Causal Influences in the Human Brain During Face Discrimination: A Short-Window Directed Transfer Function Approach

Marios G. Philiastides and Paul Sajda*, *Senior Member, IEEE*

***Abstract*—In this letter, we considered the application of parametric spectral analysis, namely a short-window directed transfer function (DTF) approach, to multichannel electroencephalography (EEG) data during a face discrimination task. We identified causal influences between occipitoparietal and centrofrontal electrode sites, the timing of which corresponded to previously reported EEG face-selective components. More importantly we present evidence that there are both feedforward and feedback influences, a finding that is in direct contrast to current computational models of perceptual discrimination and decision making which tend to favor a purely feedforward processing scheme.**

***Index Terms*—Causal influences, directed transfer function (DTF), electroencephalography (EEG), face discrimination, feedback, feedforward.**

I. Introduction

THE neural correlates of face processing in the human brain have been extensively explored. Several studies, using electroencephalography (EEG) and magnetoencephalography (MEG), have already identified two distinct face-selective components during similar face discrimination tasks [1]–[5]. Specifically, these components were found at 100 and 170 ms (the well known N170) after the onset of visual stimulation.

Though the implication of these components in face processing is widely accepted, little is known about the causal relationships of the areas involved in their generation. Popular models of perceptual discrimination and the subsequent decision making (i.e. accumulator and diffusion models) [6]–[9] propose a purely feedforward processing stream between sensory and decision areas.

The directed transfer function (DTF) method has been shown to be a robust tool in estimating direction of information flow in biological systems as a function of frequency [10]–[14]. In this paper we use a short-time window DTF approach [15], [16] to analyze EEG data collected during a face discrimination task in order to identify the flow of neural information during face discrimination.

Manuscript received June 18, 2006; revised July 29, 2006. This work was supported in part by the Office of Naval Research (ONR) under Grant N00014-01-1-0625, in part by the National Geospatial-Intelligence Agency (NGA) under Grant HM1582-05-C-0008, and in part by the National Institutes of Health (NIH) under Grant EB004730. *Asterisk indicates corresponding author.*

M. G. Philiastides is with the Laboratory for Intelligent Imaging and Neural Computing, Department of Biomedical Engineering, Columbia University, New York, USA (e-mail: mgp2101@columbia.edu).

*P. Sajda is with the Laboratory for Intelligent Imaging and Neural Computing, Department of Biomedical Engineering, Columbia University, New York, USA (e-mail: ps629@columbia.edu).

Digital Object Identifier 10.1109/TBME.2006.885122

II. Methods

A. Behavioral Paradigm

Six subjects performed a simple face-versus-car discrimination task. Images of faces and cars at two different difficulty levels were presented in random order. We adjusted the difficulty level by changing the percent phase coherence of our images [17] [see Fig. 1(a)]. All subjects performed nearly perfectly at the highest phase coherence and just above psychophysical threshold for the lower coherence level (i.e. $\approx$85% correct). Images were presented for 30 ms and followed by an interstimulus-interval which was randomized in the range of 1500–2000 ms. Subjects reported their decision by pressing a mouse button using their right hand.

B. Data Acquisition

EEG data was acquired from 60 Ag/AgCl scalp electrodes. All channels were referenced to the left mastoid with input impedance $<$ 15 kΩ and chin ground. Data was sampled at 1000 Hz with an analog pass band of 0.01–300 Hz using 12 dB/octave high-pass and eighth-order Elliptic low pass filters. Subsequently, a software-based 0.5 Hz high-pass filter was used to remove DC drifts and 60 and 120 Hz (harmonic) notch filters were applied to minimize line noise artifacts. Artifacts such as eye blinks, and horizontal and vertical eye-movements were removed using a principle component analysis approach [18].

C. Data Analysis

1) Directed Transfer Function: In the context of a multivariate autoregressive model (MVAR) we defined the measurement from p channels at time t as $\mathbf{X}(t) = [X_1(t), X_2(t), \ldots, X_p(t)]^T$ and assumed this is adequately described by the following mth-order equation:

$$\mathbf{X}(t) + \mathbf{A}(1)\mathbf{X}(t-1) + \ldots + \mathbf{A}(m)\mathbf{X}(t-m) = \mathbf{E}(t) \quad (1)$$

$\mathbf{E}(t)$ denotes a zero-mean uncorrelated noise vector $[E_1(t), E_2(t), \ldots, E_p(t)]^T$. Matrices $\mathbf{A}(i)$ are of size $p \times p$. To estimate $\mathbf{A}(i)$ we multiplied (1) by $\mathbf{X}^T(t-k)$, where $k = 1, 2, \ldots m$, to obtain the Yule-Walker equations

$$\mathbf{R}(-k) + \mathbf{A}(1)\mathbf{R}(-k+1) + \ldots + \mathbf{A}(m)\mathbf{R}(-k+m) = \mathbf{0} \quad (2)$$

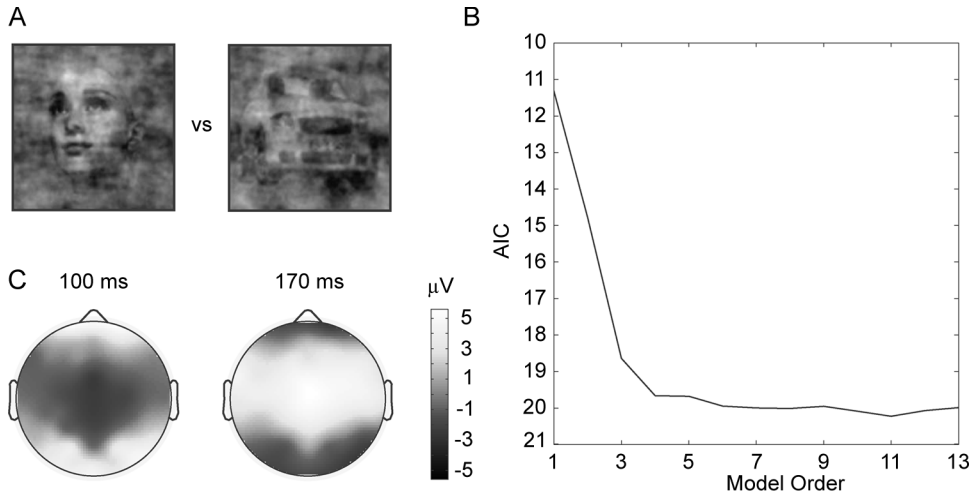

Fig. 1. (a) Sample images used in our face-versus-car discrimination task. (b) AIC as a function of model order computed in a representative 50-ms window. The shape of the curve is similar for all other windows. (c) Average ERP scalp maps for the two face-selective components at 100 and 170 ms after stimulus onset.

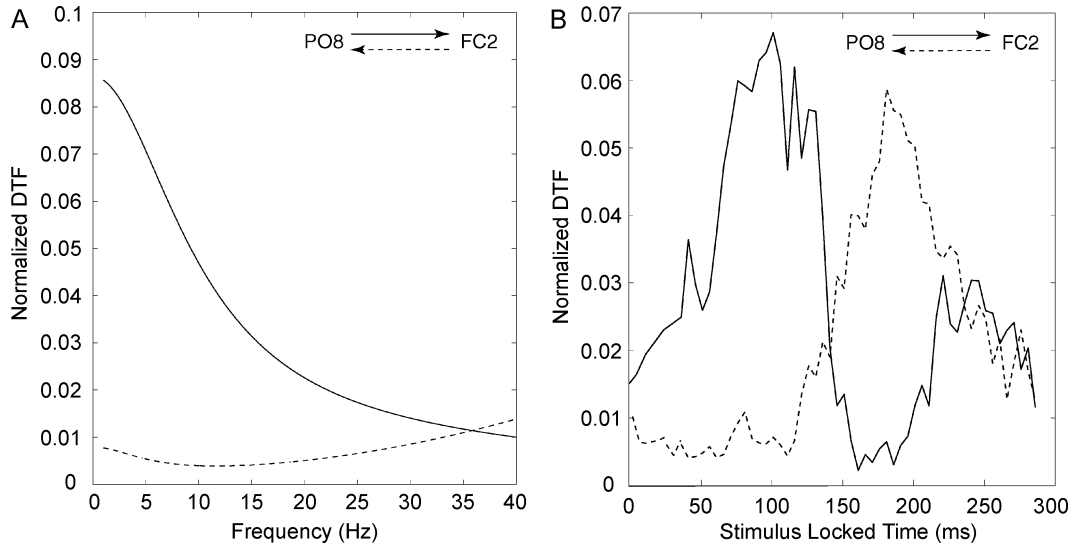

Fig. 2. (a) Normalized DTF as a function of frequency for a representative electrode pair (PO8-FC2) 100 ms after stimulus onset. (b) Normalized DTF, averaged in the frequency range 1–12 Hz, as a function of time. Each data point was computed by sliding a 50-ms window across time in 5-ms increments. For both figures, the solid lines indicate the feedforward influence, while the dotted lines indicate feedback.

$\mathbf{R}(n) = \langle \mathbf{X}(t)\mathbf{X}^T(t+n) \rangle$ are the covariance matrices of all $\mathbf{X}(t)$ of lag n. We solved these equations using a least squares approach based on QR-factorization [19].

Transforming (1) to the frequency domain yields $\mathbf{A}(f)\mathbf{X}(f) = \mathbf{E}(f)$, where $\mathbf{A}(f) = -\sum_{j=0}^{p} \mathbf{A}(j) e^{-i2\pi fj}$ with $\mathbf{A}(0) = -\mathbf{I}$ ($\mathbf{I}$: identity matrix). Equation (1) can then be rewritten as

$$\mathbf{X}(f) = \mathbf{A}^{-1}(f)\mathbf{E}(f) = \mathbf{H}(f)\mathbf{E}(f). \tag{3}$$

The transfer function of the system is denoted by $\mathbf{H}(f)$ and the DTF which represents the causal influence from channel j to channel i is defined as

$$\theta_{ij}^2(f) = |H_{ij}(f)|^2 . \tag{4}$$

The normalized DTF [20] then becomes

$$\gamma_{ij}^2(f) = \frac{|H_{ij}(f)|^2}{\sum_{r=1}^{p} |H_{ir}(f)|^2}. \tag{5}$$

To estimate the correct model order we used the Akaike Information Criterion (AIC) [21] defined as, $\mathrm{AIC}(m) = 2log[det(\mathbf{\Sigma})] + 2p^2 m/N_{\mathrm{total}}$, where $\mathbf{\Sigma}$ denotes the covariance matrix of the noise vector $\mathbf{E}(t)$ and N_{total} the total number of data points from all trials. We determined that order $m = 7$ is sufficient since there is little change in the AIC beyond that value [Fig. 1(b)].

2) Short-Time Window Directed Transfer Function: As mentioned above, to estimate $\mathbf{A}(i)$ the covariance matrices ($\mathbf{R}$'s) of the measurement $\mathbf{X}(t)$ are required. However, for a short-time window, typically 50 ms in length [10], [16], there are not enough data points to provide good estimates. To overcome this

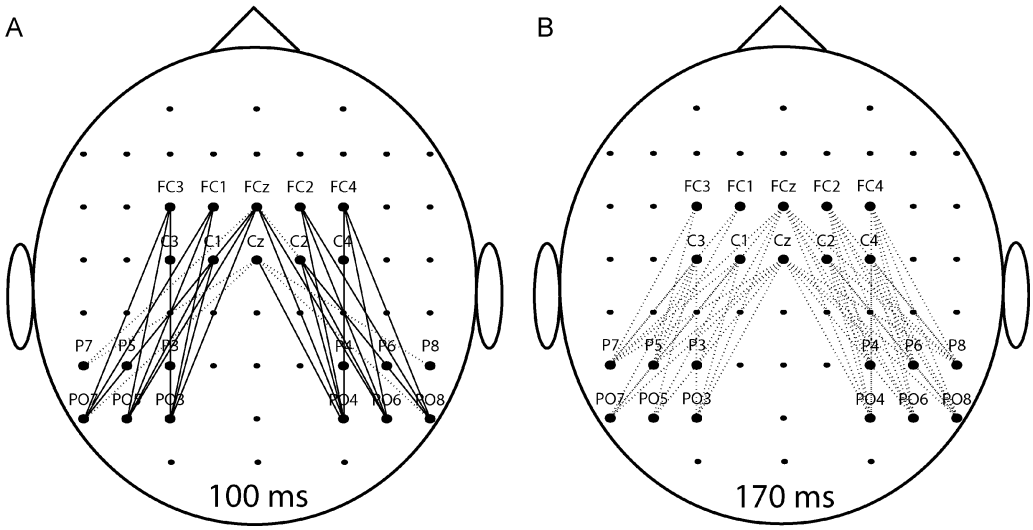

Fig. 3. Group results summarizing the directional influences between all electrode sites included in the MVAR analysis for (a) the 100-ms face component and (b) the 170-ms face component. Only significant connections (as assessed by the bootstrap technique) are shown. Note that for the 100-ms component the majority of the connections are feedforward. For the 170-ms component all significant connections are feedback.

problem several realizations of the same process are required (i.e. several trials) in order to average across all of them to obtain final estimates for the $\mathbf{R}$'s [16]. Therefore we used several hundred trials (across both difficulty levels and across all subjects) to obtain reliable estimates of the $\mathbf{R}$'s and subsequently of $\mathbf{A}(i)$.

III. Results

We computed event-related potentials (ERPs) by averaging all face trials across all subjects. In order to visualize the spatial extent of the face-selective ERP components we constructed average ERP scalp maps [Fig. 1(c)]. As expected we were able to identify the two face-selective components at 100 and 170 ms poststimulus. Note that the spatial distribution of activity is very similar between the two time points, however, the polarity of the ERPs is reversed. Specifically, strong activations are seen at several centrofrontal and occipitoparietal electrode sites. We only included electrodes from these sites into the MVAR process. Specifically, we used a total of 22 electrodes; 10 from centrofrontal locations and 12 from occipitoparietal sites (see Fig. 3).

A. DTF Analysis of the EEG

We used the short-time window approach outlined above to compute the normalized DTF [i.e., $\gamma_{ij}^2(f)$] across time by sliding a 50-ms-long window (in 5–ms increments) starting at the onset of visual stimulation. For this analysis we used a total of 960 trials across all subjects and both difficulty levels. Prior to carrying out the MVAR modeling, however, we performed two preprocessing steps on the EEG data. First, we aligned all trials to the onset of the stimulus and subtracted the temporal mean of the entire trial and divided by the standard deviation (on an electrode-by-electrode basis). This step ensured that each trial was given an equal weight in the model estimation. Second, we subtracted the ensemble mean from each trial and we divided by the ensemble standard deviation (on a point-by-point basis). This step ensured that the non-stationarities embodied in the ensemble mean and standard deviation respectively were removed.

Fig. 2(a) shows an example of the causal influence from channel PO8 to channel FC2 and *vice versa* at 100 ms poststimulus. In the frequency range 1–12 Hz there is a significant influence from the occipitoparietal site (PO8) onto the centrofrontal location (FC2) (solid line) whereas the opposite (dotted line) is not true. Note that, in the context of the EEG, this frequency range, can be considered physiologically realistic. Specifically, it includes the delta (1–4 Hz), theta (4–8 Hz), and alpha (8–12 Hz) bands which have been associated with different cognitive events such as attention, memory, and target detection [22]–[24].

In order to visualize the temporal evolution of the causal relationships between all electrode locations in our MVAR model we computed the average normalized DTF in the frequency range 1–12 Hz for each of the several 50-ms sliding windows. Fig. 2(b) shows one such example for the same electrode pair. Interestingly, we were able to identify a causal influence from occipitoparietal sites to centrofrontal locations which began about 100 ms poststimulus. The feedback influence did not arise until 70 ms later, around 170 ms. In principle, it is expected that higher processing areas will begin influencing lower processing areas after first receiving information from them. Our results not only appear to be consistent with this hypothesis but the timing of these effects are perfectly consistent with the two face-selective components previously reported for similar behavioral tasks [1]–[5].

B. Statistical Analysis and Group Results

In order to establish statistical significance for the DTF peaks seen in all of the electrode-pair influences we used a bootstrap resampling technique. Specifically, we randomly resampled all of the 960 trials 100 times, each time re-computing the normalized DTF. We compared the maximum value of each peak (at a

particular time point) against the corresponding DTF bootstrap distribution to determine significant peak values ($p < 0.01$).

We used this technique to determine all of the significant electrode-pair influences at the two times of interest (i.e. 100 and 170 ms poststimulus) in order to summarize the overall causal relationships between centrofrontal and occipitoparietal sites during face processing. Fig. 3(a) and (b) represents the directional influences for the 100- and 170-ms time windows respectively. Solid lines represent the forward influence from occipitoparietal electrode sites to centrofrontal locations while the dotted lines indicate the feedback influence. It is clear that the majority of the directional influence at the earlier time (100 ms) are feedforward and for the later time (170 ms) are feedback.

In addition the relative number of these influences suggests that there are more significant relationships at the later time (and with a right hemispheric bias) than there are for the earlier one. This is not surprising since many EEG/MEG studies have shown the importance of the N170 over the 100-ms component during face processing [2], [3], [5]. There are also several reports of a right-hemisphere bias for the N170 component [25], [26].

IV. Conclusion

In this work we took advantage of a short-window DTF approach, to identify the directional influences between different EEG electrode sites during a face discrimination task. Using this technique we were able to infer the causal relations of EEG components that have already been implicated in face processing (the 100- and 170-ms poststimulus components) [1]–[5]. Specifically, we have presented evidence that the majority of causal influences 100 ms after the onset of visual presentation follow a feedforward direction from occipitoparietal to centrofrontal electrode sites. It is not until 70 ms later (N170 component) that centrofrontal locations begin to exert influences, that can be interpreted as feedback, on occipitoparietal sites. These findings are in direct contrast to popular computational models of visual discrimination and the subsequent decision making [6]–[8] that favor purely feedforward processing.

References

[1] E. Halgren, T. Raij, K. Marinkovic, V. Jousmaki, and R. Hari, "Cognitive response profile of the human fusiform face area as determined by MEG," *Cereb. Cortex*, vol. 10, pp. 69–81, 2000.

[2] J. Liu, A. Harris, and N. Kanwisher, "Stages of processing in face perception: an MEG study," *Nature Neurosci.*, vol. 5, no. 9, pp. 910–916, 2002.

[3] B. Rossion, C. Joyce, G. Cottrell, and M. Tarr, "Early laterization and orientation tuning for face, word, object processing in the visual cortex," *NeuroImage*, vol. 20, pp. 1609–1624, 2003.

[4] V. Goffaux, I. Gauthier, and B. Rossion, "Spatial scale contribution to early visual differences between face and object processing," *Cogn. Brain Res.*, vol. 16, pp. 416–424, 2003.

[5] M. Philiastides and P. Sajda, "Temporal characterization of the neural correlates of perceptual decision making in the human brain," *Cereb. Cortex*, vol. 16, no. 4, pp. 509–518, 2006.

[6] R. Ratcliff, "A theory of memory retrieval," *Psychol. Rev.*, vol. 85, pp. 59–108, 1978.

[7] M. Mazurek, J. Roitman, J. Ditterich, and M. Shadlen, "A role for neural integrators in perceptual decision making," *Cereb. Cortex*, vol. 13, pp. 1257–1269, 2003.

[8] A. Huk and M. Shadlen, "Neural activity in the macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making," *J. Neurosci.*, vol. 25, no. 45, pp. 10 420–10 436, 2005.

[9] M. Philiastides, R. Ratcliff, and P. Sajda, "Neural representation of task difficulty and decision-making during perceptual categorization: a timing diagram," *J. Neurosci.*, vol. 26, no. 35, pp. 8965–8975, 2006.

[10] M. Kaminski, M. Ding, W. Truccolo, and S. Bressler, "Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance," *Biol. Cybern.*, vol. 85, pp. 145–157, 2001.

[11] A. Korzeniewska, M. Manczak, M. Kaminski, K. Blinowska, and S. Kasicki, "Determination of information flow direction among brain structures by a modified directed transfer function method," *J. Neurosci. Meth.*, vol. 125, pp. 195–207, 2003.

[12] L. Gennaro, F. Vecchio, M. Ferrara, G. Curcio, P. Rossini, and C. Babiloni, "Changes in fronto-posterior functional coupling at sleep onset in humans," *J. Sleep Res.*, vol. 13, pp. 209–217, 2004.

[13] R. Kus, M. Kaminski, and K. Blinowska, "Determination of EEG activity propagation: pair-wise versus multichannel estimate," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 9, pp. 1501–1510, Sep. 2004.

[14] M. Kaminski and H. Liang, "Causal influence: advances in neurosignal analysis," *Crit. Rev. Biomed. Eng.*, vol. 33, no. 4, pp. 347–430, 2005.

[15] H. Liang, M. Ding, R. Nakamura, and S. Bressler, "Causal influences in primate cerebral cortex during visual pattern discrimination," *Neuroreport*, vol. 11, no. 13, pp. 2875–2880, 2000.

[16] M. Ding, S. Bressler, W. Yang, and H. Liang, "Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data preprocessing, model validation, and variability assessment," *Biol. Cybern.*, vol. 83, pp. 35–45, 2000.

[17] S. Dakin, "What causes non-monotonic tuning of fMRI response to noisy images?," *Curr. Biol.*, vol. 12, pp. 476–477, 2002.

[18] L. Parra, C. Spence, A. Gerson, and P. Sajda, "Recipes for the linear analysis of EEG," *NeuroImage*, vol. 28, no. 2, pp. 326–341, 2005.

[19] G. Strang, *Introduction to Linear Algebra*. Cambridge, MA: Wellesley Cambridge Press, 2003.

[20] M. Kaminski and K. Blinowska, "A new method of the description of the information flow in the brain structures," *Bio.l Cybern.*, vol. 65, pp. 203–210, 1991.

[21] H. Akaike, "A new look at statistical model identification," *IEEE Trans. Autom. Control.*, vol. AC-19, pp. 716–723, 1974.

[22] W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis," *Brain Res. Rev.*, vol. 29, pp. 169–195, 1999.

[23] E. Basar, C. Basar-Eroglu, S. Karakas, and M. Schurmann, "Gamma, alpha, delta and theta oscillations govern cognitive processes," *Int. J. Psychophysiol.*, vol. 39, pp. 241–249, 2001.

[24] S. Makeig, A. Delorme, M. Westerfield, T. Jung, J. Townsend, E. Courchesne, and T. Sejnowski, "Electroencephalographic brain dynamics following manually responded visual targets," *PLoS Biol.*, vol. 2, no. 6, p. e176, 2004.

[25] N. Kanwisher, "Domain specificity in face perception," *Nature Neurosci.*, vol. 3, pp. 759–763, 2000.

[26] U. Hasson, I. Levy, M. Behrmann, T. Hendler, and R. Malach, "Eccentricity bias as an organizing principle for human high-order object areas," *Neuron*, vol. 34, pp. 479–490, 2002.