Available online at www.sciencedirect.com
SCIENCE DIRECT®
PERGAMON
Vision Research 43 (2003) 2583–2589
Vision Research
www.elsevier.com/locate/visres

Electrophysiological correlates of purely temporal figure–ground segregation

Farid I. Kandil *, Manfred Fahle

Human Neurobiology, University of Bremen, Argonnenstraße 3, D-28211 Bremen, Germany

Received 13 May 2002

Abstract

Inhomogenous displays, in contrast to homogenous ones, evoke a specific potential in the VEP (tsVEP) which appears across different classical visual stimulus dimensions defining figure–ground segregation, such as luminance, orientation, (first-order) motion, and stereoscopic depth. This negative potential has a peak latency of about 200–300 ms and a peak amplitude of about −3 to −10 μV [Doc. Ophthalmol. 95 (1998) 335]. Previously, we demonstrated that human subjects reliably segregate figure from ground, even in the absence of the classical cues, leaving time of change as the only cue for segregation. The results of the present study demonstrate that also purely temporally defined checkerboards evoke a tsVEP resembling the motion-defined tsVEP regarding polarity (negative), latency (two peaks at 180 and 270 ms, respectively), amplitude of the first negativity (−5.6 μV), and overall form of its components.

Keywords: Figure–ground segregation; Temporal texture; VEP

1. Introduction

During the last decade, a specific component in the pattern VEP has been associated with pre-attentive texture segregation. This texture-segregation component, which has been named "tsVEP" (Bach & Meigen, 1997), was found in cortical responses to checkerboards defined by spatial gradients along various dimensions: orientation of line elements (Bach & Meigen, 1990; Bach, Schmitt, Quenzer, Meigen, & Fahle, 2000), symbols (Meigen & Bach, 1993), motion (Lamme, Van Dijk, & Spekreijse, 1993), as well as luminance, stereo-depth, and color (Bach & Meigen, 1997; Fahle, Quenzer, Braun, & Spang, 2003).

On the basis of sum potential recordings in humans and monkeys as well as single cell recordings in monkeys, Lamme and coworkers (Lamme, 1995; Lamme, Van Dijk, & Spekreijse, 1992) proposed that the primary visual cortex may be able to detect texture segregation and hence to produce the tsVEP. This suggestion has recently been supported by an fMRI study by Skiera, Petersen, Skalej, and Fahle (2000). Using checkerboards defined by color, luminance and motion, they found segregation-specific activation even in V1, whereas Kastner, De Weerd, and Ungerleider (2000) using orientation-defined checkerboards found enhanced activity in areas V4 and TEO, but not in V1 or V2.

Figure–ground segregation can rely on temporal cues alone, not just on classical cues such as luminance, color, motion and orientation (Kandil & Fahle, 2001). Prior studies had shown that humans are able to segregate figure from ground when these areas are flickered asynchronously (Fahle, 1993; Usher & Donnelly, 1998; however, see Kiper, Gegenfurtner, & Movshon, 1996) or when motion direction of circular sine wave gratings in figure and ground reverses at different points in time (Lee & Blake, 1999; however, cf. Adelson & Farid, 1999; Farid & Adelson, 2001). Our stimuli differ from those previously used in that artifacts arising from differences in luminance and motion are largely eliminated, and temporal frequency as well as phase delay are controlled. Much like the stimuli used in the present study, the displays used in our previous experiment (Kandil & Fahle, 2001) consisted of an array of randomly oriented 'colons' that changed their orientation at a fixed

* Corresponding author. Present address: Department of Psychology II, University of Münster, Fliednerstrasse 21, D-48149 Münster, Germany. Tel.: +49-251-8334178; fax: +49-251-8334173.
E-mail address: kandil@psy.uni-muenster.de (F.I. Kandil).

0042-6989/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0042-6989(03)00456-5

frequency. Changes occurred synchronously among colons within both figure and ground, but phase-delayed in the ground relative to the figure. Subjects had to locate the figure in a four-alternative forced choice task while the frequency was modulated between trials. They perceived the segregation of the image reliably up to an alternation frequency of 23 Hz, corresponding to delays of approximately 22 ms between changes in figure and ground.

In the present study, using the same electrophysiological paradigm as Bach and Meigen (1997) and the same visual stimuli as before (Kandil & Fahle, 2001), we find a new tsVEP resembling the tsVEP for motion-defined segregation obtained in earlier studies (Fahle et al., 2003).

In our first experiment as well as in previous studies, clearly perceivable checkerboards were used. It seemed worthwhile to explore whether the tsVEP requires the perception of a clearly segmented form (e.g. a checkerboard consisting of squares) or else the mere perception of inhomogeneity suffices. This was tested in a second experiment using smaller checkerboard fields.

2. Methods

2.1. Temporal texture-segregation VEP

2.1.1. Subjects

Ten students (five males and five females, aged between 23 and 29 years) served as subjects in these experiments. All had normal or corrected to normal visual acuity. Prior to the electrophysiological recordings they participated in a psychophysical test of temporal grouping (for details see Kandil & Fahle, 2001) to ensure that they clearly perceived the targets under the temporal conditions of the electrophysiological experiment.

2.1.2. Setting

All stimuli were presented binocularly on a 20″ color monitor (EIZO T-662T) with a spatial resolution of 1280×1024 pixels and a frame rate of 75 Hz, driven by an AMD Duron 800 MHz PC via an Asus V7700 graphics board.

2.1.3. Stimuli

The stimuli consisted of a regular array of 16×16 colons that was subdivided into squares of 4×4 colons (Fig. 1a). In heterogenous conditions, every second square belonged to the "figure", the others to the "ground", thus forming a checkerboard of 4×4 fields. The affiliation to layer, figure or ground, was solely defined by the temporal protocol of the colons' local motions (see below).

In the course of 80 ms, each colon altered its orientation by flipping, i.e. rotating instantaneously, by 90 deg around its imaginary midpoint (Fig. 1b) creating an apparent motion stimulus of ambiguous direction. All colons in the figure flipped synchronously, and so did those in the ground, yet either with a phase delay of 40 ms (counter phase) or without any delay (in-phase). Counter phase flipping of figure and ground resulted in texture segregation and in the percept of the checkerboard (heterogenous condition) whereas in-phase flipping resulted in the percept of a homogenous field.

Viewed from a distance of 1.15 m, the full stimulation field covered approximately 11×11 deg, colons were separated by 40′ from midpoint to midpoint, a colon measured 9′ from dot to dot and a single dot had a diameter of 3′ (Fig. 1a). Colons and a central red fixation point of the same size as the dots appeared bright (40 cd/m^2) on a dark monitor screen (0.01 cd/m^2) in an otherwise dark room.

2.1.4. Paradigm

We used the same paradigm and the same rationale as Bach and Meigen (1997) when evoking and computing the tsVEP to ensure comparability with earlier studies. Homogenous displays, presenting exclusively either one or the other specification of the feature under investigation, evoke a low-level VEP associated with the mechanisms of the visual dimension stimulated, here: apparent motion. On the other side, heterogenous displays, presenting checkerboards defined by alternating squares of the two specifications, evoke both, low-level VEP and tsVEP components. Hence, the tsVEP component is the difference between the VEPs evoked by homogenous versus heterogenous displays.

There were two homogenous displays (B and D in Fig. 1c) and two heterogenous displays (A and C in Fig. 1c) in our study. In display B, all colons flipped synchronously at 0, 80, 160, 240, 320, 400 and 480 ms whereas in display D, all colons flipped synchronously at 40, 120, 200, 280, 360, 440 and 520 ms. Colons in displays A and C flipped asynchronously, half of them at 0, 80, 160, ... ms (indicated by white arrays), the other half at 40, 120, 200, ... ms (symbolized by gray arrays).

The rationale is that both, the averaged recordings of conditions B and D on one hand, and those of A and C on the other hand, contain the "low-level" cortical responses to the stimulation at every single point of time (0, 40, 80, 120, 160, 200, ...). Yet, only the responses to the heterogenous conditions A and C contain the additional texture segregation components (lower part of Fig. 1c).

2.1.5. Electrophysiological procedure

We presented each of the four displays (A–D) for a total number of 120 times, subdivided into four blocks, each of about 3 min length. Between blocks the subjects

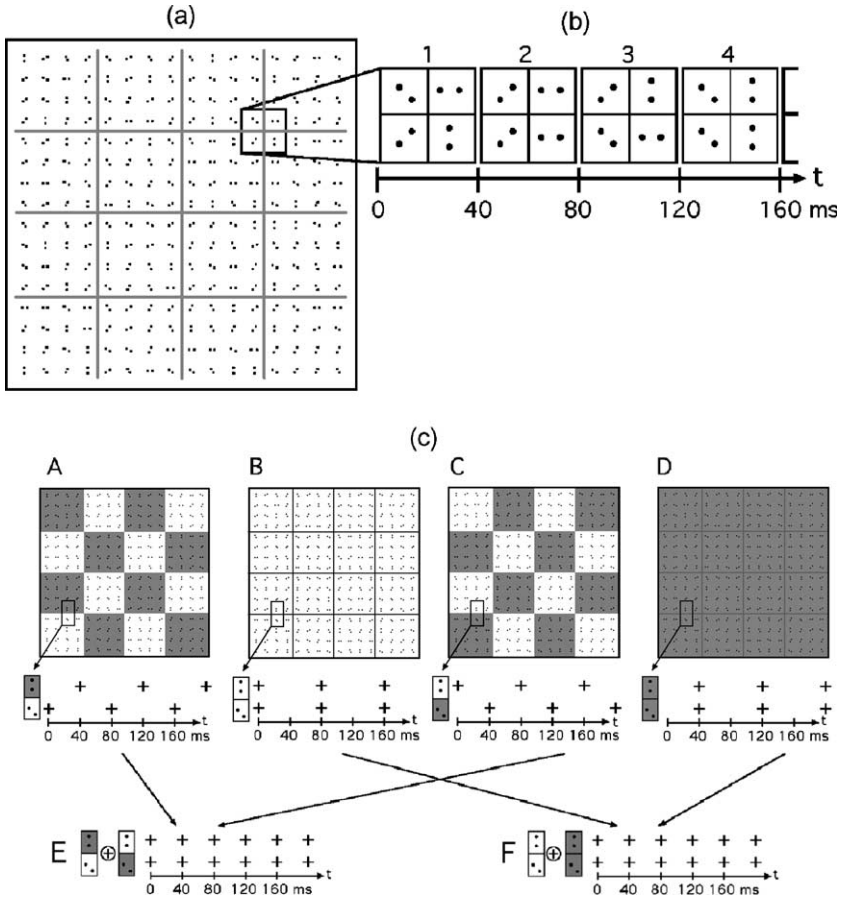

Fig. 1. (a) The display consists of 16×16 colons with starting orientation independently randomized. In contrast to the display shown here, the original display used in the experiments consisted of bright dots on a dark background. Grey lines in this figure demarcate the edges between the 4×4 fields of the checkerboard in the heterogenous conditions. These lines were not part of the original displays but appear only here for clarity. (b) A small clipping of the display exemplifies the time scheme of the local motion in the heterogenous conditions. Between frames 1 and 2, as well as between frames 3 and 4, colons in the upper-left and lower-right patches flip (that is, rotate instantaneously by 90 deg) around their imaginary midpoints. Since frames 1 through 4 are shown repeatedly, these flips occur at intervals of 40, 120, 200, ... ms after stimulus onset. Colons in the other two patches flip between frames 2 and 3, as well as between 4 and 1, corresponding to 0, 80, 160, ... ms after sweep onset. (c) For each of the two homogenous (B and D) and heterogenous (A and C) displays, flip times are marked by grayscales (white or gray), and depicted more explicitly in the lower part for an example pair of colons. All colons in each of the homogenous displays (B and D) flip at the same points in time. However, there is an offset between flip times in B and D. In contrast, colons in both heterogenous displays (A and C) all flip according to either one or the other time schedule. This difference in flip times is the only difference discriminating the squares shown in A and C. Summing flip activities over space and time for either heterogenous displays ($E = A + C$) or homogenous ones ($F = B + D$) results in analogous flip activities. Thus, there should be no difference between summed activity of low-level VEP components evoked by homogenous displays on one side and heterogenous ones on the other side. Any resulting difference would represent segregation.

were allowed a resting period of 3 min. Within every sweep dynamic displays were shown for 600 ms. Thereafter, the last frame was presented (stationary) for the next 600 ms to dishabituate the motion perception system before the next sweep started.

2.1.6. Electrophysiological recording

The VEP was recorded from a set of five active electrodes positioned above the occipital cortex at O3, O1, Oz, O2, and O4 in five subjects and only from O1, Oz, and O2 in the remaining five subjects versus a reference placed at FPz and a ground electrode at the right earlobe using gold-cup electrodes. Two further electrodes below and above the right eye were used to control for eye movements and blinks. Signals were amplified and filtered (first-order band pass, 0.3–130 Hz, Toennies "Physiologic Amplifier") and digitized with a resolution of 12 bits at a sampling rate of 400 Hz using a Maclab/8 AD converter and a Macintosh G3 computer driven by Chart 4.5 software. The recording computer continuously stored the amplified signals along with a signal from the stimulus computer indexing the actual stimulus condition for later analysis.

2.1.7. Data analysis

Data handling and analysis were performed on a PC using Igor Pro 4.0 software. Sweeps containing eye blinks, or artifacts exceeding the amplitude bandwidth

Table 1
Individual amplitude and latency values for the tsVEP component found in experiment 1 for all subjects, along with means and standard errors (SE) across ten subjects (Panel A)

Subject	Amplitude	Latency
Panel A		
ALWO	−6.1	195.0
AUWA	−2.8	172.5
CAMO	−6.7	180.0
HEST	−6.0	157.5
JABO	−6.0	182.5
KAGR	−3.8	172.5
KASP	−6.6	167.5
LAPL	−6.2	192.5
SAKN	−7.8	190.0
THEN	−4.5	175.0
Mean	−5.6	178.5
SE	0.47	3.75

Subject	Amplitude		Latency	
	Experiment 1	Experiment 2	Experiment 1	Experiment 2
Panel B				
ALWO	−5.7	−3.9	190	190
CAMO	−6.1	−7.1	180	180
KASP	−6.0	−6.7	163	160

Comparison between individual amplitude and latency values for the tsVEP components found in experiments 1 and 2 (Panel B).

of the AD converter, were excluded from the analysis. The evoked potentials were averaged individually for each of the four stimulus types of each subject. Mean responses evoked by heterogenous (A and C) and homogenous (B and D) displays were calculated as well as the differences between these means to obtain the specific potential associated with the segregation.

Peak amplitudes (as presented in Table 1 and used in the statistical tests) were measured from baseline to the peak.

2.2. Reducing the size of the checkerboard fields

To explore whether evoking a tsVEP requires the perception of a clearly segmented form or else the mere perception of inhomogeneity suffices, we performed a second experiment using the same electrophysiological paradigm but employing stimuli different from the ones of experiment 1. Here, heterogenous checkerboards consisted of 16×16 small squares, each containing only a single colon, rather than the 4×4 large squares used in experiment 1 (cf. Fig. 3a). As a consequence, subjects were still aware of the heterogenous nature of the displays, but were unable to detect any specific form (such as the squares they perceived in experiment 1). Only three subjects participated in this experiment. All other specifications corresponded to the first experiment.

3. Results

3.1. Temporal texture-segregation VEP at Oz

Fig. 2a shows averaged potentials for all four conditions (A–D) for a single subject. Homogenous displays (B and D) evoked motion onset potentials, representing the 'low-level' components in response to this stimulus. Waveforms in B and D look similar to each other, apart from a shift of about 40 ms which reflects the delayed motion onset in condition D. Heterogenous displays (A and C) evoke the same motion onset responses and, additionally, components specific for the texture being segregated. They are rather alike but differ strongly from waveforms in B and D. Conditions A and C closely resemble their mean (E in Fig. 2a), whereas B and D deviate more from their mean (F in Fig. 2a). The putative texture segregation component is revealed by subtracting F from E, yielding G in Fig. 2a. It mainly consists of a sharp negativity (N2) with a peak amplitude of about −6 μV at a latency of about 180–200 ms and another negativity (N3) with an amplitude around −3 μV and a latency of approximately 250–350 ms above the Oz position in this subject. Table 1(panel B) shows N2 peak amplitudes and latencies for all 10 subjects individually. Latencies range from 157.5 to 195 ms and amplitudes vary from −2.8 to −7.8 μV.

The grand average across 10 subjects (Fig. 2b) shows the same general waveform, albeit with smaller amplitudes. In the average, the N3 is followed by another negativity with a latency of approximately 350–400 ms. Since we recorded from an occipital region referencing against a frontal position, it is likely that negative potentials with latencies above 300 ms are indeed positive potentials originating from the frontal cortex.

3.2. Temporal texture-segregation VEP at other occipital positions

In five of the 10 subjects we recorded from five different electrode positions (see Section 2.1.6). The data show that the putative texture segregation component is largest and clearest at the central occipital position Oz, while being smaller and noisier at the more lateral occipital positions, especially at O3 and O4 (Fig. 2c). An ANOVA performed between the positions Oz, O1/O2, and O3/O4 revealed a significant difference ($F = 5.12 > F_{(2,22,0.95)}$, $p < 0.015$).

3.3. Effect of smaller checkerboard fields

Fig. 3b–e show the results of the second experiment. Each graph plots both, the tsVEPs elicited by checkerboards with 4×4 medium sized squares (thick lines) and 16×16 small squares (thin lines). We did not find any clear difference between these tsVEPs in any of the three

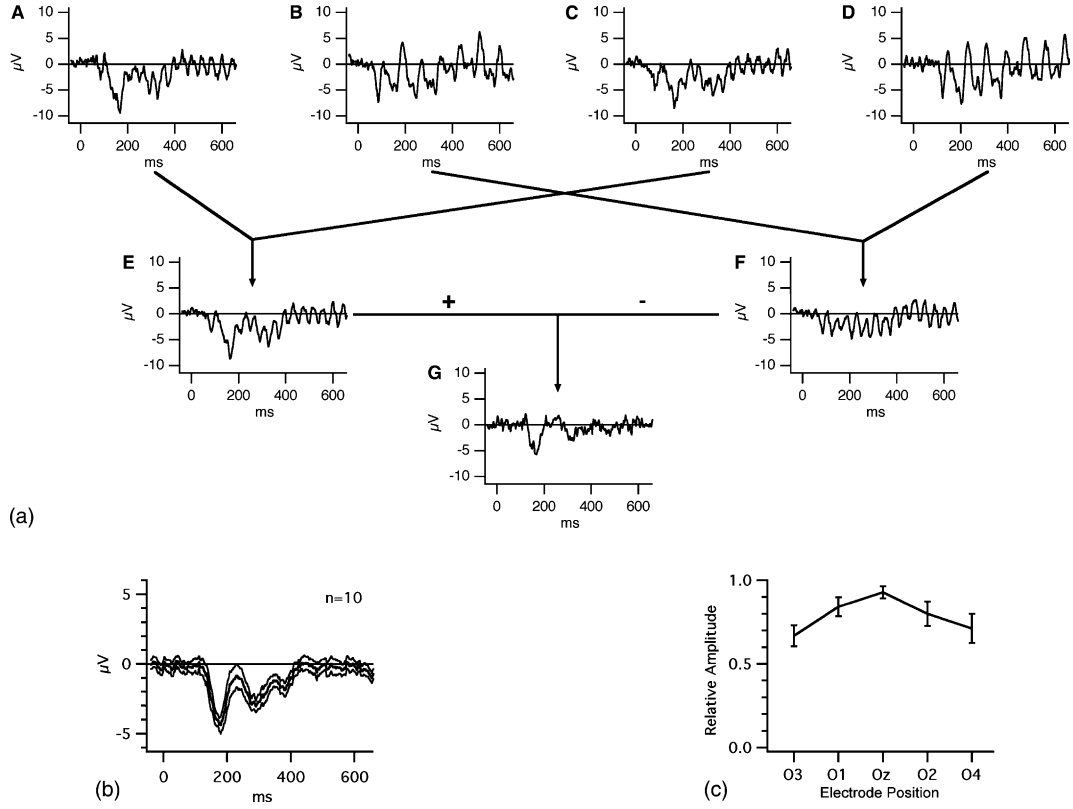

Fig. 2. (a) Graphs (A–D) show averaged cortical responses for one subject for each of the four conditions. Graph E is the average of graphs A and C (heterogenous conditions) and contains both, low-level components and tsVEP. Graph F is the average of graphs B and D (homogenous conditions) and thus presents merely low-level components. The difference between E and F, as shown in G, eliminates low-level components and exposes the tsVEP. (b) Grand average (mean ± standard error) for the tsVEP across 10 subjects. As in graph G in Fig. 2a, two negativities, a N2 at 180 ms and a N3 at 270 ms, are the most prominent components specific for texture segregation. Another negativity, N4 at approximately 380 ms, appears rather late, this component probably reflects cognitive events. (c) Relative amplitudes of the tsVEP at occipital positions O3, O1, Oz, O2, and O4 across five subjects. The mean amplitude decreases significantly from the central occipital to the lateral occipital positions.

subjects (Fig. 3b–d), nor in the mean across them (Fig. 3e). A comparison using the paired *t*-test and the randomization test did not reveal significant differences for the amplitudes ($p \approx 0.97$ and 0.48, respectively) or the latencies ($p \approx 0.42$ and 0.84, respectively). Raw data are given in Table 1(panel B).

4. Discussion

The results of experiment 1 demonstrate that time-based segregation such as the one used in a previous study (Kandil & Fahle, 2001) evokes a specific potential in the human VEP which mainly consists of two negativities. The earlier and narrower N2 has a peak latency of about 180–200 ms and a peak amplitude of about −5.6 μV, whereas the later N3 (latency: 280 ms) has a smaller amplitude of −3 μV.

The overall appearance of this potential and the specific parameters of its components resemble the tsVEPs for motion-defined checkerboards as described in earlier publications. Bach and Meigen (1997, 1998) found tsVEPs to be "similar across [visual] dimensions", all consisting of (a) negativities that (b) peak at latencies around 200 ms and with (c) peak amplitudes between −3 and −10 μV. In contrast, Fahle and coworkers (Fahle et al., 2003) found tsVEPs to be dissimilar between some of the visual dimensions investigated (color, luminance, motion and orientation). The tsVEP found here resembles the one they found for the motion condition as far as the overall form and the latencies and amplitudes of the peaks are concerned. As Fig. 4 demonstrates, the tsVEP obtained here has rather typical latency and amplitude values, as compared to tsVEPs evoked by checkerboards using differences in other visual dimensions (Bach & Meigen, 1997; Fahle et al., 2003).

The comparison between the amplitudes obtained at five different occipital recording sites indicates that the tsVEP originates from the cortex below the central occipital electrode (Oz). This finding is in accordance with earlier experiments cited in the introduction using single-cell recordings, VEPs and fMRI, and confirms that the tsVEP recorded here does not emanate from frontal areas (with reversed polarity).

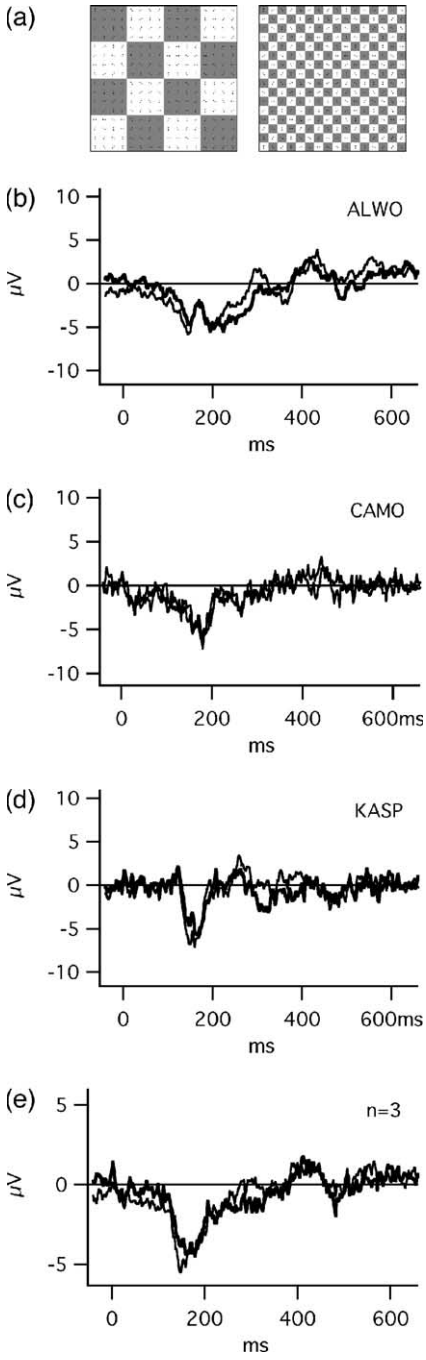

Fig. 3. Comparison between tsVEPs evoked by checkerboards with large versus small squares (a). In all three subjects (b–d) and in their mean (e), the VEPs elicited in both conditions are of comparable amplitude and latency. However, subjects readily detect the figure for large checkerboard squares but do not perceive any clear figure when the squares are small. Hence, the tsVEP reflects the perception of inhomogeneity of the display rather than detection of definite (cognitive) forms.

In experiment 2 we explored whether the perception of a display as inhomogenous suffices to evoke the tsVEP or else the perception of a distinct form is required. To investigate this possibility, we used heterogenous displays with a structure clearly visible to our

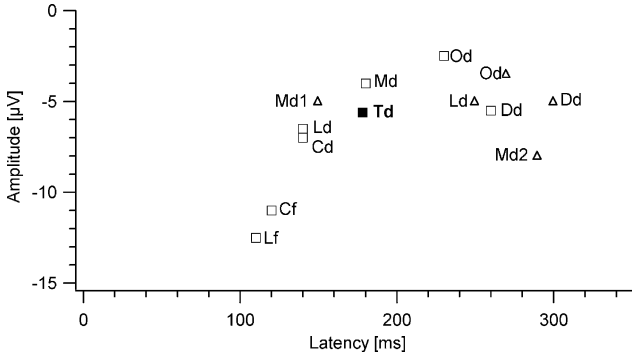

Fig. 4. Comparison of amplitude and latency values between the temporal tsVEP found here (filled square) with the tsVEPs found previously (line triangles: Bach & Meigen, 1997, line squares: Fahle et al., 2003) using differences in other visual dimensions to define the stimuli: L = luminance, C = color, D = disparity, M = motion, O = orientation, T = time; f = squares are completely filled in, d = squares merely contain a set of identical dots or line elements.

subjects and others which our subjects merely perceived as inhomogenous but not as segmented into a clear pattern. The tsVEPs obtained in the two conditions do not appear to differ from each other. Hence, we conclude that the appearance of the tsVEP based on temporal delays between different parts of a stimulus is a robust component reflecting the perception of inhomogeneity of the displays rather than requiring the (conscious) detection of any clear form, or figure.

Acknowledgements

We would like to thank Marc Repnow for his help with the computer code and technical assistance and Michael Bach for helpful discussions. Supported by Deutsche Forschungsgemeinschaft (SFB 517).

Animated QuickTime® demonstrations of the stimuli can be requested from the corresponding author.

References

Adelson, E. H., & Farid, H. (1999). Filtering reveals form in temporally structured displays. *Science, 286*, 2231a.

Bach, M., & Meigen, T. (1990). Electrophysiological correlates of texture-segmentation in human observers. *Investigative Ophthalmology and Visual Science (ARVO Suppl.), 31*, 104.

Bach, M., & Meigen, T. (1997). Similar electrophysiological correlates of texture segregation induced by luminance, orientation, motion and stereo. *Vision Research, 37*, 1409–1414.

Bach, M., & Meigen, T. (1998). Electrophysiological correlates of human texture segregation, an overview. *Documenta Ophthalmologica, 95*, 335–347.

Bach, M., Schmitt, C., Quenzer, T., Meigen, T., & Fahle, M. (2000). Summation of texture segregation across orientation and spatial frequency: electrophysiological and psychophysical findings. *Vision Research, 40*, 3559–3566.

Fahle, M. (1993). Figure–ground discrimination from temporal information. *Proceedings of the Royal Society of London Series B: Biological Sciences, 254*, 199–203.

Fahle, M., Quenzer, T., Braun, C., & Spang, K. (2003). Feature-specific electrophysiological correlates of texture segregation. *Vision Research, 43*, 7–19.

Farid, H., & Adelson, E. H. (2001). Synchrony does not promote grouping in temporally structured displays. *Nature Neuroscience, 4*, 875–876.

Kandil, F. I., & Fahle, M. (2001). Purely temporal figure–ground segregation. *European Journal of Neuroscience, 13*, 2004–2008.

Kastner, S., De Weerd, P., & Ungerleider, L. G. (2000). Texture segregation in the human visual cortex: A functional MRI study. *Journal of Neurophysiology, 83*, 2453–2457.

Kiper, D. C., Gegenfurtner, K. R., & Movshon, J. A. (1996). Cortical oscillatory responses do not affect visual segmentation. *Vision Research, 36*, 539–544.

Lamme, V. A. (1995). The neurophysiology of figure–ground segregation in primary visual cortex. *Journal of Neuroscience, 15*, 1605–1615.

Lamme, V. A., Van Dijk, B. W., & Spekreijse, H. (1992). Texture segregation is processed by primary visual cortex in man and monkey. Evidence from VEP experiments. *Vision Research, 32*, 797–807.

Lamme, V. A., Van Dijk, B. W., & Spekreijse, H. (1993). Organization of texture segregation processing in primate visual cortex. *Visual Neuroscience, 10*, 781–790.

Lee, S. H., & Blake, R. (1999). Visual form created solely from temporal structure. *Science, 284*, 1165–1168.

Meigen, T., & Bach, M. (1993). Perceptual ranking versus visual evoked potentials for different local features in texture segregation. *Investigative Ophthalmology and Visual Science, 34*, 3264–3270.

Skiera, G., Petersen, D., Skalej, M., & Fahle, M. (2000). Correlates of figure–ground segregation in fMRI. *Vision Research, 40*, 2047–2056.

Usher, M., & Donnelly, N. (1998). Visual synchrony affects binding and segmentation in perception. *Nature, 394*, 179–182.