A new variant of the Ouchi illusion reveals
Fourier-component-based processing

Hiroshi Ashida
Graduate School of Letters, Kyoto University, Kyoto 606-8501, Japan;
e-mail: ashida@bun.kyoto-u.ac.jp; also, ATR Human Information Science Laboratories,
2-2-2 Hikaridai, Seika-cho, Kyoto 619-0288, Japan

Akiyoshi Kitaoka
Department of Psychology, Ritsumeikan University, Kyoto 603-8577, Japan

Kenzo Sakurai
Department of Psychology, Tohoku Gakuin University, 2-1-1 Tenjinzawa, Izumi-ku, Sendai 981-3193,
Japan

Received 17 March 2003, in revised form 15 April 2004; published online 15 April 2005

Abstract. We report that anomalous motion illusion in a new variant of the Ouchi figure is well
predicted by the strength of its Fourier fundamentals and harmonics. The original Ouchi figure
consists of a rectangular checkerboard pattern surrounded by an orthogonal rectangular checker-
board pattern, in which illusory relative motion between the two regions is perceived. Although
this illusion has been explained in terms of biases in integrating one-dimensional motion signals
to determine the two-dimensional motion direction, the physiological mechanism has not been
clarified. With our new stimuli, which consisted of thin lines instead of rectangles, we found that
the perceived illusion is drastically reduced when the position of each line element is randomly
shifted. This is not predicted by simple models of local motion integration along the visible edges.
We demonstrate that the relative amplitude of the relevant Fourier fundamentals and harmonics
leads to a quantitative prediction. Our analysis was successfully applied to other variants of the
Ouchi figure (Khang and Essock 1997 Perception 26 585–597), closely predicting the reported
rating. The results indicate that the underlying physiological mechanism is sensitive to the Fourier
components of the stimuli rather than the visible edges.

1 Introduction
The Ouchi illusion (Ouchi 1977; Spillmann et al 1986), which is illusory relative motion
seen in a figure that consists of rectangular checkerboard patterns, as seen in figure 1a,
is considered to reflect the brain function that integrates local motion signals into
two-dimensional motion vectors (Hine et al 1995; Fermüller et al 2000; Mather 2000).
Retinal motion that originates in either image movement or eye movement can cause
this illusion (Spillmann et al 1993). Mather (2000) discussed that the integrated two-
dimensional motion direction is biased towards the normal vector of the longer edge
of the rectangular element, because motion signals are less variable along the longer
edge than along the shorter edge. Fermüller et al (2000) discussed the statistical basis
of such biases, which is not dependent upon a specific model but is inherent in the
geometry in calculating the optic flow from one-dimensional outputs of the motion
sensors. These studies provided the computational basis of the illusion, characterising
the illusion as an outcome of ecologically valid operation of motion integration.

The physiological mechanism, however, is not necessarily an instantiation of
computational theories (see Marr 1982). In this case, the mechanism might not directly
respond to the visible-edge structure. As it is generally conceived that the visual-
motion mechanism has a low-pass spatial-frequency tuning, which defines the coarse
structure rather than the fine details, it is worth considering the spatial-frequency profile
of the stimuli. Fourier series expansion reveals sinusoidal components in diagonal
orientations in the checkerboard pattern (figure 2). The two fundamentals form a
so-called plaid pattern. Ashida (2002) suggested that the Ouchi illusion is understood
Figure 1. (a) A variant of the Ouchi figure similar to the original one. The inner disk appears to float around as one moves the page or the eyes. The best illusion is seen with diagonal movement. (b) The new figure, made of lines instead of rectangles.

Figure 2. (a) A checkerboard pattern used in the Ouchi figure. (b) Sum of the two fundamentals that form a plaid pattern. (c) and (d) The fundamentals in two orientations. The white arrows represent the motion direction of the overall pattern (to the top right); the black arrows represent the component motion vectors. Note that the two component vectors lie on the same side of the pattern motion vector (type-II plaid).
in terms of the perceived motion direction of type-II plaid patterns, in which the normal motion vectors of the two components lie outside the pattern motion vector (see figures 2c and 2d) because the perceived motion direction of type-II plaid is biased toward the component vectors (Ferrera and Wilson 1987). To create an Ouchi pattern that moves diagonally, which is optimal for this illusion (Mather 2000), the centre and the surrounding regions in the Ouchi figure should be biased in opposite directions, leading to robust relative motion signals.

However, it has not been easy to prove this Fourier-component-based explanation because of difficulty in manipulating the Fourier components without changing the edges or other element structures. To overcome this difficulty, we have designed a new variant of the Ouchi figure, which consists of line elements instead of rectangles (figure 1b). With this pattern, we can alter the Fourier spectrum by randomly shifting the positions of the elements, while the elements themselves are not changed.
Figures 3c and 3d show the Fourier amplitude spectrum of the central patterns of figures 3a and 3b, respectively, derived by using two-dimensional fast Fourier transform (FFT). The random shift alters none of the local-edge structure, the number of elements, or the stimulus contrast. The average energy of the image, which is defined as the sum of the squared deviations of contrast (Khang and Essock 1997a), is therefore preserved. Changes in the perceived illusion, if any, would not be explained by edge-based models but could, instead, be explained in terms of the changes in the Fourier spectrum, especially the reduction of the amplitude of the relevant components.

In fact, as one can see in figure 3b, the illusion is largely reduced by the random shifts. We confirmed this observation by a rating experiment, and compared the results with the computed sum of amplitude of the relevant Fourier components in the images. (An interactive demonstration can be viewed on the Perception website at http://www.perceptionweb.com/misc/p5060/.)

2 Rating experiment
2.1 Methods

2.1.1 Stimuli. We created black-and-white images by printing the stimuli onto white A4 paper using a laser printer (600 dpi), the contrast being nearly 100%. We decided to use printed images as these can have high spatial and temporal resolution, and smooth motion, which is not possible on a computer screen. Also, we allowed the observers to inspect the images without constraint, except for instructions on how to hold the paper and swing it, to avoid disturbing deliberate judgment on rating for naïve observers. The results were actually consistent across observers, as indicated by the small confidence intervals, and also across the two laboratories.

For convenience, we will describe the size of the stimuli in pixels according to the original computer-generated images. A pixel formed a 0.25 mm × 0.25 mm square when printed. The reference stimuli were the Ouchi figures, like figure 1a, that consisted of rectangular elements 32 × 8 pixels in size. The test stimuli were the figures with thin black rectangles (32 × 2 pixels) on a white background. For both standard and test stimuli, the element orientation was vertical in the central disk and horizontal in the outer region. The size of the whole square region was 512 × 512 pixels (129 mm × 129 mm), and that of the central disk was 256 pixels (64 mm) in diameter. The elements were truncated at the borders. At a distance of 40 cm, the stimulus area subtended approximately 18 deg of visual angle.

The positions of the line elements in the test stimuli were randomly shifted from the regular pattern (such as that shown in figure 1b), within a certain range that was manipulated as an independent variable. For convenience, the amount of shift was described by the maximum shift size in the direction in which the line was pointing (i.e., vertical for the inner disk, and horizontal for the outer region). We tested four shift sizes: 0, 8, 16, and 24 pixels. The maximum shift in the orthogonal direction was a quarter of the shift in the direction in which the line was pointing. In the last case, for example, each random shift was within 24 × 6 pixels. Figures that included overlaps or connections of elements were discarded. Eight patterns were created for each shift size. The initial spatial phases inside and outside the central disk were independently set for each figure, so that the eight patterns were not identical, even without the random shifts.

2.1.2 Procedure. We asked the observers to rate the magnitude of each illusion. First, the experimenter handed a standard stimulus to the observer and asked him/her to move it back and forth diagonally, from top right to bottom left, at a frequency of about once a second and at a distance of approximately 40 cm. The observers viewed the central area of the stimuli binocularly without concertedly pursuing or fixating.
After confirming that they saw illusory relative motion, the observers were given a test sheet and were instructed to move it in the same way. Each observer was then asked to rate the illusion strength of the stimulus by giving it a natural number score, compared with a score of 10 for the standard stimulus. Observers were allowed to view standard stimuli and test stimuli as many times as they wished, but only one at a time.

The four shift sizes were tested one by one in a random order. Four repetitions of this test block were conducted for each observer. Different sets of stimuli were used for each presentation, randomly chosen from the eight prepared sets. Initial practice trials, with separately prepared stimuli, confirmed that the observers understood the method. Experiments were conducted under normal room lighting conditions (500 – 1000 lx).

2.1.3 Observers. Seventeen students from Kyoto University and Tohoku Gakuin University (twelve female and five male, aged between 21 and 28 years with an average age of 22.9 years), who were naïve to the purpose of the experiment, participated as paid volunteers.

2.2 Results

Figure 4 shows the average rating scores as a function of the maximum shift size. As expected, the magnitude of illusion decreased as the maximum shift size increased. A repeated-measures analysis of variance (ANOVA) revealed a significant effect of the shift size ($F_{3,48} = 76.6, p < 0.001$), with significant differences between all pairs of shift sizes (Tukey’s WSD, $p < 0.01$). Without random shifts, the rating score was not significantly different from 10 ($t_{16} = 1.18$), indicating that our stimulus was as effective as the original Ouchi pattern.

The sharp decrease in the magnitude of the illusion with the random shifts is not predicted by local motion signals for each line segment, and calls for a more global analysis. In the next section, we examine how the results are consistent with predictions according to the amplitude of the relevant Fourier components. We do not propose a specific biological model, but reveal constraints on the biological algorithm that is considered sensitive to the Fourier components of the images.

![Figure 4. Results of the rating experiment (filled circles) plotted with the results of the image analysis (the computed amplitude scores, normalised to the maximum rating score) for various low-pass cut-off frequencies (open symbols with lines). Both ratings and computed amplitude scores are plotted against the maximum shift size (see text for definition). The error bars show the 95% confidence intervals of the rating scores.](image-url)
3 Image analysis

3.1 Method

The line pattern has a structure of Fourier fundamentals and odd harmonics similar to that of a checkerboard, except for the nonoblique components that, in this case, are vertical (figure 3c). These fundamentals and harmonics in oblique orientations are considered the basis of the illusion, as we discussed in section 1. We, therefore, computed the strength of these oblique components to allow comparison with the rating results. For convenience, we refer to the Fourier components as fundamentals and harmonics according to the spatial frequencies in the original nonshifted image (see figures 3a and 3c), although the shifted image is no longer periodic and does not have genuine fundamentals.

MATLAB® software (The MathWorks, Inc.) was used for the computation. The size of the images was 256×256 pixels, and they contained only vertical elements to fit the global Fourier analysis. The size and the separation of the line elements were the same as in the rating experiment. An image with shift \( \hat{\theta} \) therefore contained 16 horizontal cycles and 4 vertical cycles. Four images were created for each shift size, in which the starting spatial phase, as well as the shift of each element, were randomly set for each image. The amplitude score was the sum of the amplitudes of fundamental and relevant harmonics, computed as follows. First, the amplitude spectrum of each image was computed by using two-dimensional FFT. Let \( a(u, v) \) be the discrete amplitude spectrum, where \( u \) and \( v \) denote horizontal and vertical spatial frequencies in cycles per image, respectively, and let \( f_x \) and \( f_y \) be the horizontal and vertical fundamental frequencies, respectively (see figure 3c). In all the images here, \( f_x \) = 16 and \( f_y \) = 4. Then, the amplitude score \( A_s \) is defined as:

\[
A_s = \sum_{i=-n_u}^{n_u-1} \sum_{j=-n_v}^{n_v-1} \left\{ a((2i+1)f_x, (2j+1)f_y)w([2i+1]f_x, [2j+1]f_y) \right\}
\]

where

\[
w(u, v) = \begin{cases} 
1 & \text{if } \sqrt{u^2 + v^2} \leq r \text{ and } v \leq u \\
0 & \text{otherwise}
\end{cases}
\]

The Nyquist frequency was \( 256/2 = 128 \) cycles per image, and the numbers of horizontal and vertical odd harmonics (in each quadrant) were thus \( n_u = ((128/f_x)/2 = 4 \) and \( n_v = ((128/f_y)/2 = 16 \), respectively. The summation area was further confined within the spatial frequency \( r \) by multiplying the binary weight \( w(u, v) \). As seen in the definition of \( w \), the harmonics whose orientation was not within 45° from the vertical were excluded, because they do not form type-II plaid in the same way as the fundamentals do. We confirmed that this operation was not crucial, at least qualitatively. To assess the contribution of harmonics, \( r \) was varied between 18 (fundamentals alone) and 200 (all relevant harmonics) cycles per image. The computed scores were averaged across four different images for each shift size, and were normalised to the maximum rating score.

3.2 Results

In figure 4, the computed amplitude scores are plotted together with the rating scores. In general, the amplitude scores decreased as the shift size increased, like the rating scores. The fundamentals alone did not give a good fit to the rating scores, but adding the harmonics led to improvement, which indicates that the harmonics significantly contribute to the illusion. This result is consistent with the finding by Khang and Essock (1997a) that a rectangular pattern yields stronger illusion than the fundamentals alone.

The best fit, judged from the least root-mean-square error, was obtained when the harmonics below 110 cycles per image were added. This limit of the highest spatial frequency corresponds to 12 cycles deg\(^{-1}\) at the distance of 40 cm, which is roughly...
consistent with the human spatial-frequency sensitivity for slowly moving stimuli (e.g., Kelly 1979). The correlation between the rating and the computed scores was very high ($r^2 = 0.995$) in this best case, and the correlation remained fairly high in all cases except for the fundamentals-alone condition ($r^2 > 0.9$). The estimated limit of 12 cycles deg$^{-1}$ might be imprecise owing to our simple analysis, but the robustness of the overall results confirms the dominant role of the fundamentals and some harmonics.

3.3 Application to previous results
The above analysis took into account the amplitude of fundamentals and harmonics in the same spatial frequency and orientation. It is therefore valid only when the stimuli have the same basic structure. We applied the same analysis to some of the images used by Khang and Essock (1997a) that satisfy this condition. From the first group of stimuli in their experiment 1, we chose rectangular, trapezoidal, triangular, and sinusoidal patterns (their figures 1a–1d). The other two could not be compared, because the sawtooth pattern has fundamentals whose spatial frequency along the luminance gradient is doubled, and the added sinusoidal pattern does not contain oblique Fourier components. We also tested the second group of stimuli used by Khang and Essock (1997a) (a rectangular pattern, its fundamentals, and its harmonics) (their figure 2a).

Images were created to reproduce the pattern in the central region of the above stimuli. Images were $240 \times 240$ pixels in size, with $f_x = 20$ and $f_y = 4$. The amplitude score was computed in the same way as described in the previous section. The cut-off spatial frequency ($r$) was 40 cycles per image which corresponded to 12 cycles deg$^{-1}$ according to their configuration.

Figure 5 shows the computed scores, plotted together with the results of Khang and Essock (1997a) that were read from their figures 3a and 3b. The computed scores were scaled to the rating score of the rectangular stimulus, which had the highest score of all the patterns. In figure 5a, it is clear that the computer scores closely simulate the rating data. In figure 5b, the computed scores showed a pattern of decreasing illusion, which is consistent with the rating data. The rating score for the scaled rectangular stimulus was particularly high compared with the computer score, but we can interpret this discrepancy as follows. The contrast of the scaled rectangular stimulus was lowered

![Figure 5](image-url)
by a factor of 0.62 from that of the rectangular stimulus. Although the averaged energy (Khang and Essock 1997a) and the sum of Fourier components both predict a proportional decrease of the illusion, the rating score was lowered only by a factor of about 0.8, which caused the deviation from the computed score. We speculate that this was due to the contrast normalisation over the stimulus area. In a typical normalisation model, the output of a detector is divided by the sum of the outputs of the neighbour-hood detectors (Heeger 1994). The output is less suppressed if the overall contrast is lowered, resulting in larger responses than predicted by a linear assumption as in our analysis. This effect, if any, should be negligible for the other stimuli, in which the range of luminance was constant.

4 General discussion

In summary, we found that the illusion decreases when the positions of the line elements are randomly shifted, as predicted by the computed amplitude scores of the Fourier fundamentals and harmonics. This finding challenges explanations based on the motion response to the edges, but is favourable for the idea that the physiological mechanism underlying the illusion is sensitive to the relatively low spatial-frequency components of the image. In this sense, the results bridge the gap between the phenomenology, or computational theories, and physiologically motivated models of two-dimensional motion perception (eg Wilson et al 1992; Alais et al 1997; Bowns 2002). As the visual motion mechanism in our brain mostly receives input from the magnocellular pathway (eg Thompson 1993), which has a low-pass spatial-frequency tuning, it is more likely that the sharp edges are blurred for motion processing. Also, bandpass spatial filtering is assumed for the motion detector (eg Adelson and Bergen 1985; van Santen and Sperling 1985), which implies that each detector responds to the specific range of Fourier components. Also, note that the finding that the illusion is almost abolished for isoluminant colour-defined stimuli (Khang and Essock 1997a) is consistent with the known property of neurons in brain area MT that is almost colour-blind (eg Gegenfurtner and Hawken 1996).

We demonstrated above that the analysis can be applied to the results of Khang and Essock (1997a). Also, our idea is qualitatively consistent with the results which showed that the illusion became stronger at higher contrasts (Khang and Essock 1997a, figures 1c and 5g), because the amplitude of Fourier components is proportional to the contrast. We did not attempt a quantitative comparison because Khang and Essock used nonstandard test patterns (their figures 2c and 4n), but their result of nearly linear increase of magnitude (their figure 5g) is in good agreement with our results. We cannot, however, gain insights into properties such as optimal size and spatial-frequency tuning (Hine et al 1995, 1997; Khang and Essock 1997a, 1997b; Ashida 2002), because we compared only the patterns that had the same basic structure. As we did not consider the integration across the two areas, ie the inside and the outside of the disk, we cannot discuss the results relating to the relative orientation of the two regions (Hine et al 1995, 1997; Khang and Essock 1997a) either. For a more complete understanding of this illusion, these properties should be explained in further studies with a model of global motion integration and segmentation.

We should note that the linear Fourier analyses have limited use when higher-order characteristics of the image have significant contribution. We found such a case when the illusory contours were prominent. In the central area of figure 1b, illusory horizontal contours that connect the end points of the abutting vertical lines are visible (Soriano et al 1996). There is physiological evidence that this type of illusory contour is processed in area V2 (von der Heydt and Peterhans 1989) or even in V1 (Ramsden et al 2001) of a macaque brain, and it is probably processed in similar early-level cortical sites in the human brain (Ohtani et al 2002). It is therefore likely that such illusory contour signals
are fed into the motion integration mechanism and contribute to the motion illusion (e.g. Wilson et al. 1992). In the stimuli we used, the illusory contours are gradually disordered by the random shifts and, thus, do not seem to have crucial effects. If, however, the elements are shifted only in the direction orthogonal to the line orientation, the illusory contour is preserved or even enhanced (figure 6). Although the amplitude of the relevant Fourier components radically decreases, one can perceive relatively strong illusion. An informal rating experiment suggested that the rating score was in between the best and the second best scores in our experiment, while the image analysis score was as low as the third one. We cannot specify a quantitative model, but it is quite likely that the illusory contour is the source of the illusion in this case. Note that illusory contours are prominent in our line stimuli, but these are not clear and probably not effective in the original rectangular pattern or in the variants by Khang and Essock (1997a).

As the computational theories suggest (Fermüller et al. 2000), the Ouchi-type anomalous-motion illusion should be the outcome of proper adaptation to the environment through evolution, rather than just revealing the defects or imperfections of the mechanism. Though, as far as we know, only a few attempts have been made to link the Ouchi illusion directly to the existing knowledge about basic motion detection either in psychophysics or physiology, the results of this study indicate that we can expect further understanding of our motion-processing mechanism by further examination of motion illusions. Our analysis would not directly apply to all anomalous motion illusions, some of which do not even contain asymmetry in orientation (Pinna and Spillmann 2002). Other types of anomalous illusion that occur, based on different principles, are expected to reveal other aspects of motion processing.

Acknowledgments. We thank Hiroshi Ono and Johannes Zanker for their valuable help. This research was supported in part by the Telecommunications Advancement Organization of Japan, 21st-century Center of Excellence program (D-2 Kyoto University), MEXT, Japan, and Hayao Nakayama Foundation for Science, Technology and Culture. Portions of the results were presented at the 25th European Conference on Visual Perception (August 25–29, Glasgow, Scotland, UK).

References
Alais D, Wenderoth P, Burke D, 1997 “The size and number of plaid blobs mediate the misperception of type-II plaid direction” Vision Research 37 143–150
Ashida H, 2002 “Spatial frequency tuning of the Ouchi illusion and its dependence on stimulus size” Vision Research 42 1413–1420


Heeger D J, 1994 “The representation of visual stimuli in primary visual cortex” *Current Directions in Psychological Science* **3** 159 – 163


Hine T, Cook M, Rogers G T, 1997 “The Ouchi illusion: an anomaly in the perception of rigid motion for limited spatial frequencies and angles” *Perception & Psychophysics* **59** 448 – 455


Marr D, 1982 *Vision* (San Francisco, CA: W H Freeman)


Ouchi H, 1977 *Japanese Optical and Geometrical Art* (New York: Dover)


© 2005 a Pion publication
Conditions of use. This article may be downloaded from the Perception website for personal research by members of subscribing organisations. Authors are entitled to distribute their own article (in printed form or by e-mail) to up to 50 people. This PDF may not be placed on any website (or other online distribution system) without permission of the publisher.