34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Kinect-based detection of self-paced hand movements: Enhancing Functional Brain Mapping paradigms

Reinhold Scherer[1,2], Johanna Wagner[1], Günter Moitzi[1], and Gernot Müller-Putz[1]

Abstract— **Monitoring and interpreting (sub)cortical reorganization after stroke may be useful for selecting therapies and improving rehabilitation outcome. To develop computational models that predict behavioral motor improvement from changing brain activation pattern, we are currently working on the implementation of a clinically feasible experimental set-up, which enables recording high quality electroencephalography (EEG) signals during inpatient rehabilitation of upper and lower limbs. The major drawback of current experimental paradigms is the cue-guided repetitive design and the lack of functional movements. In this paper, we assess the usability of the Kinect device (Microsoft Inc., Redmond, WA, USA) for tracking self-paced hand opening and closing movements. Three able-bodied volunteers performed self-paced right hand open-close movement sequences while EEG was recorded from sensorimotor areas and electromyography (EMG) from the right arm from extensor carpi radialis and flexor carpi radialis muscles. The results of the study suggest that the Kinect device allows generation of trigger information that is comparable to the information that can be obtained from EMG.**

I. INTRODUCTION

Directed and early rehabilitation after stroke aims to promote neuroplasticity, i.e., inducing (sub)cortical reorganization for minimizing motor impairment [1], [2]. However, little is known about the interplay between therapeutic interventions, functional improvements and related changes in brain activity. We are interested in exploring this functional interaction with the aim of developing computational models that predict functional improvement from an individual's current brain activation pattern as a function of therapy. This may lead to a dynamic treatment regime for stroke rehabilitation.

In order to develop such predictive models, we need a longitudinal study design and the possibility of monitoring the brain activity of individuals during rehabilitation. Multichannel electroencephalography (EEG) has emerged as the most important non-invasive signal source for functional brain mapping (fBM) and brain-computer interfacing (BCI) in humans [3], [4]. EEG is, compared to magnetoencephalography (MEG) or functional magnetic resonance imaging (fMRI), widely available, inexpensive, compact, and offers a reasonable trade-off between temporal and spatial resolution. We are currently working on the implementation of a clinically feasible experimental set-up, which enables recording high quality EEG signals during inpatient rehabilitation of upper and lower limbs.

The major drawback of EEG is the low signal-to-noise (SNR) ratio. The low SNR and the predominant use of statistical models and machine learning algorithms for characterizing the causal relationships between behavior and brain activity patterns in fBM and BCI [5], usually requires that monitored individuals repeat a given behavioral task (cognitive or motor) a number of times. Experimental paradigms are consequently designed for recording large number of trials of stereotypical, usually isolated, movements (e.g. index finger tapping). However, the resulting time-constraints and the limited use of functional movements does not fit the needs of individuals during rehabilitation. In this case, we need an easy to use and unobtrusive system that is capable of tracking body limb movements, enabling patients to perform the task according to their own abilities and timing. Moreover, since experiments will be repeated several times, the paradigm should remain motivating. The use of game-based rehabilitation paradigms, i.e., embedding activities of daily life in an entertaining game environment, which allows adapting task difficulty to the individual's motor function repertoire, will more likely keep users engaged and compliant with the requested task [6].

Motion tracking can be realized in several ways [7]. Tracking systems are generally categorized as either non-visual tracking systems (e.g. based on inertial sensors or data gloves) or visual systems (e.g. camera). The latter can further be subdivided into marker-based and marker-free systems (conventional video camera). Each method has different merits and limitations. Please refer to [7] for a survey on different techniques. Considering our interest in game-based rehabilitation, one inexpensive, flexible and powerful marker-free visual motion tracking device is the Kinect (Microsoft Inc., Redmond, WA, USA). The Kinect motion sensing input device consists of depth sensor, color image camera (640x480 pixel resoluion at 30 Hz) and 3-D microphone, and does not require additional sensors that need to be placed on monitored individuals that may further limit their range of motion.

In this paper, as a first step towards the development of enhanced functional brain mapping paradigms, we assess the usability of the Kinect to track self-paced hand opening and closing movements (postures) and compare our Kinect-based movement onset detection algorithm with electromyography (EMG) signals recorded from finger extensor and flexor carpi radialis, and present event-related spectral perturbation (ERSP) maps [8] from EEG electrodes placed over sensori-

*This work was supported by the Land Steiermark Project bci4rehab.

[1] Reinhold Scherer, Johanna Wagner, Günther Moitzi, and Gernot Müller-Putz are with the Institute for Knowledge Discovery, Graz University of Technology, Graz, Austria.

[2] Reinhold Scherer is with the rehabilitation center Judendorf-Strassengel, Judendorf-Strassengel, Austria.

motor areas.

II. METHODS

A. Hand posture detection

Subject recognition and skeleton calculation were based on the openNI^{TM}/NITE framework (http://www.openni.org). For this study only the position of upper right extremity was tracked, i.e., shoulder, elbow and hand position. The tracking algorithm was calibrated by performing a Click-gesture.

Due to varying lightning conditions, surface characters and limited options in positioning the Kinect sensor during therapy in the clinical environment, the depth information provided by the Kinect sensors was noisy and could not be considered for hand posture detection. Hence, the implemented algorithm was based only on the red-green-blue (RGB) color information. Please note that the hand tracking was still based on the Kinect depth sensor.

The implemented hand posture detection algorithm has two assumptions. Firstly, the hand position is quasi-stationary, i.e., hand displacements are limited, and secondly, changes of interest are expected only in a close area around the estimated hand position. As region of interest (ROI) a 80×80 pixel rectangle centered around the hand position of the calculated skeleton was selected. The detection algorithm requires an initial calibration, during which reference images I_i, i.e., the RGB values in the ROI, of the number of i different hand postures P_i are acquired and stored (Fig.1). In this study, we only used hand open and hand closed postures ($i = [1, 2]$); The system, however, is scalable. Features used to characterize each posture P_i were computed by subdividing the 80×80 pixel bitmap of the ROI into groups of 2×2 pixels, and by averaging the RGB information over the 4 corresponding pixels ($\mu_{2\times2}^{P_i}$). During posture detection, in every frame the algorithm compares the $\mu_{2\times2}^{P_i}$ from the reference image I_i with $\mu_{2\times2}^{Current}$ of the current hand image $I_{Current}$. Hand posture P_i is detected if at least 90% of the $\mu_{2\times2}^{Current}$ fall in the range of $\mu_{2\times2}^{P_i} \pm 15\%$. The choice of detection parameters was made empirically from pilot experiments, where we observed that the selected parameters achieved stable detection across subjects. As long as this criterion was met, posture P_i detected events were triggered. If the criterion was not met, then no events were triggered.

To reduce false positive posture detections, a dwell time was implemented. Triggered posture events were only valid if the same event was triggered consecutively for a given period of time. We used a dwell time of 100 ms (three image frames). The corresponding hand posture detection time t_{P_i} was defined as the the time of the first occurrence of a valid event series. Fig. 1 shows images of detected hand close and hand open postures. As visual feedback, hand contours are highlighted.

Hand opening and closing movements inevitably lead to small displacements of the hand position. To compensate for such small deviations, after each P_i detection, the reference image I_i was updated according to $I_i = 0.8 \cdot I_i + I_{Current}$.

The developed detector software framework is compliant with the TOBI specifications and was

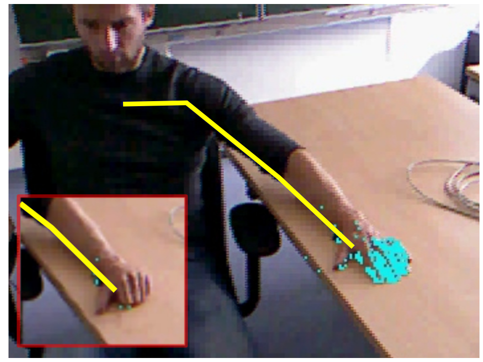

Fig. 1. Picture from the Kinect RGB camera showing a user and, superimposed, the calculated skeleton of the right upper limb (yellow line). The image is used for calibrating the hand open posture; The inlay shows the hand closed posture.

integrated into the TOBI SignalServer framework [9] (http://tools4bci.sourceforge.net/signalserver.html), which is distrubited under the GPL version 3.0. Hence, events are synchronized with biosignal data acquisition.

B. Subjects, Data Acquisition and Experimental Paradigm

Three able bodied volunteers (CE4, BX2 and CC1, all male, 26 ± 1.5 years old, all right handed) participated in this study. EEG was recorded from 6 Ag/AgCl electrodes placed over sensorimotor hand and feet areas. Electrode positions included FC_3, FC_Z, FC_4, C_3, C_z, and C_4 (Reference and ground electrodes were placed on the left and right mastoid, respectively). EEG electrodes were mounted by using the Easy Cap (Herrsching, Germany) recording cap. Electrode impedances were below 5 $k\Omega$. Additionally, EMG was recorded from the right arm from extensor carpi radialis and flexor carpi radialis muscles using standard adhesive disposable Ag/AgCl electrodes. EMG was recorded monopolarly, with reference and ground electrodes placed at the right elbow. EMG bipolar derivations were calculated offline. All signals were filtered between $0.1 - 100$ Hz (Notch at $50 Hz$) and sampled at a rate of 512 Hz (biosignal amplifier model gUSBamp from Guger Technolgies, Graz, Austria). The TOBI Signal Server framework was used to record biosignals along with the Kinect-triggered events. The Kinect sensor was placed about 1.5 m away from participants.

Subjects sat in a comfortable chair with the hand open (fingers extended) positioned on the right arm rest. Participants were asked to relax and perform self-paced single hand close-open movements at a comfortable speed with their right hand in intervals of about 30 seconds. Two experimental runs of 20 minutes each were recorded for each participant. The experimenter was visually supervising the experiment and noting the number of performed hand movements.

C. Signal Analysis

EEG data analysis was performed in Matlab 7.11.0.584 and EEGLAB 8.0.3.5b [10]. Bipolar derivations were computed according to $FC_3 - C_3$, $FC_Z - C_Z$ and $FC_4 - C_4$. The bipolar EEG time series was high pass filtered at

1 Hz (zero-phase FIR filter order 7500), movement trials were segmented, and visually inspected for muscle artifacts. Movement trials were defined as EEG segments from -4 to $+6$ seconds relative to the movement-onset time t_{onset}, i.e., the time when participants started closing their hand. The initial posture was hand open. The time t_{onset} was defined as the time of the last Kinect-triggered hand open event. Trials with muscle artifacts at any time within the trial were excluded from further analysis. This conservative selection criterion led to the rejection of about 20-30 trials per participant. From the remaining $>$ 50 trials Event-Related Spectral Perturbations (ERSPs) [8] were computed. ERSPs are calculated, computing the power spectra over a sliding latency window and normalizing these spectograms by dividing by their respective mean baseline spectra. These normalized transforms are then averaged over trials. ERSPs were computed from 6 to 60 Hz. A reference period from -4 to -2 seconds before movement onset was selected to compute the mean baseline spectrum. Significant deviations (p $\leq$ 0.01) from the average baseline spectrum were computed with a bootstrapping method [10].

The EMG was high pass filtered at 10 Hz (zero-phase FIR filter order 300), rectified, segmented from -4 to $+6$ seconds relative to movement onset and plotted for visual inspection. For plotting, a low pass filter of 10 Hz (zerophase FIR filter order 300) was applied to smooth curves.

III. RESULTS

The mean$\pm$standard deviation EMG time course, averaged over single trial EMG traces, recorded from the extensor carpi radialis (Extensor) and flexor carpi radialis (Flexor) muscles for each subject are summarized in Fig. 2. The upper six plots show the Extensor and Flexor EMG signal segmented relative to to Kinect-based movement-onset time $t_{onset} = 0$ (start of hand closing movement); The lower six plots show the mean EMG relative to the Kinect-based movement-offset trigger ($t = 0$ equals the detection of the hand open posture). For easier interpretation of Kinect-based detection and related EMG signals, only the 500 (250) ms prior and 250 (500) ms after movement onset (offset) detection, respectively, are presented. The curves show that duration and shape have a high variability between and also within subjects. The average time lags for movement onset detection for the Flexor are <250 ms. Time lags for the Extensor are $\geq$250 ms. For movement offset detection the average time lags for both Flexor and Extensor are in the range between 250 ms and 500 ms (not shown in Fig. 2).

The ERSP time-frequency maps for each bipolar channel and subject are shown in Fig. 3. The maps shows characteristic activity patterns at electrode location C_3 over the contralateral hemisphere. Characterisic patterns are a decrease in sensorimotor rhythms in the 10-13 Hz band followed by a increase in the beta band (beta-rebound) after the movement stopped.

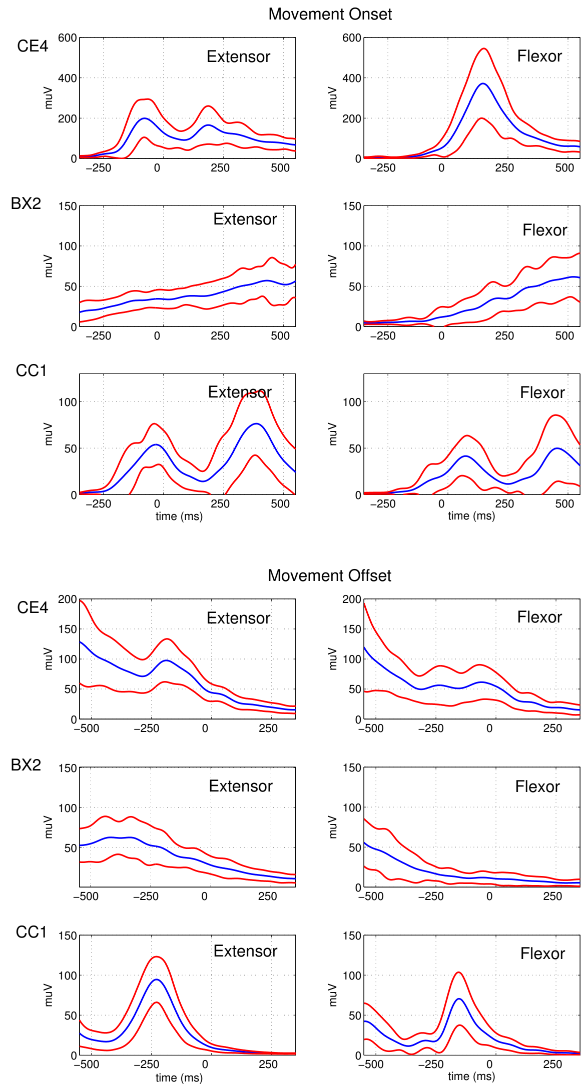

Fig. 2. Mean EMG activity of finger Extensor and Flexor carpi radialis muscles for each participant. Blue lines shows the mean and the red lines the mean$\pm$standard deviation time course. Kinect trigger for movement onset and offset detection, respectively, were issued at time $t = 0$ s.

IV. DISCUSSION

The aim of this work was to assess the usefulness of the Kinect sensor to track basic self-paced hand postures in a clinical environment. We selected a vision-based marker free system, which does not interfere with the movement abilities of monitored individuals. Placing additional sensors and calibrating the tracking system can be time-consuming and hence may be an additional burden for individuals during functional rehabilitation.

At this time, the used openNI^{TM}/NITE framework requires an initial calibration of the skeletal tracking algorithm. Provided algorithms are robust and hence, in the case of patients with limited arm movements, assistants, e.g. nursing staff, could surrogate the patient. However, alternative software development kits already support, and future releases of

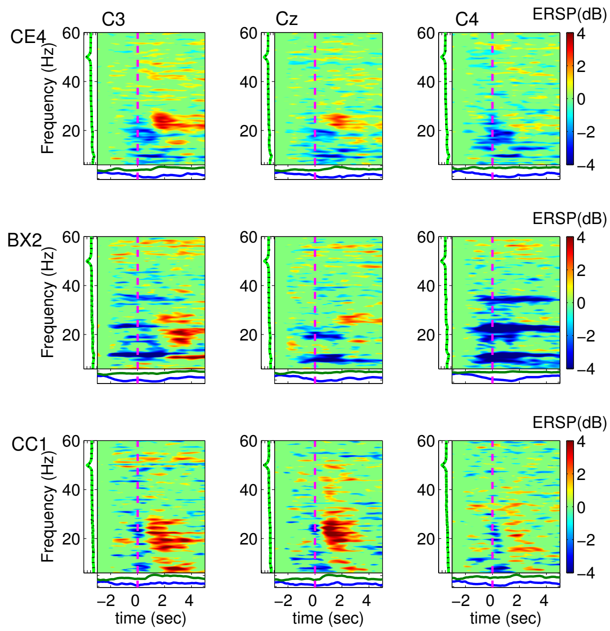

Fig. 3. ERSP time-frequency maps. Blue pixels indicate a significant power decrease compared to the reference period $[-4 - 2]$ s. Red pixels correspond to a significant power increase.

the used framework will support, calibration-less skeleton tracking.

The developed hand posture detection algorithm is not very sophisticated, however, it was able to successfully detect all hand opening and closing movement sequences (visual supervision). These good results are due on the one hand to the very basic experimental paradigm and on the other to the use of a dwell time (100 ms) for reducing false positive detections. More sophisticated methods may be required for detecting distinct hand postures while performing functional movements.

The mean time lag between the average EMG onset of the Flexor carpi radialis muscle and the Kinect-trigger was less than 250 ms for each subject. This time lag allowed us to get a satisfactory segmentation of EMG and meaningful ERSP maps. All other average time lags were ≥250 ms and have not been further researched in this study. Hence, for the selected movement task the Kinect-trigger information was more in line with the mean EMG onset of the Flexor carpi radials muscle for the hand closing movement (Fig.2). The curve shapes in Fig.2 show that each subject performed the motor task differently. Subject BX2 executed movements slower than CE4 and CC1. Subjects CE4 and BX2 performed the movement continuously, whereas CC1 showed a break (decreased EMG activity) between hand closing and opening (Flexor in Fig.2). For all subjects, however, the Extensor was activated first. Since the right arm was placed on an armrest, subjects performed a small wrist extension prior the hand open-close movement task.

The ERSP time-frequnecy maps in Fig.3 show different activity patterns for movement preparation and execution for each subject. Each subject, however, exhibited characteristic alpha band synchronization, as well as a short lasting beta band synchronization, also known as beta rebound, over the contralateral hemisphere. The ERSP maps also confirm that spectral components in the range $8 - 35\ Hz$ are most responsive. These results are in agreement with the literature (e.g. [11]). Higher frequency components did not show significant activity during movement.

Please note that the use of the Kinect device for tracking body limb movements does not make the use of EMG and other measures of muscle activity such as inertial sensors obsolete. Video based tracking will only be useful when patients have regained visible control over lost limb function. The minimum movement range depends, among others, on the distance of the Kinect sensor to the user, the image resolution and the frame rate. Studying these relationships and finding optimal parameters is left for future research.

V. CONCLUSION

We introduced the use of the Kinect device as a small, flexible, scalable, unobtrusive and inexpensive motion tracking system for functional brain mapping. These qualities and the encouraging first results represent a sound basis for further developments and improvements.

ACKNOWLEDGMENT

The Authors would like to thank the therapeutic team at the rehabilitation center Judendorf-Strassengel, Austria for their help with the system setup and the fruitful discussion during the preliminary tests, and Ian Daly for proofreading.

REFERENCES

[1] A. Turton and V. Pomeroy, “When should upper limb function be trained after stroke? evidence for and against early intervention,” *NeuroRehabilitation*, vol. 17(3), pp. 215–224, 2002.

[2] P. Hlustik and M. Mayer, “Paretic hand in stroke: from motor cortical plasticity research to rehabilitation,,” *Cogn Behav Neurol*, vol. 19(1), pp. 34–40, 2006.

[3] C. Michel and M. Murray, “Towards the utilization of EEG as a brain imaging tool,” *NeuroImage*, vol. 61, no. 2, pp. 371 – 385, 2012.

[4] S. G. Mason, A. Bashashati, M. Fatourechi, K. F. Navarro, and G. E. Birch, “A comprehensive survey of brain interface technology designs,” *Ann Biomed Eng*, vol. 35, pp. 137–169, 2007.

[5] F. Lotte, M. Congedo, A. Lcuyer, F. Lamarche, and B. Arnaldi, “A review of classification algorithms for EEG-based brain-computer interfaces,” *J Neural Eng*, vol. 4, pp. R1–R13, 2007.

[6] B. Lange, S. Flynn, R. Proffitt, C. Chang, and A. Rizzo, “Development of an interactive game-based rehabilitation tool for dynamic balance training.” *Top Stroke Rehabil*, vol. 17(5), pp. 345–352, 2010.

[7] H. Zhou and H. Hu, “Human motion tracking for rehabilitation: A survey,” *Biomed Signal Process Control*, vol. 3, no. 1, pp. 1 – 18, 2008.

[8] S. Makeig, “Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones,” *Electroencephalogr Clin Neurophysiol*, vol. 86, pp. 283–293, 1993.

[9] C. Breitwieser, I. Daly, C. Neuper, and G. Müller-Putz, “Proposing a standardized protocol for raw biosignal transmission,” *IEEE Trans Biomed Eng*, vol. 59, no. 3, pp. 852 –859, 2012.

[10] A. Delorme and S. Makeig, “EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis,” *J Neurosci Methods*, vol. 134, pp. 9–21, 2004.

[11] C. Neuper and G. Pfurtscheller, “Evidence for distinct beta resonance frequencies in human eeg related to specific sensorimotor cortical areas,” *Clin Neurophysiol*, vol. 112, pp. 2084–2097, 2001.