Hierarchical Codebook Visually Evoked Potentials for fast and flexible BCIs

Hannes Riechmann, Andrea Finke, Helge Ritter

Abstract— **Brain-Computer Interfaces provide a direct communication channel from the brain to a technical device. One major problem in state-of-the-art BCIs is their low communication speed. BCIs based on Codebook Visually Evoked Potentials (cVEP) outperform all other non-invasive approaches in terms of information transfer rate. Used only in spelling tasks so far, more flexibility with respect to stimulus structure and properties is needed. We propose using hierarchical codebook vectors together with varying color schemes to increase the stimulus flexibility. An off-line study showed that our novel hcVEP approach is capable of discriminating groups of targets after only 250ms of stimulus flickering and the final target within the group after 1s. The accuracies are 81% and 67%, respectively. Different color schemes (black/white and green/red) are equally effective.**

I. Introduction

Brain-Computer Interfaces promise to provide a communication channel for patients who have lost motor control. By directly measuring the electrical fields of the subject's brain with an EEG, all traditional neuro-muscular communication paths are circumvented. This provides a method of communication to patients who are unable to communicate by other means. But after two decades of intensive research, these devices are still quite slow for practical applications [1], [2].

Recently, the codebook visually evoked potential (cVEP) has been rediscovered [3], [4]. cVEP denotes a particular activation pattern over the primary visual cortex (PVC) triggered by a stimulus flickering according to a pseudo-random pattern. This pattern is fixed by a so-called *codebook vector*, which contains binary values that indicate whether or not a stimulus is active at a given frame. They are generated by an m-sequence which is shifted for each stimulus, firstly described in [5]. While up to 32 stimuli flicker simultaneously according to different codebook vectors, a brain wave can be detected which corresponds to the stimulus the subject gazes at. Only 1.05 s of flickering are needed to detect the target reliably ($> 90\%$ accuracy) in a spelling task. This scheme is similar to the better known SSVEP potential ([6], [7]), but cVEP allows for more simultaneous targets and does not need extra LED panels for stimulus presentation.

cVEP-based BCIs are *dependent* BCIs, because the user must necessarily fixate the flickering target with his eyes.

Fortunately, many patients suffering from neuromuscular diseases still have oculomotor control [8].

Previous studies have shown that cVEP-based BCI achieve high information transfer rates of more than $130 \, \text{bits/min}$. These approaches have in common that they utilize only spelling matrices with a fixed number of targets and are thus rather restricted in terms of possible applications beyond spelling. More complex control scenarios like wheelchair navigation or robot control demand more flexibility with respect to the structure and the properties of the stimulus presentation.

As a first step we investigate three questions. Firstly, can the codebook vector be constructed in such a way that it is possible to very rapidly (250 ms) detect the general area at which the subject gazes, while preserving the classification accuracy? This would, for example, make it possible to initiate robot movement roughly in the right direction right at the beginning of detecting the subject's intention. To accomplish this rapid detection of the general area, we split the decision of the classifier system into two parts. First, the general area and second the specific item within the area. Secondly, can a green/red flicker be detected as reliably as a black/white flicker in terms of classification accuracy? This would allow to use different colors for the stimuli wherever appropriate. And thirdly, is one of the two conditions less exhausting for the subjects? This will be measured a) by a questionnaire answered by each subject directly after the experiment. This provides a subjective measure of exhaustiveness; and b) by the Galvanic Skin Response (GSR) of the subject, which is recorded during the experiment. GSR is often used as a physiological, objective measure of the level of excitation and stress [9]. Psychologists assume that a green/red flicker is less exhausting [5], but still this has not been tested in cVEP-based BCIs.

We used 16 targets for this study and split them into 4 areas or groups, which means that both classifier stages had to solve a 4-class problem. This allows for more flexibility when constructing the codebooks. We discuss this, along with the classification scheme, in the next section.

II. Methods

A. Hierarchical Codebook and Classification

To rapidly detect the general area where the subject is gazing at, we split the 16-targets into 4 groups, according to Figure 1. Additionally, the figure shows the non-target stimuli in the surroundings which are matching neighbors for the second classification stage. A detailed discussion of the neighbor problem can be found in [5].
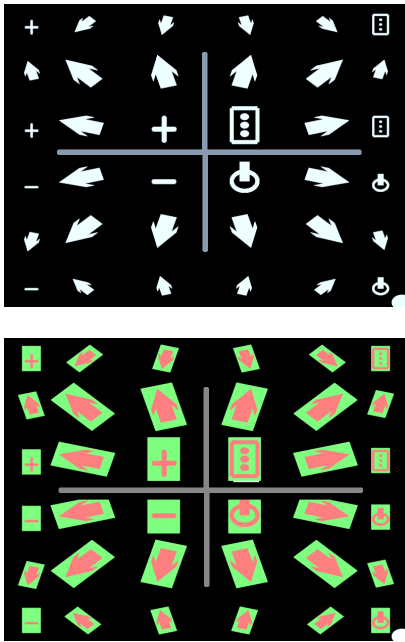
Fig. 1. Stimulus presentation of the cVEP system for navigation tasks. The 16 target symbols in the middle area show 12 navigation arrows and 4 four typical menu items: accelerate, open menu, decelerate, on/off. The thin gray lines indicate the division of the items into 4 general areas. These lines are not part of the real stimulus presentation. The surrounding stimuli are non-target stimuli which are added to have correct direct neighbors for each target stimulus. Top image shows the stimuli for the black/white condition, bottom image shows the stimuli for green/red condition.

```
[0,1,0,1,  1,0,0,1,  0,1,0,1   1,0,0,1
[0,1,0,1,  0,1,1,0,  0,1,0,1,  0,1,1,0

[1,0,1,0,  1,0,0,1,  1,0,1,0,  1,0,0,1
[1,0,1,0,  0,1,1,0,  1,0,1,0,  0,1,1,0
```

Fig. 2. Hierarchical codebook example. The four rows are four targets. In this example only 4 bits are used in each segment. Segments 1 and 3 and segments 2 and 4 are are always equal within a given target. Target 0 and 1 are in the same group and thus differ only in segment 2 and 4, where instead of shifting we inverted the code for this example. Target 2 is the first item of the second group, thus its segment 1 and 3 are inverted in relation to the two target from the first group. Target 3 accordingly is the second target of the second group.

We propose a hierarchical two-stage system for classification. The first classifier detects the group or area, the second classifier the specific item within the group. To stimulate the brain activity for this two stage system, we split the codebook. The first part of the codebook is different between groups and equal within the groups and the second part is constructed the other way round. To create the codebooks we use m-sequences according to [3]. We chose a 4-bit shift between items resulting in a 15 bit m-sequence for the first part and accordingly another 15 bit for the second part. We repeated this 30 bit codebook twice and ended up with a 60 bit codebook, schematically shown in Figure 2.

In current cVEP systems the data is shifted to obtain a one-class problem, for example in [4]. For the hierarchical codebook this is not necessary, though. On both classification stages only a 4 class classification problem has to be solved. This can be done efficiently by a multi-class Support-Vector Machine (SVM). For this study we used the SVM from LibSVM [10] with a linear kernel and the cost parameter set to 1. These parameters were found experimentally.The advantage of not shifting the data is that it allows to use arbitrary codebooks. In principle every set of binary codebook vectors would do.

Before the data is classified by the multi-class SVM, a Canonical Covariate Analysis (CCA) is applied as a spatial filter, following other state-of-the-art cVEP approaches. The CCA finds, for given data matrices X, Y, the projections A, B which maximize the correlation $p = corr(AX, BY)$. We set the EEG data as X, in the form of $(trials \cdot samples) \times channels$ and the repeated average EEG data as Y, in the form of $(k * samples) \times channels$ where k is the number of training items. After performing the CCA we used the projection matrix A as spatial filter.

B. Experimental Setup

10 healthy subjects completed 2 blocks each, one black/white, one green/red, in alternating order. Each block consisted of 480 trials. All subjects signed a written consent and were paid for their expenditure of time.

Data was acquired using a gUSBamp (Guger Technologies) amplifier. 10 channels were equipped with Ag/AgCl electrodes placed at P3, Pz, P4, PO3, POz, PO4, PO7, O1, O2, PO8 according to the extended international 10-20 system and referenced at Oz. Electrode impedances were kept below 5 kΩ. The amplifier sampled the EEG data at 600 Hz, performed high-pass filtering at 1 Hz and notch filtering at $48 - 52$ Hz.

We used 16 navigation symbols as stimuli, instead of letters for a spelling task. The stimuli were presented on a 19" CRT running at a resolution of 800x600 pixels and a frequency of 60 Hz. In the black/white condition the stimulus was rendered bright white when the codebook vector was 1 for this frame and not rendered when the codebook vector was zero. The background was completely black. In the green/red condition the background was also black. Around each stimulus, though, a green rectangle was rendered regardless of the codebook. If the codebook was 1 for the actual frame the stimulus was rendered in red onto the rectangle (see 1). Otherwise the stimulus was not rendered. Timestamps were acquired using a photo-resistor attached to the screen.

The Galvanic Skin Response was measured at the left hand to asses which condition (one black/white, one green/red) was more exhausting. Additionally, the participants were asked to fill in a questionnaire that contained questions on their subjective experience with the task.

III. RESULTS

We evaluated the study mainly with regard to the classification accuracy. To do so, the data set was split into a training and a testing set, the latter containing 160 and the former 320 items. The accuracy was measured as the
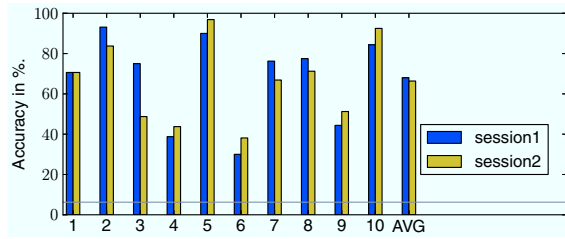
Fig. 3. Classification accuracy for the two sessions for all subjects and the average. The gray line indicates the chance level.
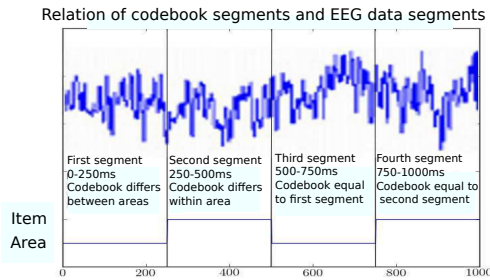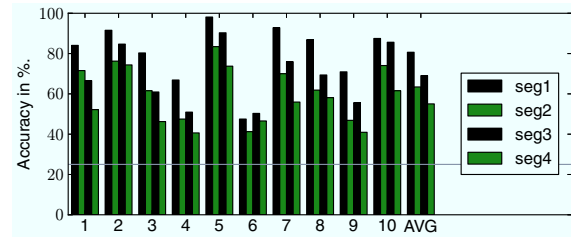

Fig. 5. Classification accuracy of the four distinct segments for all subjects and the average. The gray line indicates chance level. The first and third segment contain information about the general area, the second and fourth about the specific item within the area.


Fig. 4. The EEG data of each trial is split into four segments according to the four segments of the hierarchical codebook.

Fig. 6. Classification accuracy of the two stages separately for all subjects and the average. The gray line indicates the chance level.

percentage of correctly classified items on the testing set given an SVM trained on the training set. Training and testing was performed on a per-subject basis. Figure 3 shows the overall classification accuracies per subject and session. Three subjects achieve good classification rates (>80%) and four other subjects reach usable classification rates of about 70%. Three subjects do not achieve control over the system, however, their classification results are still way better than chance. In total, the system correctly classifies 67% of the trials on average, nevertheless, the variance is quite high. If we exclude the three non-performing subjects the average increases to 78%.

A. Hierarchical Codebook

The EEG data is split into four segments according to Figure 4 for further analysis. Classifier training and evaluation is then done for each segment seperately. The accuracies of the four distinct segments are plotted in Figure 5. The general area can be detected with an accuracy of 81% on average, using only the first 250 ms which corresponds to the first segment.

Additionally, the classifier for the general area (first stage) achieves a better classification than the classifier for the specific item (second stage). Figure 6 shows the separate accuracies of the two classifiers for the whole trials. Overall the general area is classified significantly more accurately than the specific item within the area (paired t-test $p < 0.001$). This is quite unexpected, as the two classification stages should be equally difficult. Both classifiers need to solve a 4 class problem and both parts use two codebooks segments of 15 bits each.

Figure 5 shows the individual accuracies of the data segments. Note that the respective codebooks and thus the
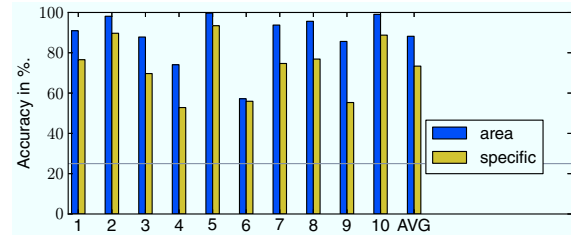
(stimulus) flickering patterns are identical for segment 1 and 3 and also for segment 2 and 4. In theory, the corresponding segments should classify equally well. In practice, however, this is not the case. Segment 1 classifies significantly better than segment 3. The same is true for segments 2 and 4. The reason for this degradation is not yet clear, but is likely related to the structure of a trial (see Fig. 4). Each segment, that is, the particular flickering pattern described by the corresponding codebook, triggers a unique activation pattern in the PVC. The exact time course of this activation is yet unknown. As we use very short segments (ca. 250 ms), the PVC activation pattern might not instantly switch from one pattern to the other upon the start of a new segment. There might be a smooth transition from one pattern to the other in terms of neural activity. Such an effect would than cause the signal of a consecutive segment to be a mixture of the true activity corresponding to the current flickering pattern and the activity corresponding to the previous pattern. In other words, a segment would contain also information on its predecessor. If so, a classifier trained on segment 2 data should be able to cope with segment 3 data, at least to a certain degree (i.e., better than chance, $> 25\%$). We evaluated this hypothesis by training a classifier only on segment 2 data and applying it to the data of segment 3. The accuracy (66%) was considerably above chance, which supports our hypothesis. On the other hand, segment 3 classifies generally better than segment 2, but should be equally effected by the mixture problem. Likely, the latter is not the only effect that is present in the data and not a sufficient explanation for the worse performance of the second classification stage. A deeper analysis taking the neurophysiological properties of cVEP PVC activation into account is needed to understand these phenomena.
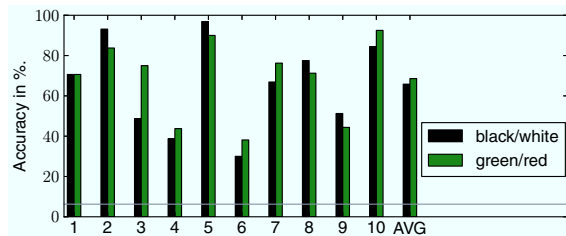
Fig. 7. Classification accuracy sorted by condition for all subjects and the average. The gray line indicates chance level.

B. Green/Red Flickering

For comparing the two stimulus color conditions, Figure 7 shows the accuracies for the two different conditions per subject. There is no large difference in classification accuracy except for subject 3. This subject was very tired during the second session, which was the black/white one. So, in terms of classification accuracy, there is on average no advantage for one condition over the other. There where small differences for some subjects and it might be valuable to evaluate this further. There might be subjects where one condition is indeed preferable, but more trials are needed to asses this in detail.

To determine the exhaustiveness of the two conditions for the subjects we evaluated the questionnaires. The study was quite exhausting and tiring for all subjects, mainly caused by the flickering. The average questionnaire score was 3.75 points out of 6 possible points for "The experiment was exhausting" and 4.13 out of 6 for "The experiment was tiring". There was no significant difference between the two color conditions. For the evaluation of the skin conduction, we compared the average measured GSR value during the trials of the two sessions for each subject. The GSR sensor data showed a reduction in skin conductance between 2% and 30% for 8 subjects from the first to the second session. One subject showed an increase of 2% and for one subject the measurement was broken. No difference related to the two conditions could be found.

IV. DISCUSSION AND CONCLUSION

The results of this study are three-fold. Firstly, we found that the codebook vector can be chosen in a way that makes it possible to very rapidly (250 ms) detect the general area at which the subject gazes, without destroying the overall classification. The classifier accuracy in detecting the general area in this short time windows was 81%. As stated in the introduction this rapid detection of the general area of the target is very useful for control of semi-autonomous devices like a BCI-controlled wheelchair. What is more, it greatly increases flexibility in constructing the codebook. There might be situations where it is useful to increase the hamming distance of the codebook vector of a specific item to all items in exchange for a decrease of the hamming distances within the other items.

Still, the total classification accuracy is lower than in other recent cVEP studies. We strongly suspect that this is mainly due to the overlapping effect between the different segments of the hierarchical codebook. Additionally, the questionnaire showed that the subjects had to make an effort to keep gazing at the targets without blinking. It is difficult to check whether all subjects were motivated to do so. This might be one reason for the high variance. Still, both session are classified roughly equally well for nearly all subjects and no degradation of accuracy within the sessions could be found. This suggests that, although the exhaustiveness is a problem for subjects' motivation and an issue for practical use, it is not a problem for the classification.

Secondly, we showed that a green/red flicker can be detected as reliably as a black/white flicker in terms of classification accuracy. The questionnaires revealed that from the subjects' point of view both conditions are quite exhausting and on average no condition is less exhausting than the other. There were differences for single subjects, though. These differences can be partly attributed to the order of the sessions. In general, the second session is perceived to be more exhausting. However, some subjects seem to prefer one color condition over the other. It might be useful to test for individual preferences before doing the study.

As the next step towards increased flexibility, we propose to include a background image and optimize the hierarchical codebook.

REFERENCES

[1] D. J. McFarland and J. R. Wolpaw, "Brain-computer interfaces for communication and control," *Commun. ACM*, vol. 54, no. 5, pp. 60–66, May 2011.

[2] O. Tonet, M. Marinelli, L. Citi, P. M. Rossini, L. Rossini, G. Megali, and P. Dario, "Defining brain-machine interface applications by matching interface performance with device requirements." *J Neurosci Methods*, vol. 167, no. 1, pp. 91–104, 2008.

[3] G. Bin, X. Gao, Y. Wang, Y. Li, B. Hong, and S. Gao, "A high-speed bci based on code modulation vep," *Journal of Neural Engineering*, vol. 8, no. 2, 2011.

[4] M. Spüler, W. Rosenstiel, and M. Bogdan, "One class svm and canonical correlation analysis increase performance in a c-vep based brain-computer interface (bci)," in *The 20 th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning - Proceedings - Bruges, Belgium from 25 to 27 April 2012*, 2012, pp. 103–108.

[5] E. E. Sutter, "The brain response interface: communication through visually-induced electrical brain responses," *J. Microcomput. Appl.*, vol. 15, no. 1, pp. 31–45, Jan. 1992.

[6] Ding, Jian, Sperling, George, Srinivasan, and Ramesh, "Attentional Modulation of SSVEP Power Depends on the Network Tagged by the Flicker Frequency," *Cerebral Cortex*, vol. 16, no. 7, pp. 1016–1029, Jul. 2006.

[7] S. Resalat, V. Saba, F. Afdideh, and A. Heidarnejad, "High-speed ssvep-based bci: Study of various frequency pairs and inter-sources distances," in *Biomedical and Health Informatics (BHI), 2012 IEEE-EMBS International Conference on*, jan. 2012, pp. 220 –223.

[8] M. Ahmadi, J.-X. Liu, T. Brännström, P. M. Andersen, P. Stål, and F. Pedrosa-Domellöf, "Human extraocular muscles in als," *Investigative Ophthalmology & Visual Science*, vol. 51, no. 7, pp. 3494–3501, 2010. [Online]. Available: http://www.iovs.org/content/51/7/3494.abstract

[9] J. T. Cacioppo, L. G. Tassinary, and G. Berntson, Eds., *Handbook of Psychophysiology*. Cambridge University Press, Mar. 2007.

[10] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011, software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.