

# 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 study improves upon a previous hypothesis which claims a connection between the existence of a certain visual apparent-motion effect and the capacity of the visual system to recognize, in a very general sense, objects that have undergone certain transformations. In the present work a sharper hypothesis is developed by making use of the notion of visual recognition defined with respect to a fixed local Lie transformation group  $G$ . Considering the apparent-motion effect between an object  $A$  and some transform  $\sigma(A)$  of  $A$  as arising from the action of a certain 1-parameter family of transformations generated by the visual system, this hypothesis asserts that if the 1-parameter family lies entirely within  $G$ , then the visual system is capable of recognizing with respect to  $G$  the transformed object  $\sigma(A)$  as the original object  $A$ .

## Introduction

In an earlier study (Foster, 1972a), a method was put forward for the investigation of those transformations of Euclidean 3-space which have the property that for a given visual object their action leaves visual recognition of that object invariant. This method depends essentially on an assertion, also put forward in this paper, connecting visual pattern recognition and a visual apparent movement effect referred to as *phi motion*; it says that given an object  $A$  and some transform  $\sigma(A)$  of  $A$ , if *phi motion* can be induced between the two objects so that  $A$  appears to change smoothly into  $\sigma(A)$ , then  $\sigma(A)$  can be visually recognized as  $A$ . Recognition is formally regarded here as a setting-up by the visual system of a smooth one-to-one correspondence between the two objects. Investigations carried out by Kolers and co-workers indicate a similar line of thought (Kolers and Perkins, 1969; Kolers and Pomerantz, 1971; see also Kolers, 1972). In keeping with the convention adopted by Kolers (1972), we shall in the present study use the appellation *beta motion* for what was previously called *phi motion* (see, however, Schureck, 1960).

It is known that *beta motion* can be induced between patterns which are, loosely speaking, quite disparate (Kolers, 1972). This fact is not necessarily inconsistent with the above assertion since the notion

of recognition associated with the latter is very general. It would be more useful, however, if the assertion took account of a more specific notion of visual recognition. In the present study we advance a hypothesis which does this; the notion of visual recognition we employ is that of recognition defined with respect to a fixed local Lie group of transformations. [With reference to the use of Lie transformation groups in the analysis of visual perception, the comprehensive writings of Hoffman (1964, 1966, 1968, 1970, 1971) should be noted. These, in part, concentrate upon relating neuron morphology to certain operations associated with Lie transformations groups, but no attempt at establishing such a relationship is made here. A review of some of Hoffman's early work is given by Julesz (1971).]

We start the argument by setting up some notation and definitions, and then formulate, precisely, our interpretation of visual recognition and the *beta motion* effect. The hypothesis connecting the two phenomena then follows as the result of a plausible conjecture. In a separate investigation (Foster, 1970), an experiment is described and data presented which support the hypothesis in the particular case of the group  $E(2)$  of Euclidean motions of the plane.

## The Connecting Hypothesis

We shall be concerned with the monocular case. Rather than consider the retina as the surface upon which the visual stimuli are to be defined (as in Foster, 1972a, b) we consider a fixed 2-dimensional plane perpendicular to the visual axis and define visual objects upon that. Extreme peripheral vision corresponding to visual angles greater than or equal to  $180^\circ$  will be ignored.

Let then  $\mathbf{R}^2$  ( $\mathbf{R}$  the reals) be a fixed 2-dimensional plane perpendicular to the visual axis. In conjunction with  $\mathbf{R}^2$  we consider at all times a fixed mapping  $C$  of  $\mathbf{R}^2$  into  $\mathbf{R}$ , the *background field*, which assigns to each point  $x \in \mathbf{R}^2$ , unless otherwise indicated, some specified luminance  $C(x) \geq 0$ . (Suppose white-light

stimuli.) We then define a visual *object* or *pattern* on  $\mathbf{R}^2$  as (at least) a mapping of a non-empty subset  $U_A$  of  $\mathbf{R}^2$  into  $\mathbf{R}$  such that  $A(x) \geq 0$  is the luminance of the object at the point  $x \in U_A$ . Neither background  $C$  nor objects  $A$  need be continuous functions and the domain  $U_A$  of  $A$ , which can coincide with  $\mathbf{R}^2$ , need not be open.<sup>1</sup>

Depending upon the occasion, we shall allow an object  $A$ , or more strictly its domain  $U_A$ , to inherit certain of the structure of the plane  $\mathbf{R}^2$  (structure taken in the general sense of Bourbaki, 1968). Consider, for example, the induced *topological* structure on  $U_A$  or the induced *metric* structure on  $U_A$  or the induced *differentiable* structure on  $U_A$ . Note that the first two induced structures exist for any  $U_A$ , but the last only when  $U_A$  is open in the topology of  $\mathbf{R}^2$ . Because of this, it will be assumed in the sequel that the domain of an object is always thus open. This causes no real loss in generality.

For a fixed open subset  $U$  of  $\mathbf{R}^2$ , we denote by  $\text{Diff}(U, \mathbf{R}^2)$  the set of all diffeomorphisms  $\sigma$  taking  $U$  onto an open subset  $\sigma(U)$  in  $\mathbf{R}^2$ . (For definiteness, let this differentiability be  $C^\infty$ .) We define the action of such a *transformation*  $\sigma \in \text{Diff}(U, \mathbf{R}^2)$  on the set  $F(U)$  of all objects with domain  $U$  by setting

$$(\sigma(A))(q) = A(\sigma^{-1}(q)) \quad \text{for } A \in F(U) \quad \text{and} \quad q \in \sigma(U) \quad (1)$$

which assigns to each point  $q$  in the domain of the transformed object  $\sigma(A)$  the luminance at its preimage. Whenever we write  $\sigma(A)$  this is what we mean.

We now turn to the formulation of our notion of recognition. All statements involving sets, mappings, etc., which relate to a visual process should be taken to be defined only to within visual indistinguishability (see Zeeman, 1962) whether or not this is explicitly declared on each occasion.

As was said in the Introduction, the visual recognition of one object  $A$  as some other  $B$  is regarded in Foster (1972a) as the establishing by the visual system of a certain smooth one-to-one correspondence between  $A$  and  $B$ . This correspondence is understood to be luminance preserving. More formally, we have a diffeomorphism  $\sigma$  between the domains  $U_A$  and  $U_B$  such that  $B(\sigma(p)) = A(p)$  for all  $p$  in  $U_A$ , or, using (1), such that  $B = \sigma(A)$ . The viewpoint is taken in the present study that recognition as a setting-up of an association which preserves differentiable structure

is the most primitive of a number of different kinds of recognition (each corresponding to a different local Lie group of transformations). [It could be argued that even this notion of "recognition at its most primitive" is not sufficiently fundamental, and that a more natural notion is that of the setting-up of an association preserving just cardinality. We keep the above definition since it is the most appropriate within the present framework. See Cassirer (1944) and Julesz (1969, 1971) for some relevant discussion.]

Let  $G$  be a connected local Lie group and let  $\tau: G \times U \rightarrow \mathbf{R}^2$  be a local Lie group of transformations defined on a fixed open subset  $U$  of  $\mathbf{R}^2$  (see Appendix). Denote by  $\tau_g$  the mapping  $p \in U \rightarrow \tau(g, p) \in \mathbf{R}^2$ , and by  $\iota$  the mapping  $g \in G \rightarrow \tau_g \in \text{Diff}(U, \mathbf{R}^2)$ . Assume  $\iota$  to be injective. Let  $G \times U \rightarrow \mathbf{R}^2$  preserve some certain structure on  $U$  and be maximal on  $U$  with respect to this structure-preserving property (that is, there exists no local Lie group of transformations  $G' \times U \rightarrow \mathbf{R}^2$  preserving the same structure on  $U$  and such that  $G$  is a proper subset of  $G'$ ). For brevity, we shall sometimes refer to the  $G$ -invariant structure of  $U$ . By way of example, consider the group  $E(2)$  of Euclidean motions of the plane. The group  $E(2)$  leaves invariant the distance between any two points in  $\mathbf{R}^2$  and hence preserves the induced metric structure on  $U$ . We shall continue to refer to each member of  $\iota(G) \subset \text{Diff}(U, \mathbf{R}^2)$  as a diffeomorphism or transformation, though whenever appropriate this should be understood to be short for *structure-preserving* diffeomorphism or transformation.

For a fixed local Lie group of transformations  $\tau: G \times U \rightarrow \mathbf{R}^2$ , *visual recognition with respect to  $G$*  of an object  $A \in F(U)$  as some other object  $B$  is defined as the establishing by the visual system of a diffeomorphism  $\tau_g$  between domain  $U$  and domain  $U_B$  for some  $g \in G$  such that  $B(\tau_g(p)) = A(p)$  for all  $p \in U$ . [The usefulness of requiring the set of structure-preserving transformations to exhibit a local Lie group structure will become apparent later.] In the experimental situation, we infer that recognition with respect to  $G$  has been effected if (and only if) by visual inspection an affirmative answer can be given to the question: Given that  $\tau: G \times U \rightarrow \mathbf{R}^2$  preserves some certain structure on  $U$  and that it is maximal on  $U$  with respect to this structure-preserving property, is the object  $B$  equal to  $\tau_g(A)$  for some  $g \in G$ ?

Obviously, a necessary condition for recognition with respect to  $G$  of object  $A$  as object  $B$  is that  $B = \tau_g(A)$  for some  $g \in G$ . It is certainly not always sufficient, see for example Dearborn (1899), Arnould (1954), and Foster (1970).

<sup>1</sup> Given an object  $A$  on  $\mathbf{R}^2$ ,  $U_A \neq \mathbf{R}^2$ , we could, according to our definition, treat the restriction  $C|_{U_A^c}$  of the background field to the complement  $U_A^c$  of  $U_A$  as the object, and the original object  $A$  as the restriction of the background field to  $U_A$ ; consider the vase-faces ambiguous figure (Rubin, 1921).

We next consider the formal description of beta motion. Note that, as indicated in the Introduction, this term is being used here to describe Wertheimer's *optimal movement* (Wertheimer, 1912; see also Kenkel, 1913, and Neuhaus, 1930). In the sense that we can replace this internal subjective phenomenon by an equivalent external real motion and that the latter is uniquely defined to within visual indistinguishability, we refer to the phenomenon taking place *on the plane*  $\mathbf{R}^2$ . Formally then, if  $F = \bigcup \{F(U') : U' \subset \mathbf{R}^2\}$  denotes the set of all objects defined locally on  $\mathbf{R}^2$ , given the sequential presentation to the visual system of some object  $A \in F(U)$  and some transform  $\sigma(A)$  of  $A$ ,  $\sigma \in \text{Diff}(U, \mathbf{R}^2)$ , *beta motion* between  $A$  and  $\sigma(A)$  is the generation by the visual system of a smooth time-parametrized curve  $\Omega(A)$  in  $F$  joining these two objects.

The curve  $\Omega(A)$  may be thought of as arising from the action of a 1-parameter family of transformations  $\psi_t : N \rightarrow \mathbf{R}^2$ ,  $t \in [0, t_1]$ , (see Appendix) satisfying  $\psi_0(p) = p$  and  $\psi_{t_1}(p) = \sigma(p)$  for all  $p \in U$ . That is,

$$\Omega(A) : t \in [0, t_1] \rightarrow \psi_t(A) \in F,$$

where the action of  $\psi_t$  on elements of  $F(U)$  is, as usual, defined by (1). We take this approach here and consider the visual system effecting the family<sup>2</sup>  $\psi_t$  on the plane  $\mathbf{R}^2$  in the sense defined above. Given the capacity to effect some family  $\psi_t$ , whether or not beta motion does in fact take place depends upon experimental conditions like the presentation (exposure) times of  $A$  and  $\sigma(A)$ , and the time lag between presentations. These will always be understood to be optimally chosen.

If a local Lie group of transformations  $\tau : G \times U \rightarrow \mathbf{R}^2$  is specified, then we shall be particularly interested in the case when for object  $A \in F(U)$  and transform  $\tau_{g_0}(A)$ ,  $g_0 \in G$ , beta motion can be induced between  $A$  and  $\tau_{g_0}(A)$  in such a way that the curve  $\Omega(A)$  in  $F$  is given by the action of a 1-parameter family  $\psi_t : U \rightarrow \mathbf{R}^2$ ,  $t \in [0, t_1]$ , lying entirely within  $\iota(G)$ , that is, when the mapping  $t \rightarrow \psi_t$  of  $[0, t_1]$  into  $\text{Diff}(U, \mathbf{R}^2)$  can be factored through the local Lie group  $G$  thus

$$\begin{array}{ccc} [0, t_1] & \xrightarrow{t \rightarrow \psi_t} & \text{Diff}(U, \mathbf{R}^2) \\ & \searrow \gamma & \nearrow \iota \\ & G & \end{array}$$

where the differentiable curve  $\gamma : [0, t_1] \rightarrow G$  joins the identity  $e$  of  $G$  to  $g_0$  in  $G$ .

We now develop our modified hypothesis connecting recognition and apparent motion. Suppose

<sup>2</sup>  $\psi_t$  short for  $\psi_t$ ,  $t \in [0, t_1]$ .

beta motion can be induced between some object  $A \in F(U)$  and some transform  $\sigma(A)$  of  $A$ ,  $\sigma \in \text{Diff}(U, \mathbf{R}^2)$ . This implies that the visual system can at least set up a bijection between  $A$  and  $\sigma(A)$  by means of the 1-parameter family  $\psi_t$ ; this bijection is just the mapping  $p \in U \rightarrow \psi_{t_1}(p) \in U_{\sigma(A)}$ . The family  $\psi_t$  is differentiable, and the visual system in setting up this bijection preserves the differentiable structure of  $A$  at every point in its passage  $\Omega(A)$  from initial to final form. We shall therefore assume this bijection to be (at least) a diffeomorphism. Let us now conjecture that the visual system's capacity to effect such a diffeomorphism in the beta motion situation implies its capacity to effect the same diffeomorphism in the recognition situation. This gives us, in essence, Proposition 1 of Foster (1972a). We take this further. Given some fixed local Lie group of transformations  $G \times U \rightarrow \mathbf{R}^2$ , suppose beta motion can be induced between the object  $A$  and some transform  $\tau_{g_0}(A)$ ,  $g_0 \in G$ , so that the situation described in the previous paragraph holds, that is, the 1-parameter family of transformations  $\psi_t : U \rightarrow \mathbf{R}^2$ ,  $t \in [0, t_1]$ , taking  $A$  into  $\tau_{g_0}(A)$  lies entirely within  $\iota(G) \subset \text{Diff}(U, \mathbf{R}^2)$ . In this case, the visual system in setting up the bijection  $p \in U \rightarrow \psi_{t_1}(p) \in U_{\tau_{g_0}(A)}$  preserves the  $G$ -invariant structure of  $A$  at every point in its passage  $\Omega(A)$  from initial to final form. As before, let us conjecture that the visual system's capacity to effect such a structure-preserving diffeomorphism in the beta motion situation implies the same for the recognition situation. Recalling the definition of recognition with respect to a local Lie group of transformations, we have the following.

#### Hypothesis

Let object  $A \in F(U)$  and local Lie group of transformations  $\tau : G \times U \rightarrow \mathbf{R}^2$  be fixed. Given transformation  $\tau_{g_0} \in \iota(G) \subset \text{Diff}(U, \mathbf{R}^2)$ ,  $g_0 \in G$ , if beta motion can be induced between  $A$  and transform  $\tau_{g_0}(A)$  so that the time-parametrized curve  $\Omega(A)$  joining  $A$  to  $\tau_{g_0}(A)$  is given by the action of a 1-parameter family of local transformations  $\psi_t : U \rightarrow \mathbf{R}^2$ ,  $t \in [0, t_1]$ , which admits a factorization through  $G$

$$\begin{array}{ccc} [0, t_1] & \xrightarrow{t \rightarrow \psi_t} & \text{Diff}(U, \mathbf{R}^2) \\ & \searrow \gamma & \nearrow \iota \\ & G & \end{array}$$

for some curve  $\gamma$  in  $G$ , then the visual system can recognize with respect to  $G$  the object  $\tau_{g_0}(A)$  as the object  $A$ .

## Discussion

Consider the case in which the set  $F(U)$  of all objects defined on a fixed open subset  $U$  of  $\mathbf{R}^2$  is restricted to a certain finite number of 2-dimensional random-dot patterns  $A_1, A_2, \dots, A_n$ , and in which the fixed local Lie group of transformations  $\tau: G \times U \rightarrow \mathbf{R}^2$  is the group  $E(2)$  of Euclidean motions of  $\mathbf{R}^2$ , with action restricted to  $U$ . The group  $E(2)$  is the semidirect product  $O(2) \times \mathbf{R}^2$  of the orthogonal group  $O(2)$  and the group of translations  $\mathbf{R}^2$ . Since the hypothesis is only concerned with those points in  $E(2)$  which can be joined to the identity by a curve in  $E(2)$ , we need just the identity component  $SO(2) \times \mathbf{R}^2$  of  $O(2) \times \mathbf{R}^2$ , where  $SO(2)$  is the group of (proper) rotations of  $\mathbf{R}^2$  about the origin. The rigid motion represented by the pair  $(\varrho_\theta, k) \in SO(2) \times \mathbf{R}^2$ , where  $\varrho_\theta$  is a rotation of  $\mathbf{R}^2$  about the origin by an angle  $\theta$ , is the transformation

$$x \rightarrow \varrho_\theta(x) + k, \quad x \in \mathbf{R}^2.$$

In this case, the hypothesis reduces to the following. Given a Euclidean motion  $(\varrho_\theta, k) \in SO(2) \times \mathbf{R}^2$ , 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 transform smoothly into  $(\varrho_\theta, k)(A_i)$  under the action of a 1-parameter family of rigid motions, then the visual system can recognize with respect to the identity component of  $E(2)$  the pattern  $(\varrho_\theta, k)(A_i)$  as the pattern  $A_i$ .

This situation, with 2-dimensional random-dot patterns as objects and the identity component of  $E(2)$  as the fixed local Lie group of transformations, was investigated experimentally in Foster (1970). The observed data were found to provide substantial support for the validity of the particular version of the hypothesis set out above.

*Acknowledgments.* The author wishes to thank Prof. W.C. Hoffman, Prof. P.A. Kolers and Dr. J.O. Robinson for providing comments on the two previous studies.

## Appendix

Let  $U$  be a non-empty open subset of the plane  $\mathbf{R}^2$ ,  $I$  an interval in  $\mathbf{R}$ , and  $G$  a local Lie group (see Cohn, 1968).

(1) A local Lie group of transformations defined on  $U$  is a differentiable mapping  $\tau$  of  $G \times U$  into  $\mathbf{R}^2$  which satisfies:

(a) For each  $g \in G$ ,  $\tau_g: p \rightarrow \tau(g, p)$  is a diffeomorphism of  $U$  onto  $\tau_g(U)$ .

(b) If  $g_1, g_2, g_1 g_2 \in G$  and if  $p, \tau_{g_2}(p) \in U$ , then

$$\tau_{g_1 g_2}(p) = \tau_{g_1}(\tau_{g_2}(p)).$$

( $G$  acts on the left.)

(c) For all  $p \in U$ ,  $\tau_e(p) = p$  ( $e$  the identity of  $G$ ).

(2) A 1-parameter family of transformations defined on  $U$  is a differentiable mapping  $\psi$  of  $I \times U$  into  $\mathbf{R}^2$  such that for each  $t \in I$ ,  $\psi_t: p \rightarrow \psi(t, p)$  is a diffeomorphism of  $U$  onto  $\psi_t(U)$ . (The 1-parameter family is also denoted by  $\psi_t: U \rightarrow \mathbf{R}^2$ ,  $t \in I$ .)

## References

- Arnoult, M.D.: Shape discrimination as a function of the angular orientation of the stimuli. *J. exp. Psychol.* **47**, 323–328 (1954)
- Bourbaki, N.: Theory of sets, Chapter IV. Paris: Hermann/Reading, Mass.: Addison-Wesley 1968
- Cassirer, E.: The concept of group and the theory of perception. *Philos. and Phenomenological Res.* **5**, 1–36 (1944)
- Cohn, P.M.: Lie groups, p. 48. London: Cambridge Univ. Press 1968
- Dearborn, G.V.N.: Recognition under objective reversal. *Psychol. Rev.* **6**, 395–406 (1899)
- 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.: An experimental examination of a hypothesis connecting visual pattern recognition and apparent motion. *Kybernetik* (in press)
- Hoffman, W.C.: Pattern recognition by the method of isoclines: A mathematical model for the visual integrative process. *Boeing Scientific Res. Labs. Math. Note No. 351*, Seattle (1964)
- Hoffman, W.C.: The Lie algebra of visual perception. *J. Math. Psychol.* **3**, 65–98 (1966)
- Hoffman, W.C.: The neuron as a Lie group germ and a Lie product. *Quart. appl. Math.* **25**, 423–440 (1968)
- Hoffman, W.C.: Higher visual perception as prolongation of the basic Lie transformation group. *Math. Biosci.* **6**, 437–471 (1970)
- Hoffman, W.C.: Memory grows. *Kybernetik* **8**, 151–157 (1971)
- Julesz, B.: Pattern discrimination. In: Reichardt, W. (Ed.): *Processing of optical data by organisms and by machines*, pp. 580–588. New York: Academic Press 1969
- Julesz, B.: *Foundations of cyclopean perception*. Chicago: The University of Chicago Press 1971
- Kenkel, F.: Untersuchungen über den Zusammenhang zwischen Erscheinungsgröße und Erscheinungsbewegung bei einigen sogenannten optischen Täuschungen. *Z. Psychol.* **67**, 358–449 (1913)
- 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 Psychophysics* **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)
- Rubin, E.: *Visuell wahrgenommene Figuren*. Copenhagen: Gyldendalske Boghandel 1921
- Schureck, P.J.: Studies in the perception of apparent visual movement. *Australian J. Psychol.* **12**, 101–115 (1960)
- Wertheimer, M.: Experimentelle Studien über das Sehen von Bewegung. *Z. Psychol.* **61**, 161–265 (1912)
- Zeeman, E.C.: The topology of the brain and visual perception. In: *The topology of 3-manifolds and related topics*. New Jersey: Prentice-Hall 1962

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