Neural mechanisms in processing of emotion in real and virtual faces using functional-near infrared spectroscopy (fNIRS)

by

Dylan Rapanan

A thesis submitted to the
School of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements for the degree of

Masters of Science in **Computer Science**

Faculty of Science
Ontario Tech University
Oshawa, Ontario, Canada

Abstract

This study investigates the neural mechanisms underlying face and emotion perception in both real and virtual stimuli using functional near-infrared spectroscopy (fNIRS), a question of growing importance given the increasing prevalence of avatars/virtual characters in our lives. We employed multiple analyses, including General Linear Model (GLM) analysis, and functional connectivity. The GLM analysis was used to identify brain regions activated by different face types and emotions, revealing differences in activation mainly in occipital regions. Functional connectivity analysis provided insights into the correlation between brain regions in the time-frequency domain, finding significantly correlated brain regions. Our findings reveal the similarities and differences in how the brain distinguishes between real and virtual emotional expressions. This research has implications for the development of virtual environments and the design of more effective virtual characters.

Acknowledgements

** Put your Acknowledgements here. ***

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Chapter 2

Methodology

2.1   Data Acquisition

fNIRS data was collected using two NIRSport2 systems (NIRx Medical Technologies, Berlin, Germany). Each NIRSport2 system was equipped with 16 source and 16 detector optodes, and daisy-chained together for a high density 32x32 optode configuration. Each neighboring pair of source and detector optode is referred to as a channel, resulting in a total of 103 HbO + 103 HbR channels (plus 16 short distance channels). The average distance between source and detector optodes was 30 mm, and 7mm for short distance channels, which were placed on a flexible fNIRS head cap (NIRScap) 58 cm in circumference. The optodes were arranged in a high density 32x32 montage with one bundle of short distance channels, as shown in Figure 2.1. This montage was designed to cover a maximally large area of the brain, given increasing evidence that emotion processing is not localized to specific discrete areas of the brain, rather distributed across the brain (Lindquist et al., 2012). The fNIRS cap and optodes were positioned following the 10-20 international coordinate system. Light was emitted at 760 nm and 850 nm wavelengths, and the sampling rate was approximately 6.105 Hz.
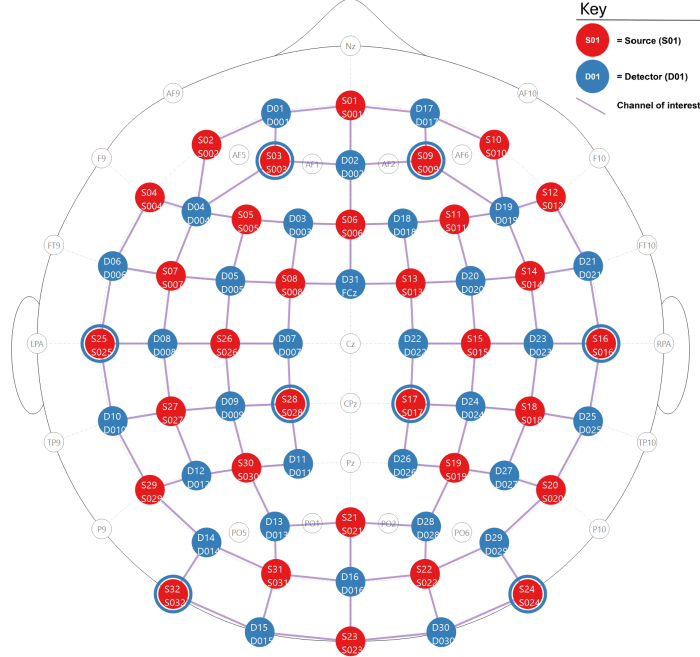
Figure 2.1: 2D depiction of the high density 32x32 optode montage, red circles represent sources, blue circles represent detectors, purple lines represent channels, and blue rings around sources represent the locations of the 8 short distance detectors.

2.2 Participants

Participants were recruited from Ontario Tech University's SONA system. 91 participants completed the study, however 3 were removed due to recording issues, leaving 88 participants. Signal quality was assessed using the following criteria, similar to (Bulgarelli et al., 2025) and (Hernandez and Pollonini, 2020): Sliding windows of 5 seconds were taken from each channel, and Peak Spectral Power (PSP) and the Scalp Coupling Index (SCI) (Pollonini et al., 2016) were calculated for each window. If PSP > 0.1 and SCI > 0.5 for more than 70% of the windows, the channel was considered to have good signal quality. If > 70% of channels were considered to have good signal quality, the participant was included in the analysis. The minimum SCI threshold was set to 0.5, used in (Holmes et al., 2024), however other studies have used SCI thresholds as low as 0.25(Zhou et al.,

2024). After applying these criteria, 53 participants were included in the analysis. All participants met the PSP threshold, but only 53/88 met the SCI threshold, as shown in Figure 2.2. The 53 participants ranged in age from 17 to 51 (M = 21.60, SD = 6.61), and no demographic information was collected.
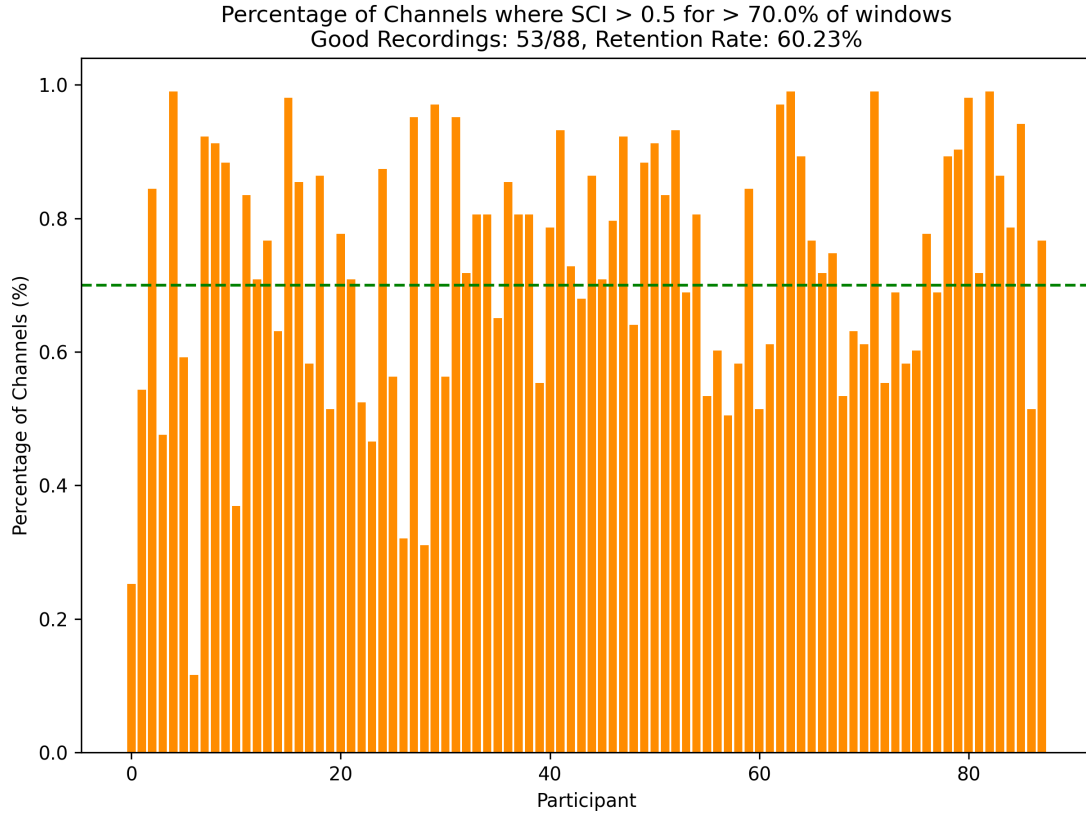
Figure 2.2: Percentage of Channels where SCI > 0.5 for > 70% of the windows. The green dashed line represents the threshold of 70% of windows that each participant must meet to be included in the analysis.

2.3 Experimental Paradigm

2.3.1 Participant Instructions

Upon entering the lab, the participant was greeted by the experimenter and asked to read and sign a consent form. Their head size was then measured using a tape measure, and the fNIRS cap was fitted to their head as the fNIRS cap was explained to them.

The participant waited patiently while the experimenter(s) checked the signal quality in Aurora fNIRS, the acquisition software for the NIRSport2 system. The experimenter(s) then attempted to move the participants' hair out of the way of the optodes, to improve signal quality. The experimenter(s) then explained the task to the participant, and notified them of the camera/microphone in the back of the room. No emphasis was given on what to focus on during the task, all the participant was told is that they will be seeing many different types of faces, and that they need to complete the memory task, which is as follows:

Memory Task: The participant will see blocks of 8 faces, and after each block, there will be a 9th face, and they will need to indicate whether the 9th face was in the previous block of 8 faces. The participant was instructed to press 'y' on the keyboard if the face was in the previous block, or 'n' if the face was not in the previous block. This task is visualized in Figure 2.3.

The experimenter(s) then turned the lights off in the room to avoid any interference with the fNIRS cap. Then the experimenter(s) started the experiment, and left the room to minimize noise and distractions. Experimenter(s) then monitored the experiment from outside the room. After the experiment was completed (about 35 minutes later), the experimenter(s) entered the room, removed the fNIRS cap, and the participant was given a debriefing letter to read and take home. The main purpose of the memory task was to keep the participant engaged and focused on the faces, the memory task was not the main focus of the study, and this was divulged to the participant if they inquired about it after the experiment was completed.

2.3.2 Facial Expression Dataset Selection

Faces were picked from two datasets: the RADIATE dataset (Conley et al., 2018) and the UIBVFED dataset (Oliver and Amengual Alcover, 2020). The RADIATE dataset is a set of 16 emotional expressions by over 100 racially and ethnically diverse real people. The

UIBVFED dataset is a set of 32 emotional expressions by 20 virtual characters that are also ethnically diverse. The faces used in the stimuli presentation were picked carefully with some criteria in mind:

- 20 models were selected, 5 males, 5 females from both datasets (RADIATE and UIBVFED).

- The corresponding models from each dataset were matched as closely as possible, in terms of each model's face shape, skin tone, and hair color.

- 7 emotional expressions (anger, disgust, fear, happiness, sadness, surprise, neutral) were selected for each model, that mostly closely align with Ekman's 6 basic emotions + neutral (Ekman, 1992).

The UIBVFED images were cropped to the same size as the RADIATE images, so now a presentation of the stimuli can be created.
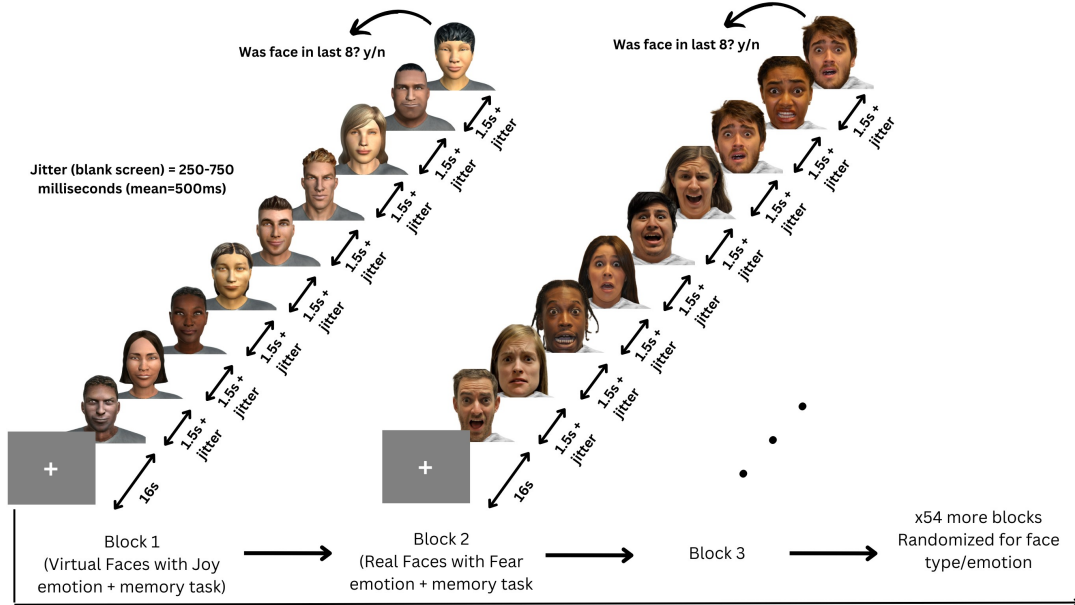
2.3.3 Stimuli Presentation

Figure 2.3: Participants viewed 56 blocks of 8 faces (4 male, 4 female) from two sets: real (RADIATE) and virtual (UIBVFED). Each face displayed one of 7 emotions (anger, disgust, fear, happiness, sadness, surprise, neutral).

The stimuli presentation was prepared using PsychoPy3 Experiment Builder (v2024.1.5) (Peirce et al., 2019). The stimuli were presented on a Dell U2415 24 inch 1920x1200 60Hz monitor, placed at eye level, and the participants were seated in a comfortable chair facing the monitor. The beginning of the stimuli presentation had instructions on the screen, which explained the task to the participant, and the participant entered the space bar when they were ready to start the experiment.

The format of the stimuli presentation was as follows: There are 56 blocks in total, each block consists of 8 faces followed by a 9th face. The faces within each block are either real (from the RADIATE dataset) or virtual (from the UIBVFED dataset), and the faces are presented in a random order. All faces within a block had one of seven emotional expressions (joy, fear, anger, disgust, sadness, neutral, surprise), and all faces within a block express the same emotion. Between every block and starting the experiment, there is a fixation cross presented for 16 seconds. The 8 faces (4 male, 4 female, randomly selected) are presented one at a time, for 1.5 seconds each, with a 250-750 ms (mean 500 ms) interstimulus interval (ISI) between each face. The 9th face is presented after the 8th face, and the participant must indicate whether the 9th face was in the previous block of 8 faces by pressing 'y' for yes or 'n' for no. This task is described in detail in the **Memory Task** section above. The 9th face has a 50% chance of either being in the previous block or not, and if the participant does not respond within 7 seconds, the presentation will continue to the next block. Every 7 blocks, the participants are given a break, and prompted to enter the space bar when they are ready to continue the experiment. The stimuli presentation lasted approximately 35 minutes in total. This paradigm is shown in Figure 2.3.
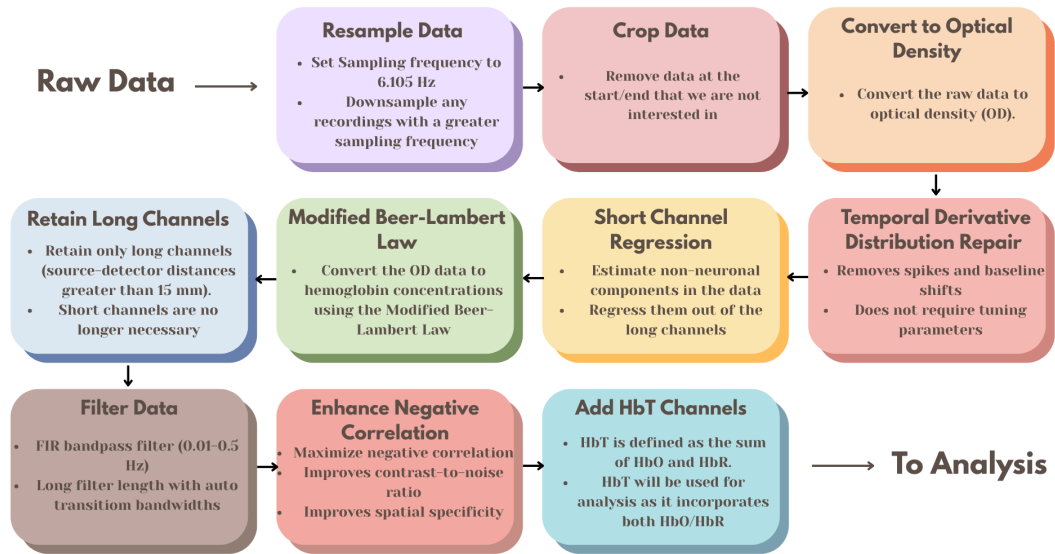
2.4   Preprocessing Steps

Figure 2.4: Preprocessing steps for fNIRS data, from the raw data to the fully processed data.

All fNIRS data was preprocessed and analyzed using MNE-Python (Gramfort et al., 2013) and MNE-NIRS (Luke et al., 2021), using Python 3.11.9. The preprocessing steps for the fNIRS data (shown in Figure 2.4) were as follows:

1. **Downsample Data:** Downsample the data if the sampling frequency is greater than 6.105 Hz, the initial two datasets were sampled higher than 6.105 Hz, and we want to keep the sampling frequency consistent across all datasets.

2. **Crop Data:** Crop the data to the first and last annotation. This gets rid of the extra data at the beginning and end of the recording that we are not interested in.

3. **Convert to Optical Density (OD):** Convert the raw data to optical density.

4. **Temporal Derivative Distribution Repair (TDDR):** Apply temporal derivative distribution repair to the OD data (Fishburn et al., 2019). TDDR is effective at removing spikes and baseline shifts from the data.

5. **Short Channel Regression:** Apply short channel regression to the OD data (Scholkmann et al., 2014). Short channels are used to estimate the superficial hemodynamics (non-evoked/extracerebral/systemic components) in the data, and then regress it out of the long channels (Tachtsidis and Scholkmann, 2016).

6. **Modified Beer-Lambert Law (MBLL):** Convert the OD data to hemoglobin concentrations using the modified Beer-Lambert law. The MBLL relates the change in light attenuation to the change in hemoglobin concentration of chromophores in the tissue (Kocsis et al., 2006).

7. **Retain Long Channels:** Retain only long channels (source-detector distance > 15 mm). Since the short channels have already been regressed out, we do not need to keep them in the data.

8. **Filter Data:** This FIR bandpass filter extracts signal components in the 0.01-0.5 Hz range, it uses a long filter length (2015 samples) with automatically determined transition bandwidths by MNE-Python (Pinti et al., 2019).

9. **Enhance Negative Correlation:** Maximizes negative correlation between HbO and HbR (Cui et al., 2010). This method removes spikes, improves contrast-to-noise ratio, and improves spatial specificity of the data.

10. **Add HbT Channels:** Add HbT (hemoglobin total) channels to the data. HbT is defined as the sum of HbO and HbR. Often, fNIRS studies will only use either one of HbO or HbR channels (more frequently HbO), leaving out one channel with no justification (Kinder et al., 2022). Therefore, we opt to use HbT channels, as technically HbT represents both HbO and HbR channels, and using both hemoglobin species improves the inferences as to where activation occurs Hocke et al. (2018).

2.5 General Linear Model (GLM) Analysis

2.6 Connectivity Analysis

2.7 Building Classifiers

2.8 Memory Task Analysis

Chapter 3

Results

3.1   General Linear Model (GLM) Results

3.2   Functional Connectivity Results

3.3   Classification Results

3.4   Memory Task Results

Chapter 4

Discussion

Bibliography

Bulgarelli, C., Blasi, A., McCann, S., Milosavljevic, B., Ghillia, G., Mbye, E., Touray, E., Fadera, T., Acolatse, L., Moore, S. E., Lloyd-Fox, S., Elwell, C. E., Eggebrecht, A. T., and Team, t. B. S. (2025). Growth in early infancy drives optimal brain functional connectivity which predicts cognitive flexibility in later childhood. Pages: 2024.01.02.573930 Section: New Results.

Conley, M. I., Dellarco, D. V., Rubien-Thomas, E., Cohen, A. O., Cervera, A., Tottenham, N., and Casey, B. (2018). The racially diverse affective expression (RADIATE) face stimulus set. *Psychiatry Research*, 270:1059–1067.

Cui, X., Bray, S., and Reiss, A. L. (2010). Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. *NeuroImage*, 49(4):3039–3046.

Ekman, P. (1992). Are there basic emotions? *Psychological Review*, 99(3):550–553. Place: US Publisher: American Psychological Association.

Fishburn, F. A., Ludlum, R. S., Vaidya, C. J., and Medvedev, A. V. (2019). Temporal Derivative Distribution Repair (TDDR): A motion correction method for fNIRS. *NeuroImage*, 184:171–179.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., and Hämäläinen, M. (2013). MEG

and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, 7. Publisher: Frontiers.

Hernandez, S. M. and Pollonini, L. (2020). NIRSplot: A Tool for Quality Assessment of fNIRS Scans. In *Biophotonics Congress: Biomedical Optics 2020 (Translational, Microscopy, OCT, OTS, BRAIN)*, page BM2C.5, Washington, DC. Optica Publishing Group.

Hocke, L. M., Oni, I. K., Duszynski, C. C., Corrigan, A. V., Frederick, B. D., and Dunn, J. F. (2018). Automated Processing of fNIRS Data—A Visual Guide to the Pitfalls and Consequences. *Algorithms*, 11(5):67. Number: 5 Publisher: Multidisciplinary Digital Publishing Institute.

Holmes, M., Aalto, D., and Cummine, J. (2024). Opening the dialogue: A preliminary exploration of hair color, hair cleanliness, light, and motion effects on fNIRS signal quality. *PLOS ONE*, 19(5):e0304356. Publisher: Public Library of Science.

Kinder, K. T., Heim, H. L. R., Parker, J., Lowery, K., McCraw, A., Eddings, R. N., Defenderfer, J., Sullivan, J., and Buss, A. T. (2022). Systematic review of fNIRS studies reveals inconsistent chromophore data reporting practices. *Neurophotonics*, 9(4):040601.

Kocsis, L., Herman, P., and Eke, A. (2006). The modified Beer–Lambert law revisited. *Physics in Medicine & Biology*, 51(5):N91.

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., and Barrett, L. F. (2012). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences*, 35(3):121–143.

Luke, R., Larson, E. D., Shader, M. J., Innes-Brown, H., Yper, L. V., Lee, A. K. C., Sowman, P. F., and McAlpine, D. (2021). Analysis methods for measuring passive

auditory fNIRS responses generated by a block-design paradigm. *Neurophotonics*, 8(2):025008. Publisher: SPIE.

Oliver, M. M. and Amengual Alcover, E. (2020). UIBVFED: Virtual facial expression dataset. *PLOS ONE*, 15(4):e0231266.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., and Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1):195–203.

Pinti, P., Scholkmann, F., Hamilton, A., Burgess, P., and Tachtsidis, I. (2019). Current Status and Issues Regarding Pre-processing of fNIRS Neuroimaging Data: An Investigation of Diverse Signal Filtering Methods Within a General Linear Model Framework. *Frontiers in Human Neuroscience*, 12.

Pollonini, L., Bortfeld, H., and Oghalai, J. S. (2016). PHOEBE: a method for real time mapping of optodes-scalp coupling in functional near-infrared spectroscopy. *Biomedical Optics Express*, 7(12):5104–5119. Publisher: Optica Publishing Group.

Scholkmann, F., Metz, A. J., and Wolf, M. (2014). Measuring tissue hemodynamics and oxygenation by continuous-wave functional near-infrared spectroscopy—how robust are the different calculation methods against movement artifacts? *Physiological Measurement*, 35(4):717. Publisher: IOP Publishing.

Tachtsidis, I. and Scholkmann, F. (2016). False positives and false negatives in functional near-infrared spectroscopy: issues, challenges, and the way forward. *Neurophotonics*, 3(3):031405. Publisher: SPIE.

Zhou, X., Hong, X., and Wong, P. C. M. (2024). Autistic Traits Modulate Social Synchronizations Between School-Aged Children: Insights From Three fNIRS Hyperscanning Experiments. *Psychological Science*, 35(8):840–857. Publisher: SAGE Publications Inc.