch often occurs in preschool education. It is necessary to clarify how well the trainee can make use of synesthesia in order to evaluate the effect of the training. To the best of our knowledge, there are no existing methods for objectively confirming whether the subject is using synesthesia. Therefore, in this study, we present a method to distinguish the use of color-auditory synesthesia from the separate use of color and audition during absolute pitch training. This method measures blood volume in the prefrontal cortex using functional Near-infrared spectroscopy (fNIRS) and assumes that the cognitive step has two parts, a non-linear step and a linear step. For the linear step, we assume a second order ordinary differential equation. For the non-linear part, it is extremely difficult, if not impossible, to create an inverse filter of such a complex system as the brain. Therefore, we apply a method based on a self-organizing map (SOM) and are guided by the available data. The presented method was tested using 15 subjects, and the estimation accuracy is reported.

Keywords—Absolute pitch, functional near-infrared spectroscopy, prefrontal cortex, synesthesia.

I. INTRODUCTION

ABSOLUTE pitch is the ability to identify a musical note without a reference tone [1]-[3]. Although absolute pitch is not essential for musical performance, training absolute pitch for children is one of the popular preschool educations. Gregersen et al. point out that absolute pitch exhibits phenotypic and genetic overlap with synesthesia [4]. Another research exhibits that using color-auditory synesthesia can effectively train absolute pitch for children [5]. They show that by using the combination of color and a major triad, children under seven years old can successfully be trained to have absolute pitch. In their method, for example, color red is trained to link to C major chord (C, E, G). To evaluate the effect of the training, clarifying how well the trainee makes use of synesthesia would help both trainer as well as trainee. Unfortunately, to the best of our knowledge, an objective method to distinguish whether a subject is using synesthesia is not widely known. In this study, a method to distinguish whether a subject is using synesthesia or not is proposed. Among the variety of synesthesia, we focus especially on color/auditory synesthesia. The presented method aims to distinguish

color-auditory synesthesia from color/auditory sensory alone during the memorization and recall of absolute pitch learning method. The blood volume of the prefrontal cortex is observed in our presented system using fNIRS. The presented system assumes that the cognitive function within the brain has two parts. The first part is a non-linear part. In this step, the input from the sensory is selected and distributed to the corresponding parts of the brain. The second step is a linear part. The step input from the non-linear part will cause the blood volume to increase. Here, we aim to construct a model to reverse engineer such non-linear and linear part of the cognitive system. For the linear part, we assume a second order ordinary differential equation. An inverse filter for such system can easily be obtained. For the non-linear part, since it is difficult to design such complex system, the system adopts a machine learning approach. A method based on SOM is utilized and the system learns itself from the available data. A preliminary experiment using 15 subjects has been conducted and the performance of the presented method is reported.

II. PROPOSED METHOD

Fig. 1 shows the flow of our method. Generally, when a brain activates a certain input from the sensory system, the cerebral blood flow increases. The amount of the blood flow increase depends on the amount of sensuous information, such as hearing, visual, and/or color- auditory synesthesia.

Here, the amount of sensuous information input is defined by $u(t)$. The input $u(t)$ passes through an unknown function $f\{u(t)\}$ which converts the sensuous input to the corresponding positions of the brain to activate. Here, $u_1(t).u_2(t) \dots u_n(t)$ indicate the information stimuli input to each part of the brain. When the input is amount of sensuous information $u_1(t).u_2(t) \dots u_n(t)$ and the output is cerebral blood flow at absolute pitch learning $x_1(t).x_2(t) \dots x_n(t)$ is assumed to be as:

$$M \frac{d^2 x_{1,2 \dots n}(t)}{dt^2} + \mu \frac{dx_{1,2 \dots n}(t)}{dt} + T x_{1,2 \dots n}(t) = T u_{1,2 \dots n}(t) \quad (1)$$

where t is continuous time, M, μ, T are coefficients.

The presented method aims to estimate $u(t)$ mentioned above from observed cerebral blood flow. When dt is a time difference and k is a discrete time, the amount of sensuous information can be expressed from the above model equation as follows: In order to represent the second order lag model, let us define two state variables $\mathbf{y}_1(t)$ and $\mathbf{y}_2(t)$ as:

T. Komura, T. Kaburagi, and Y. Kurihara are with the Department of Industrial and Systems Engineering, Aoyama Gakuin University, Sagamihara, Kanagawa 252-5258 Japan (phone: +81-42-759-6372; e-mail: kaburagi@ise.aoyama.ac.jp).

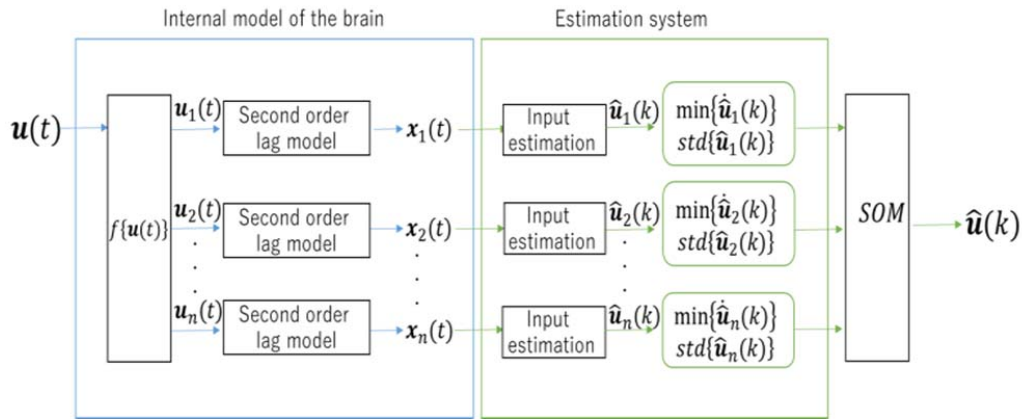


Fig. 1 Overview of the presented system

$$\begin{cases} \mathbf{y}_1(t) = \mathbf{x}_{1,2,\dots,n}(t) \\ \mathbf{y}_2(t) = \frac{d\mathbf{x}_{1,2,\dots,n}(t)}{dt} \end{cases} \quad (2)$$

The estimated input $\hat{\mathbf{u}}(k) := \hat{\mathbf{u}}_{1,2,\dots,n}(k)$ can be obtained by:

$$\hat{\mathbf{u}}(k-1) = \frac{1}{a_3} \mathbf{y}_1(k) - \frac{a_1}{a_3} \mathbf{y}_1(k-1) - \frac{a_2}{a_3} \mathbf{y}_2(k-1) \quad (3)$$

where

$$\begin{cases} a_1 = \left(1 - \frac{Tdt^2}{2M}\right) \\ a_2 = \left(t - \frac{\mu dt^2}{2M}\right) \\ a_3 = \frac{T}{M} \left(\frac{dt}{2} - \frac{\mu dt^3}{6M}\right) \end{cases} \quad (4)$$

Then, the concentration of oxygenated hemoglobin, which is the output, is measured by NIRS and the amount of sensuous information $\hat{u}_1(k)$ is obtained from the above model equation. From the obtained $\hat{u}_1(k)$, a minimum value of the time difference and the standard deviation *STD* are used as a feature value from the following equation to create a feature value vector.

$$STD = std\{\hat{\mathbf{u}}(k)\} = \frac{\sum[\hat{\mathbf{u}}(k) - mean\{\hat{\mathbf{u}}(k)\}]}{N} \quad (5)$$

where N is the number of data.

$$MIN = min\{\hat{\mathbf{u}}(k)\} \quad (6)$$

where $min\{\}$ gives the the minimum value, and $\hat{\mathbf{u}}(k+1) := \frac{\hat{\mathbf{u}}(k+1) - \hat{\mathbf{u}}(k)}{\Delta t}$ is the differential input.

A self-organized map (SOM) based on this feature quantity vector is created. A feature vector where input sensory is unknown is also mapped to the same two-dimensional map. The system estimates if the subject is using color-auditory synesthesia or not by this map.

III. EXPERIMENT

In this experiment, the subjects were divided into three groups; a) color only b) auditory only, and c) color-auditory synthesis. Each group was given a task to memorize and recall

their sensory. Cerebral blood flow was monitored during the task. Using the presented model, we judged whether colored-hearing synesthesia was used during absolute pitch learning, and the validity of the presented method was studied.

The chord used in this experiment is 1. CEG, 2. CFA, 3. CFG. CEG is a major triad with root position as C. CFA is a major triad with root position as C, first expansion. CFG is a major triad with root position as C. CFA is a major triad with root position as C. The colors (R: G: B) used in this experiment are red (255: 0: 0), yellow (255: 255: 0), and blue (0: 112: 192). The correspondence of chord and color is 1. Red (255: 0: 0) to CEG, 2. (255: 255: 0) to CFA, 3. Blue to CFG (0: 112: 192).

In this experiment, there are two phases; namely the memorization phase and recall phase. At 30 seconds rests and when the screen of the video is blank, participants watch + mark on the center of the screen. One flow of this experiment is memorization (first half) → recollection (first half) → memorization (second half) → recollection (second half). The interval of flows is around 1 minute. The order of tasks in both phases is arranged randomly in order not to remember based on the order. Participants receive the explanation of the flow of this experiment and concrete explanation of what to answer in advance. These procedures are shown in Figs. 2 and 3.

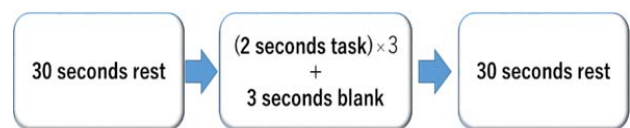


Fig. 2 Flow of experiment at the time of memorization

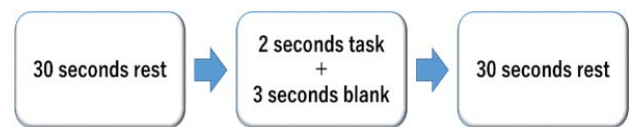


Fig. 3 Flow of experiment at the time of recall

Each group of subjects was given one of the three kinds of tasks.

The group using color only was given a visual task. Three types of colors were randomly presented for two seconds at a time. The subjects were to memorize the presented color itself.

In the recall phase, the subjects were presented a color for 2 seconds followed by a 3 seconds of a blank. During this 3 second period, the subjects are to answer if the color was previously presented in the memorization task or not.

The group using auditory only was given a hearing task. The chord sounded three times in 2 seconds randomly. The subjects were to memorize the presented chord itself. In the recall phase, the subjects were presented a chord for 2 seconds followed by a 3 second of silence. During this 3 seconds period, the subjects are to answer if the chord was previously presented in the memorization task or not.

The group using color-auditory synthesis was given a task of colored-hearing synesthesia. The subject was presented a color on a screen while a chord sounded three times in 2 seconds randomly. The subjects were to memorize the combination of the presented chord and color. In the recall phase, the subjects were presented a combination of a chord and color for 2 seconds followed by a 3 second of blank and silence. During this 3 second period, the subjects are to answer if the combination of a chord and the color was previously presented in the memorization task or not.

We used the Hitachi Kokusai Yagi Solutions Inc WOT-100 (Fig. 4) as an NIRS device. The device has 10 measuring positions, channels 7 to 16. In order to focus on the auditory system, we took measurements of the right hemisphere of the brain by recording ch7-9 as shown in Fig. 5. The experiment was held in a silent room. Only subjects who agreed to informed consent participated in the experiment. The number of subjects was 15 of the age range was 19–23 years old.

The correct answer rate using SOM and SVM was evaluated using leave-one-out-cross-validation. The correct answer rate

was evaluated by dividing the number of correct responses by the total number of responses.



Fig. 4 Hitachi Kokusai Yagi Solutions Inc WOT-100

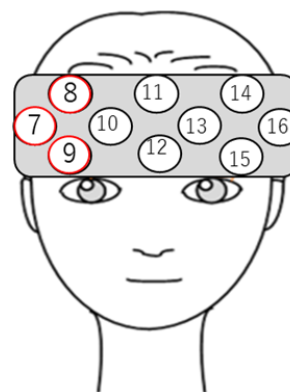


Fig. 5 The placement of the observed points

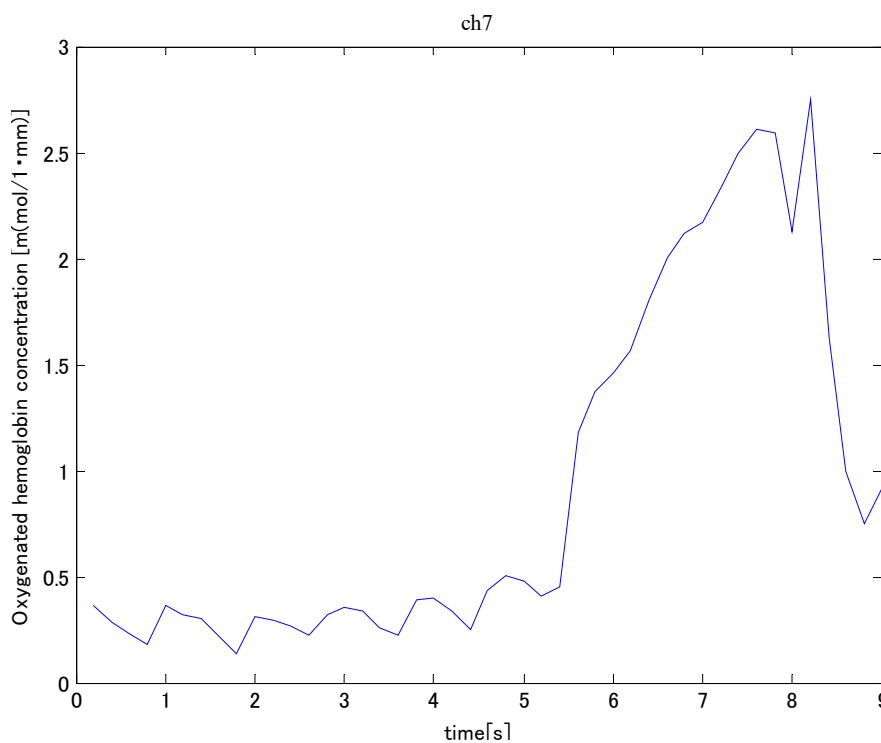


Fig. 6 A typical observed oxidized hemoglobin concentration

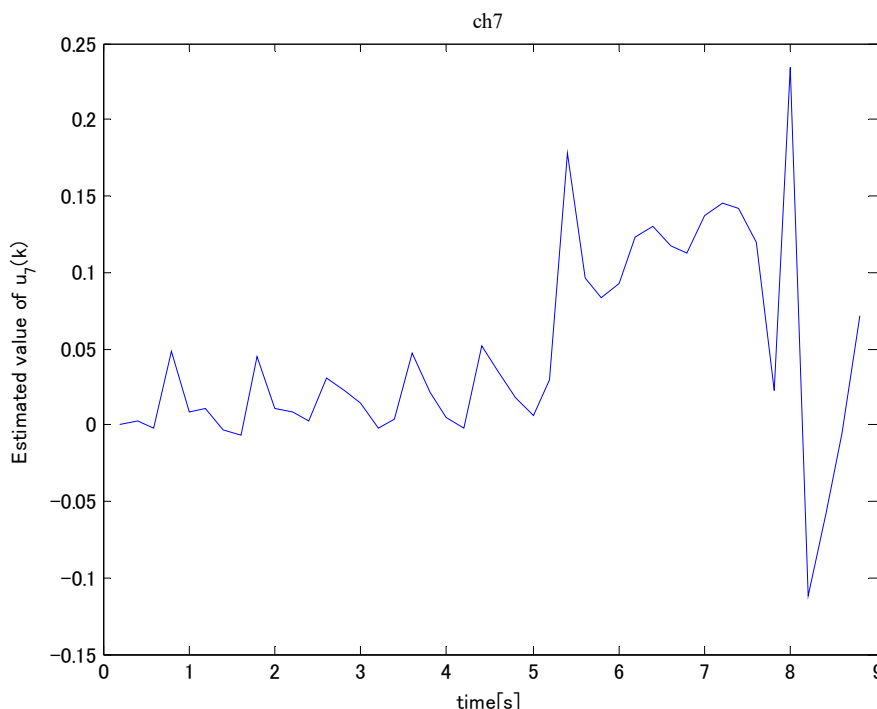


Fig. 7 A typical result of input estimation

IV. RESULTS

Fig. 6 shows the oxidized hemoglobin concentration over time, for one subject at ch7, when synesthesia was used.

A graph of the sensory amount corresponding is shown in Fig. 6 of the oxyhemoglobin concentration is shown in Fig. 7.

The result of creating a SOM using the presented feature quantity is shown in Fig. 8.

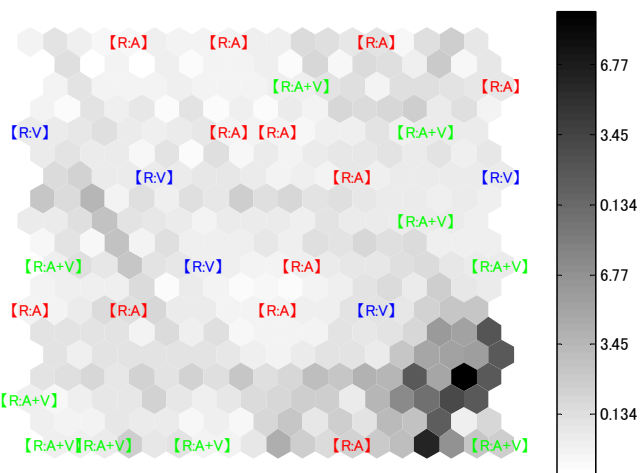


Fig. 8 The created self-organized map [R:A]: for auditory sense, [R:V]: for visual sense, and [R:A+V]: for auditory-visual synesthesia

The correct answer rate with SOM was higher than the correct answer rate of SVM for both 2 class classification and 3 class classification, 83.0% and 70.0% for 2 class discrimination, 63.3% and 50.0% for 3 class discrimination, respectively. The results are shown in Table I.

TABLE I
 CORRECT ANSWER RATE (%)

| Number of classes | SOM | SVM |
|-------------------|------|------|
| 2 classes | 83.3 | 73.3 |
| 3 classes | 63.3 | 53.3 |

The percentage incorrectly identified as synesthesia (a false positive rate) is 0.0% in the case of using SOM, 10.0% in the case of SVM, it can be seen that the SOM is more accurately discriminated. The results are shown in Table II.

TABLE II
 ERROR RATE (%)

| SOM | SVM |
|-----|------|
| 0.0 | 10.0 |

V. DISCUSSION

From the map in Fig. 8, we can see that the visual data are scattered around the hearing data. This could be attributed to the fact that the difference between the amount of information at the time of hearing and the amount of information at the time of vision is small when considered in terms of the amount of sensuous information, and it can be considered to have similar properties.

From Table II, it can be seen that in the case of SVM, visual and auditory errors were erroneously discriminated as colored-hearing synesthesia data. The data at colored-hearing synesthesia sensation are not erroneously determined as data at the time of vision or audition. It is thought that it is not good in that the opposite is considered not good in thinking that the wrong learning method is correct when considering the actual situation, therefore, we consider that the proposed analysis method is effective.

VI. CONCLUSION

In our experiment, we obtained a high correct answer rate of 83% in two class discrimination and we could discriminate synesthesia during absolute psychoacoustic learning by using the model equation of sensory quantity and cerebral blood flow rate. However, it can be said that the credibility is low because it is small, so it is a future prospect to increase the number of data.

ACKNOWLEDGMENT

A part of this work was supported by JSPS KAKENHI Grant Number JP 16K16392.

REFERENCE

- [1] Miyazaki Kenichi: "How well do we understand absolute pitch?", *Acoustical Science and Technology* Vol. 25, No. 6, pp. 270-282 (2004).
- [2] L.L. Cuddy: "Practice effects in the absolute judgment of pitch", *Journal of the Acoustical Society of America*, Vol. 43, No. 5, pp.1069-1076 (1968).
- [3] L.L. Cuddy: "Training the absolute identification of pitch", *Perception & Psychophysics*, Vol. 8, No. 5, pp.265-269 (1970).
- [4] P. K. Gregersen, E. Kowalsky, A. Lee, S. Baron-Cohen, S.E. Fisher, J. E. Asher, D. Ballard, J. Freudenberg, W. Li: "Absolute pitch exhibits phenotypic and genetic overlap with synesthesia", *Human Molecular Genetics*, Vol. 22, No. 10, pp. 2097-104 (2013).
- [5] A. Sakakibara "Why are people able to acquire absolute pitch only during early childhood?: Training age and acquisition of absolute pitch", *Japanese Journal of Educational Psychology*, Vol. 52, No4, pp. 485-496 (2004).