Kybernetik 14, 63—70 (1973)

An Experimental Examination of a Hypothesis Connecting Visual Pattern Recognition and Apparent Motion

David H. Foster

Imperial College of Science and Technology, Applied Optics Section, Department of Physics, London, Great Britain

Received: April 1, 1973

Abstract

The validity of a hypothesis connecting the existence of a certain visual apparent-motion effect and the capacity of the visual system to recognize, in a special way, objects that have undergone certain transformations is subjected to experimental test. The objects under consideration are random-dot patterns and the transformations under consideration are rigid motions of the plane. For the two subjects examined, the hypothesis is verified, with confidence coefficient 90%, in 278 cases out of 279.

Introduction

The purpose of this work is to examine, experimentally, the validity of a hypothesis developed in another study (Foster, 1973) connecting the existence of a certain visual apparent-motion effect and the capacity of the visual system to recognize, in a special way, visual objects that have undergone certain transformations.

The hypothesis is not presented here in its original general form [for this see Foster (1973)], since we shall, in fact, be concerned with a quite particular version of it. Before formulating this version, we have to specify the objects and transformations under consideration and define the notions of visual recognition and apparent motion we employ here. Because this is an experimental investigation, our approach throughout will be less formally oriented than that in Foster (1973).

Let $\boldsymbol{R}^2$ denote a fixed 2-dimensional plane perpendicular to the visual axis (monocular situation), and let $\boldsymbol{R}^2$ have a fixed luminance distribution defined upon it. This constitutes the background field. The set F of visual objects that we shall consider consists of a finite number of 2-dimensional random-dot patterns $A_1, \ldots, A_n$ superimposed upon this background field. Their precise dimensions are given later. The transformations of $\boldsymbol{R}^2$ that we shall consider will be drawn from the group $E(2)$ of Euclidean (rigid) motions of $\boldsymbol{R}^2$. We actually use just the identity component of $E(2)$, that is, the set of those points in $E(2)$ which can be joined to the identity by a smooth curve in $E(2)$. Each element in this set has a unique representation as a pair (ϱ_θ, k) thus

$$(\varrho_\theta, k)(x) = \varrho_\theta(x) + k, \quad x \in \boldsymbol{R}^2,$$

where ϱ_θ is a rotation of the plane $\boldsymbol{R}^2$ about the origin by an angle θ, and k is an element of $\boldsymbol{R}^2$. Denoting by $SO(2)$ the group of all rotations ϱ_θ, we can thus represent the identity component of $E(2)$ as $SO(2) \times \boldsymbol{R}^2$.

Suppose that we are given two random-dot patterns A_i and A_j from F. We shall say that *visual recognition with respect to* $SO(2) \times \boldsymbol{R}^2$ of the pattern A_j as the pattern A_i has occurred if by visual inspection an affirmative answer can be given to the question: Given that the group $SO(2) \times \boldsymbol{R}^2$ preserves the metric structure of an object (i.e., the distances between points) and that it is the largest connected group having this property, is object A_j equal to $(\varrho_\theta, k)(A_i)$ for some $(\varrho_\theta, k) \in SO(2) \times \boldsymbol{R}^2$? (The patterns A_i will sometimes be referred to as *initial* patterns and the patterns A_j as *final* patterns.)

The apparent-motion effect we are concerned with is the following. If objects A_i and A_j (not necessarily random-dot patterns) are presented in rapid temporal sequence to the visual system, then, depending upon A_i, A_j, and the experimental conditions, a visual impletion effect occurs in which A_i appears to change smoothly into A_j (see Wertheimer, 1912; Kenkel, 1913; Neuhaus, 1930; Kolers, 1964). Sometimes referred to as phi motion (Schureck, 1960; Foster, 1972a, b), we here follow Kolers' convention (Kolers, 1972) and use the term *beta motion*.

We can now state the particular version of the hypothesis of Foster (1973) relevant to the above.

Hypothesis (Special Case)

Given an Euclidean motion $(\varrho_\theta, k) \in SO(2) \times \boldsymbol{R}^2$, if beta motion can be induced between one random-dot pattern A_i and its transform $(\varrho_\theta, k)(A_i)$, so that A_i appears to change smoothly and rigidly into $(\varrho_\theta, k)(A_i)$, then the visual system can recognize with respect to the identity component $SO(2) \times \boldsymbol{R}^2$ of $E(2)$ the pattern $(\varrho_\theta, k)(A_i)$ as the pattern A_i.

It is emphasized that the beta motion should preserve the metric structure of A_i throughout.

Our procedure in testing this hypothesis is straightforward: first see whether object $(\varrho_\theta, k)(A_j)$ is recognized with respect to $SO(2) \times \boldsymbol{R}^2$ as object A_i, and then see whether beta motion of the specified kind can be induced between A_i and $(\varrho_\theta, k)(A_j)$. (We consider cases in which both $i=j$ and $i \neq j$.) The recognition experiments involve stimulus exposure times about one-tenth of those of the apparent-motion experiments. Interaction between the two experiments through "pattern-learning" is consequently less with the above ordering than with the reverse; moreover, any biasing of the eventual outcome by this interaction is towards failing the hypothesis rather than verifying it.

It is remarked that there have been several investigations aimed specifically at determining the visual system's capacity to recognize rotated objects. See Arnoult (1954), French (1953), and, for a general review, Hake (1966). It is also pointed out that Kolers and Pomerantz (1971) have investigated the conditions under which "rigid-motion" beta motion and "deformation" beta motion occur, and in a separate study Kolers and Perkins (1969) have examined recognition rates for various orientations of letters. Discussion of these and related experiments is given in Kolers (1972).

Experimental Apparatus

A diagram of the experimental apparatus is shown in Fig. 1. It is similar to that employed in Foster (1972b). The channels labelled RLH, LLH and RH each formed Maxwellian view systems: RLH gave rise to the initial test-objects A_i, LLH the final test-objects $(\varrho_\theta, k)(A_j)$, and RH the fixed background field.

The five initial test-objects $A_1, \ldots, A_5$ were formed from the dot-pattern masks shown in Fig. 2 (these also being labelled $A_1, \ldots, A_5$). The rotation $A_j \to \varrho_\theta(A_j)$ was determined by the Dove prism DP (Fig. 1), and the translation $\varrho_\theta(A_j) \to (\varrho_\theta, k)(A_j)$ by the position of the mirror M to the left of the biprism B (Fig. 1). With both patterns A_i and $(\varrho_\theta, k)(A_j)$ visible, the stimulus field (apart from the background) appeared as in Fig. 3.

An infrared corneal-reflection fixation-monitor was employed. This allowed continuous inspection, by

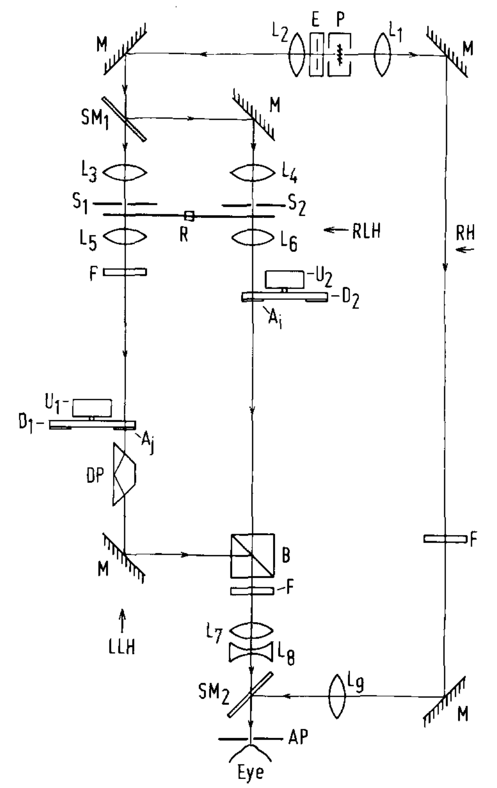

Fig. 1. The experimental apparatus: P light source; E electromagnetic shutter; M mirror; SM semi-reflecting plate; L lens; S stop; F filter; R rotating sector; A pattern mask; D disc; U electromagnetic rachet assembly; DP Dove prism; B biprism; AP artifical pupil

the experimenter, of the subject's fixation and also provided the subject with some feedback. The latter took the form of an audio signal with varying pitch. The system (not shown here) is described fully in Foster (1972b).

Detailed Description

The single light source P was a 12 V, 100 W quartz-iodine lamp with a compact coiled filament. This was run from a 12 V stabilized power supply (fluctuations of the light level being less than 0.25% of the mean). Light was taken from both sides of P and rendered parallel by the collimating lenses L_1 and L_2. The left-hand beam was split (amplitude-division in all cases) by the semi-reflecting plate SM_1, and the two resulting beams focussed by the lenses L_3 and L_4 onto the stops S_1 and S_2, respectively. The light from S_1 and S_2 was then recollimated by the lenses L_5 and L_6. The parallel light beam in channel LLH transilluminated the pattern mask A_j, and the parallel light beam in channel RLH transilluminated the pattern mask A_i. A_j was followed by the Dove prism DB. The two beams were brought together by the biprism P, after which each was brought, by the lenses L_7 and L_8, to a focus at the 2 mm artificial pupil AP. By means of the lens L_9 and semi-reflecting plate SM_2, the parallel beam of the single right-hand channel RH was also brought to a focus at AP. The aperture was

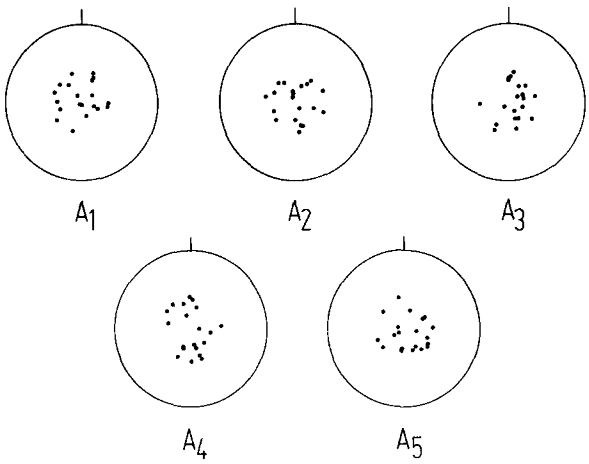

Fig. 2. The dot-pattern masks $A_1, \ldots, A_5$

completely filled with light. With the colour-correcting filters F inserted, the channels RLH, LLH and RH matched (in colour and brightness) from SM_2 onwards. The brightness of LLH and RLH was then raised by 1 log unit.

The exposure time of the whole left-hand side of the system, LLH and RLH, was determined by the electromagnetic shutter E, and could be set from 10 msec to 10 sec.

The two sets of five pattern masks were carried on discs D_1 and D_2, which were each attached to electromagnetic rachet assemblies U_1 and U_2. This facilitated the rapid interchange of each pattern mask in each channel. The Dove prism DP was mounted in a cylindrical bearing concentric with the optic axis of LLH. Rotation of the prism was by means of an electric motor, accurate orientation being achieved through a light-activated feedback system.

The alternate presentation of channels LLH and RLH in Experiment 2 (next section) was effected by a rotating 180° sector R, which was driven by an electric motor with electronic feedback stabilization. In Experiment 1 (next section), R was dispensed with.

The five dot-pattern masks shown in Fig. 2 are to scale. The locating circles (not visible when the masks were in position) gave angular subtenses at the eye of 1.04°. Each mask consisted of twenty dots (the dots being the transparent parts of the masks). The Cartesian coordinates of each dot were fixed from random-number tables with the constraint imposed that each point was to lie within a circle subtending 0.43° at the eye and concentric with the locating circle. Rotation-correlograms of the objects A_k, that is, the functions

$$c_{ij}(\theta) = \frac{\iint A_i(x, y) \cdot (\varrho_\theta(A_j))(x, y)\, dx\, dy}{[\iint |A_i(x, y)|^2\, dx\, dy]^{\frac{1}{2}}\, [\iint |A_j(x, y)|^2\, dx\, dy]^{\frac{1}{2}}}, \quad 0 \leqq \theta < 2\pi,$$

where the A_k are considered as luminance distributions defined locally on the plane $\boldsymbol{R}^2$, were evaluated, optically, from the masks (Fig. 2) using a laser-photodetector system. Rotation of the patterns was in all cases about the centre of the locating circles.

The full stimulus field reproduced in Fig. 3 is also to scale. The angular subtense at the eye of the centres of the two dot-patterns (i.e., the centres of the locating circles) was 0.76°. (The translation k was thus fixed in these experiments.) As viewed, the right dot-pattern underwent the rotation, with clockwise rotation being taken as positive. (The patterns were transposed at the biprism B, Fig. 1.) Fixation spots (not shown) were placed above and below the midpoint of the configuration (Fig. 3). These were visible at all times, in both Experiments 1 and 2.

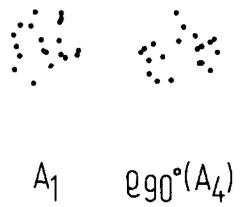

Fig. 3. The stimulus field for the case $(A_i, (\varrho_\theta, k)(A_j)) = (A_1, (\varrho_{90°}, k)(A_4))$

The angular subtense of the steady spatially-uniform background field was 13°. The retinal illumination of the background field was 2400 trolands, and the retinal illumination of the dots was 24000 trolands (superimposed). Colour temperature was 3200 °K.

Experimental Procedure

In both Experiments 1 and 2, the subject, using a dental bitebar, monocularly fixated a point midway between the two fixation spots. The fixation monitor was adjusted at the beginning of each experiment to give maximum response in this situation.

Experiment 1: Pattern recognition

For a fixed rotation-angle θ, a preselected pair of patterns $(A_i, (\varrho_\theta, k)(A_j))$ (as in Fig. 3) was presented to the subject for an exposure time of 200 msec. A forced-choice technique was employed: the subject was required to indicate (by means of a hand-buzzer) whether or not the pattern $(\varrho_\theta, k)(A_j)$ could be obtained from the pattern A_i by application to A_i of a rigid motion drawn from $SO(2) \times \boldsymbol{R}^2$. The subject was cued before each presentation, and test patterns were only presented when fixation, as indicated by the monitor, was judged to be good.

Twenty-five different pairs of patterns $(A_i, (\varrho_\theta, k)(A_j))$, $i, j = 1, \ldots, 5$, were presented for each fixed angle θ. The order of these presentations was chosen at random. Eighteen different fixed values of θ, namely 0°, 20°, 40°, ..., 340°, were used, with the order of selection also chosen at random. This full procedure was performed five times with different random orderings in each case.

For each angle θ and each pair of patterns (A_i, A_j), five trials were thus carried out to determine whether or not recognition with respect to the group of rigid motions $SO(2) \times \boldsymbol{R}^2$ of pattern $(\varrho_\theta, k)(A_j)$ as pattern A_i took place.

Experiment 2: Beta motion

For a fixed rotation-angle θ, two preselected patterns A_i and $(\varrho_\theta, k)(A_j)$ were presented alternately (in the configuration of Fig. 3) to the subject for

500 msec each. The time-lag between the alternating presentations was zero. Observation of the stimulus was allowed for 5 sec. As in Experiment 1, a forced-choice technique was employed: at the end of the 5 sec observation period, the subject was required to indicate (by means of the hand-buzzer) whether or not the pattern A_i appeared to change smoothly and rigidly (in the plane) into the pattern $(\varrho_\theta, k)(A_j)$. The subject was cued before each observation period, and observations were only performed when fixation, as indicated by the monitor, was judged to be good. The number and ordering of the trials, $i, j = 1, \ldots, 5$, $\theta = 0°$, $20°$, $40°, \ldots, 340°$, were the same as in Experiment 1.

Two subjects were employed: FMF, who was slightly myopic and aged twenty-six years, and DHF (the author), who was myopic and aged twenty-seven years. In both cases, the apparent distance of the test stimuli (Fig. 1) was within the range of accomodation of the subject's naked eye. FMF was unaware of the purpose of the study.

Results

Some remarks on the approach to the analysis of the data must first be made.

For a fixed pair of patterns (A_i, A_j) and rotation angle θ, let $p(\theta)$ and $q(\theta)$ be the respective average relative frequencies of affirmative responses in the recognition and beta motion experiments; that is, if $x_N(\theta)$ and $y_N(\theta)$ are the respective numbers of affirmative responses in N trials, then

$$p(\theta) = \lim_{N\to\infty} \hat{p}_N(\theta) \quad \text{and} \quad q(\theta) = \lim_{N\to\infty} \hat{q}_N(\theta),$$

where $\hat{p}_N(\theta) = x_N(\theta)/N$ and $\hat{q}_N(\theta) = y_N(\theta)/N$.

The hypothesis under consideration is concerned solely with the existence or nonexistence of certain kinds of recognition and apparent motion; no allowance is made for a graded or probabilistic response. An appropriate interpretation of the results $0 < p(\theta) < 1$ and $0 < q(\theta) < 1$ must therefore be fixed upon. The following decision rule is adopted. If $p(\theta) > 0.5$, then we shall say that the visual system can achieve the specified recognition, and if $p(\theta) < 0.5$, then we shall say that it cannot. If $p(\theta) = 0.5$, then no decision will be made. Similarly, if $q(\theta) > 0.5$, then we shall say that the visual system can effect the specified beta motion, and if $q(\theta) < 0.5$, then we shall say that it cannot. If $q(\theta) = 0.5$, then no decision will be made.

Including the cases $p(\theta) = 0.5$, $q(\theta) = 0.5$, there are in this way nine possible outcomes, at each (A_i, A_j), θ, for the paired Experiments 1 and 2. Of these, the outcome

$$p(\theta) < 0.5 \quad \text{and} \quad q(\theta) > 0.5 \tag{1}$$

contradicts the hypothesis, the outcomes

$$\left.\begin{array}{l} p(\theta) = 0.5 \quad \text{and} \quad q(\theta) \geqq 0.5\,, \\ p(\theta) < 0.5 \quad \text{and} \quad q(\theta) = 0.5 \end{array}\right. \tag{2}$$

allow no comment, and all others support the hypothesis.

Since, in fact, only the estimates $\hat{p}_5(\theta)$ and $\hat{q}_5(\theta)$ of $p(\theta)$ and $q(\theta)$ are available, the above statements need to be modified. Set

$$\hat{p}_{15}(\theta_k) = \tfrac{1}{3}[\hat{p}_5(\theta_{k-1}) + \hat{p}_5(\theta_k) + \hat{p}_5(\theta_{k+1})]$$

and

$$\hat{q}_{15}(\theta_k) = \tfrac{1}{3}[\hat{q}_5(\theta_{k-1}) + \hat{q}_5(\theta_k) + \hat{q}_5(\theta_{k+1})]\,,$$

where $\theta_k = 20k°$, $k = 0, 1, \ldots 17$. (We assume that $p(\theta)$ and $q(\theta)$ do not vary too rapidly with θ.) Let $[p_1(\theta), p_2(\theta)]$ and $[q_1(\theta), q_2(\theta)]$ be 90% confidence intervals for $p(\theta)$ and $q(\theta)$, based on $\hat{p}_{15}(\theta)$ and $\hat{q}_{15}(\theta)$ (see, for example, Hoel (1962)). We now replace the statements of the previous paragraph with the following.

The outcome

$$p_1(\theta) < p_2(\theta) < 0.5 \quad \text{and} \quad 0.5 < q_1(\theta) < q_2(\theta) \tag{1'}$$

contradicts the hypothesis with confidence coefficient 90%, the outcomes

$$\left.\begin{array}{lll} p_1(\theta) < 0.5 \leqq p_2(\theta) & \text{and} & 0.5 < q_1(\theta) < q_2(\theta)\,, \\ p_1(\theta) \leqq 0.5 \leqq p_2(\theta) & \text{and} & q_1(\theta) \leqq 0.5 \leqq q_2(\theta)\,, \\ p_1(\theta) < p_2(\theta) < 0.5 & \text{and} & q_1(\theta) \leqq 0.5 \leqq q_2(\theta) \end{array}\right\} \tag{2'}$$

allow no comment, and all others support the hypothesis with confidence coefficient 90%.

Hence, if at some (A_i, A_j), θ the hypothesis is to be disproved (with confidence coefficient 90%), then (1') must be shown to apply. Graphs of the upper bound $p_2(\theta)$ and the lower bound $q_1(\theta)$ as functions of θ for all the cases (A_i, A_i), $i = 1, \ldots, 5$, and for a representative sample of five from the cases (A_i, A_j), $i \neq j$, are shown for each subject in Fig. 4. Linear interpolation was allowed for intermediate values of θ, that is, $\theta \neq 0°$, $20°$, $40°, \ldots, 340°$. The corresponding rotation-correlograms $c_{ij}(\theta)$ (see *Experimental Apparatus*) are also displayed for each pair (A_i, A_j).

In all, the different patterns (A_i, A_j) and rotation angles θ define for each observer 150 independent cases. For FMF it was found that 142 of these cases gave outcomes which do not satisfy (2'), and that for none of these is (1') true. Hence, for FMF, assertions on the validity of the hypothesis can be made in 142 cases.

All 142 support the hypothesis with confidence coefficient 90%. For DHF, it was found that 137 cases gave outcomes which do not satisfy (2′), and that for only one is (1′) true (see Fig. 4d, DHF, $\theta = 300°$). Hence, for DHF, assertions on the validity of the hypothesis can be made in 137 cases. 136 cases support the hypothesis with confidence coefficient 90% and one case contradicts the hypothesis with confidence coefficient 90%.

Although the outcome giving rise to the last assertion is significant in the sense defined, the failure of the hypothesis at this point is, in a qualitative way, not very marked (see Fig. 4d). We shall discuss the relative shapes of the recognition and apparent-motion curves of Fig. 4 in the next section.

Discussion

We have shown that on an "all-or-none" basis the observed data provide almost complete support for the validity of the hypothesis set out in the Introduction. On a less delimited basis, it is interesting to consider the variations of the recognition variable $p_2(\theta)$ and the apparent-motion variable $q_1(\theta)$ with angle θ at values of θ near 0° (Fig. 4). In those cases in which $p_2(\theta)$ is varying rapidly, $p_2(\theta)$ and $q_1(\theta)$ follow similar courses; see, for example, Figs. 4a and 4b. [We now concentrate upon matched-pair cases (A_i, A_i).] In Foster (1972a) we discussed the possibility of the hypothesis going in the opposite direction, that is, *if* recognition, *then* beta motion. In that $p_2(\theta)$ and $q_1(\theta)$ exhibit similar dependencies upon θ in the situations described, it does indeed appear that the hypothesis goes both ways. At values of θ away from 0°, however, this conjecture breaks down with the introduction of subsidiary peaks in the recognition variable which are not reflected in the apparent-motion variable. Observe, for example, Fig. 4d, θ close to 180°.

We have not considered the possibility that the improvement, in some cases, in the capacity to effect recognition with respect to $SO(2) \times \boldsymbol{R}^2$ at values of θ other than those near 0° is in fact due to the particular

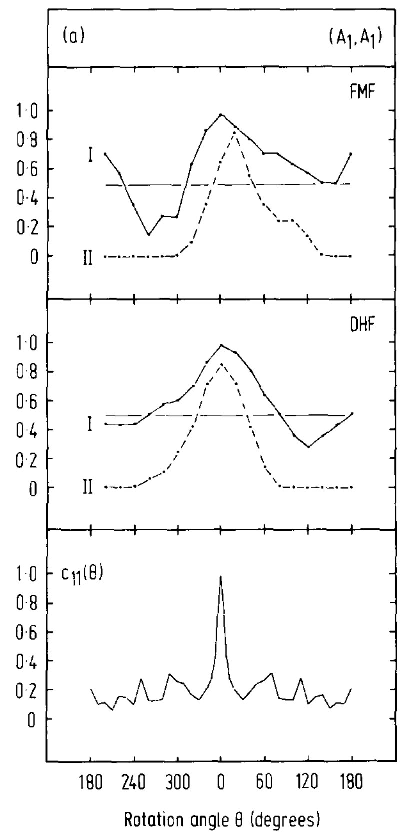

Fig. 4a

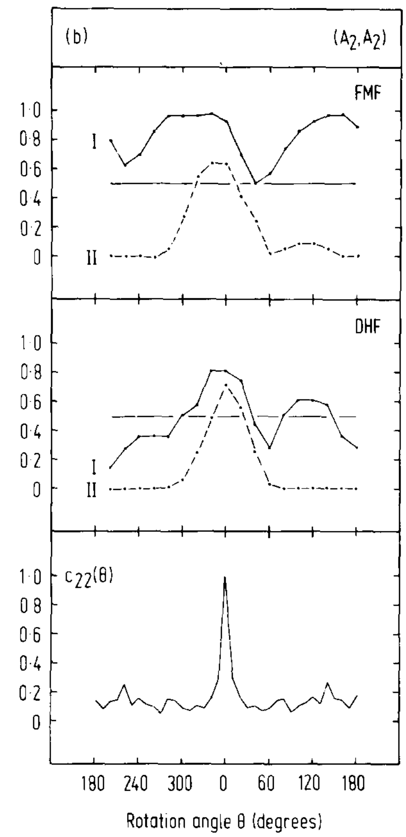

Fig. 4b

Fig. 4a–j. For various pattern pairs (A_i, A_j), the variables $p_2(\theta)$ (curve I) and $q_1(\theta)$ (curve II) for FMF and DHF, and the rotation correlograms $c_{ij}(\theta)$

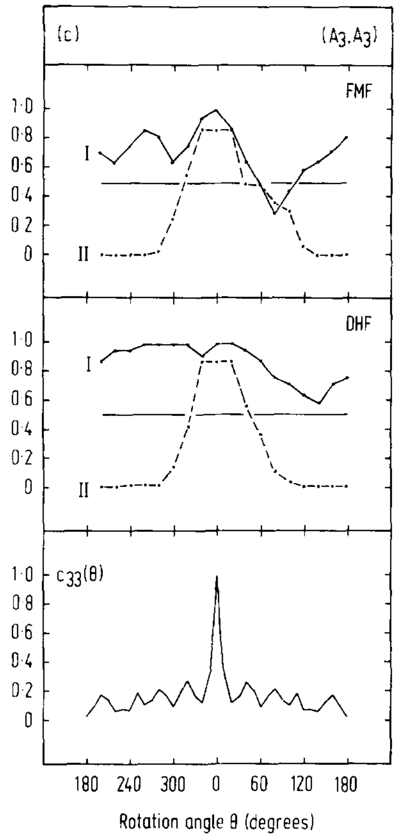

Fig. 4c

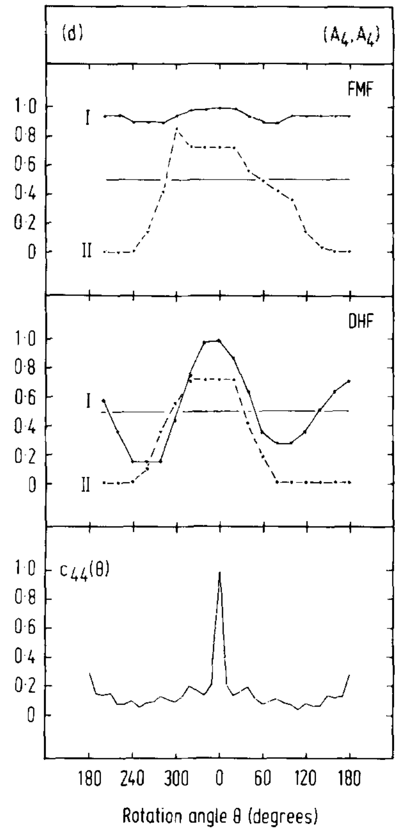

Fig. 4d

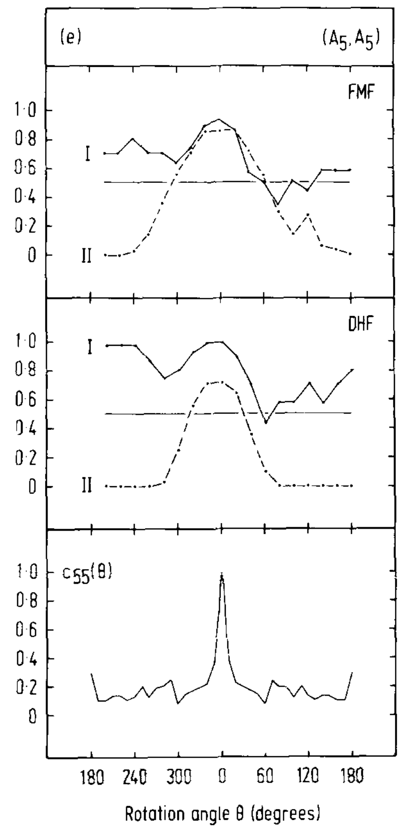

Fig. 4e

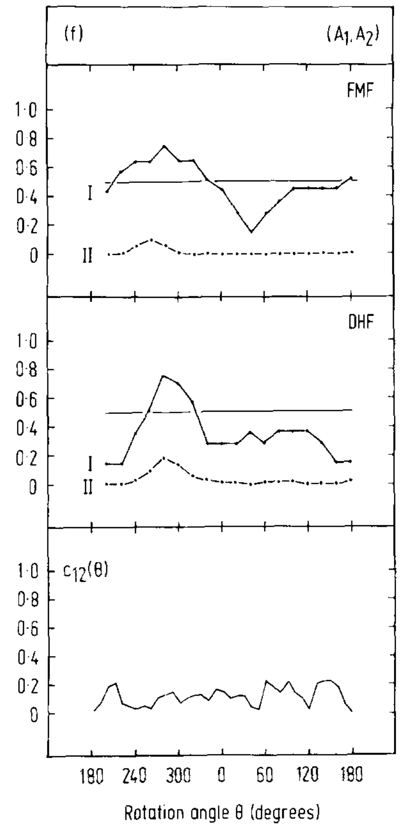

Fig. 4f

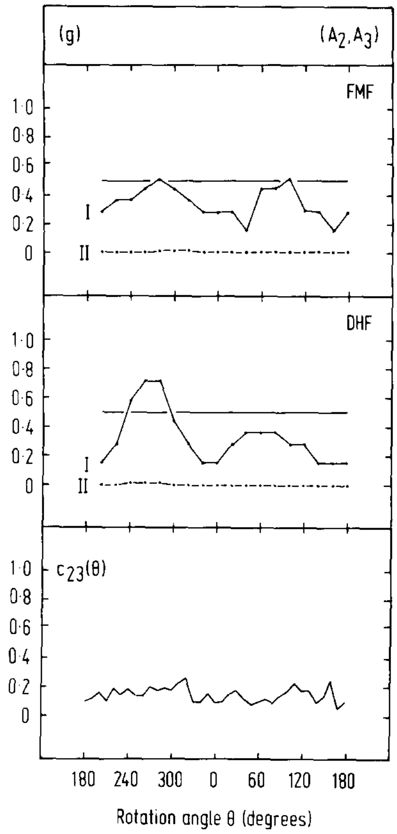

Fig. 4g

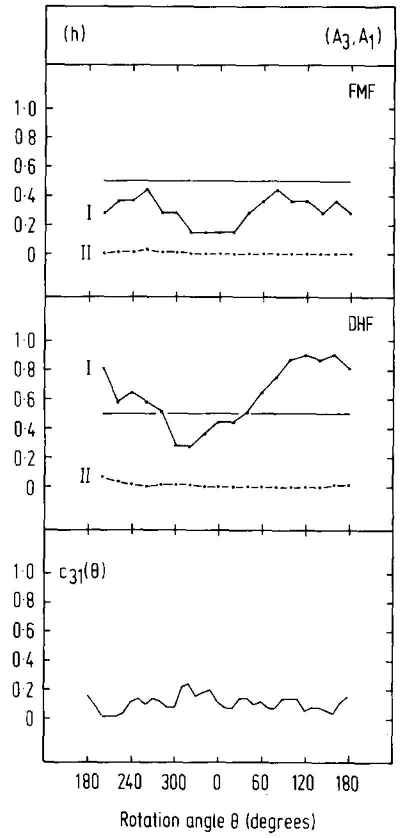

Fig. 4h

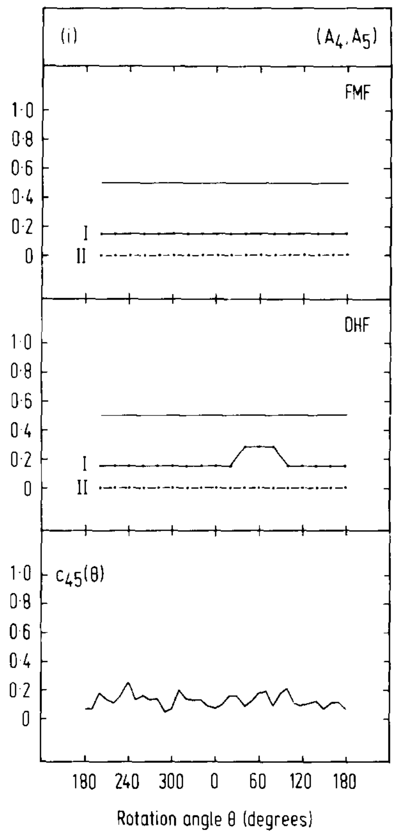

Fig. 4i

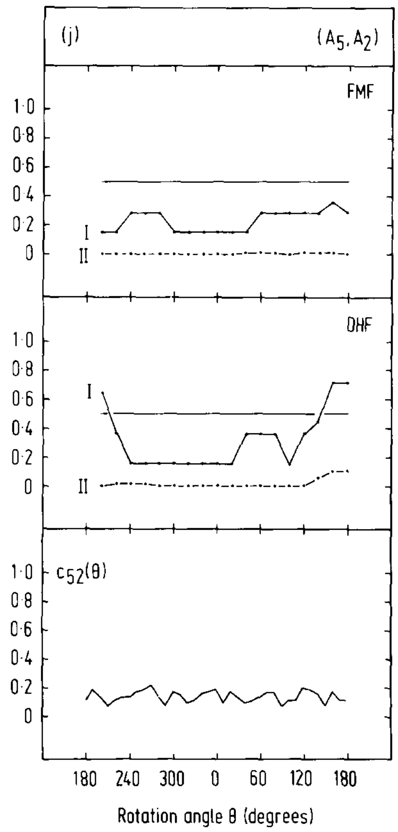

Fig. 4j

form of the pattern used. For example, the rotation-correlogram $c_{44}(\theta)$ associated with the pair (A_4, A_4) is seen to have relatively low values at all values of θ except at those near 0° and 180°; since A_4 is then similar to $\varrho_{180°}(A_4)$ we might well expect the observed variation of $p_2(\theta)$ with θ at θ near 180° to be like that at θ near 0°. [This need not be the situation for $q_1(\theta)$ with the ($\times$ 10) longer viewing period.] Nevertheless, the existence of extra peaks in the $p_2(\theta)$ curves cannot be explained by the shape of the rotation-correlograms in all cases (see Fig. 4a, FMF, and Fig. 4b, FMF and DHF.)

The preceding considerations do not affect the basic results of this investigation. In conclusion, we summarize what has been done. We set out to test the validity of a hypothesis connecting the existence of a certain rigid-motion apparent-motion effect and the capacity of the visual system to effect recognition with respect to the identity component of the group of rigid motions of the plane. For the two subjects examined, this hypothesis was verified, with confidence coefficient 90%, in 278 cases out of 279 that permitted a well-defined decision to be made.

References

Arnoult, M. D.: Shape discrimination as a function of the angular orientation of the stimuli. J. exp. Psychol. **47**, 323–328 (1954)

Foster, D. H.: A method for the investigation of those transformations under which the visual recognition of a given object is invariant: I. The theory. Kybernetik **11**, 217–222 (1972a)

Foster, D. H.: A method for the investigation of those transformations under which the visual recognition of a given object is invariant: II. An example experiment: the group of rotations $SO(2)$ acting on a Landolt ring. Kybernetik **11**, 223–229 (1972b)

Foster, D. H.: A hypothesis connecting visual pattern recognition and apparent motion. Kybernetik **13**, 151—154 (1973)

French, R. S.: The accuracy of discrimination of dot patterns as a function of angular orientation of the stimuli. USAF Air Training Command, HRRC, Res. Bull. 53-3, March 1953

Hake, H. W.: Form discrimination and the invariance of form. In: Uhr, L. (Ed.): Pattern recognition. New York: Wiley 1966

Hoel, P. G.: Introduction to mathematical statistics. 3rd Ed. New York: Wiley 1962

Kenkel, F.: Untersuchungen über den Zusammenhang zwischen Erscheinungsgröße und Erscheinungsbewegung bei einigen sogenannten optischen Täuschungen. Z. Psychol. **67**, 358–449 (1913)

Kolers, P. A.: The illusion of movement. Sci. Amer. **211**, 98–106 (1964)

Kolers, P. A.: Aspects of motion perception. Oxford: Pergamon Press 1972

Kolers, P. A., Perkins, D. N.: Orientation of letters and their speed of recognition. Perception and Psychophys. **5**, 275–280 (1969)

Kolers, P. A., Pomerantz, J. R.: Figural change in apparent motion. J. exp. Psychol. **87**, 99–108 (1971)

Neuhaus, W.: Experimentelle Untersuchung der Scheinbewegung. Arch. ges. Psychol. **75**, 315–458 (1930)

Schureck, P. J.: Studies in the perception of apparent visual movement. Austr. J. Psychol. **12**, 101–115 (1960)

Wertheimer, M.: Experimentelle Studien über das Sehen von Bewegung Z. Psychol. **61**, 161–265 (1912)

Dr. D. H. Foster
Imperial College of Science
and Technology
Dept. of Physics
Prince Consort Road
London SW7 2BZ, Great Britain