

Visual perception of biological motion and a model for its analysis*

GUNNAR JOHANSSON

University of Uppsala, S:t Larsgatan 2, S-752 20 Uppsala, Sweden

This paper reports the first phase of a research program on visual perception of motion patterns characteristic of living organisms in locomotion. Such motion patterns in animals and men are termed here as biological motion. They are characterized by a far higher degree of complexity than the patterns of simple mechanical motions usually studied in our laboratories. In everyday perceptions, the visual information from biological motion and from the corresponding figurative contour patterns (the shape of the body) are intermingled. A method for studying information from the motion pattern per se without interference with the form aspect was devised. In short, the motion of the living body was represented by a few bright spots describing the motions of the main joints. It is found that 10-12 such elements in adequate motion combinations in proximal stimulus evoke a compelling impression of human walking, running, dancing, etc. The kinetic-geometric model for visual vector analysis originally developed in the study of perception of motion combinations of the mechanical type was applied to these biological motion patterns. The validity of this model in the present context was experimentally tested and the results turned out to be highly positive.

With very few exceptions, research on visual motion perception has dealt with simple patterns of mechanical motions. As a rule, rigid objects in rotary or translatory motion have been chosen as the distal stimuli, and the proximal stimulus patterns investigated have been projections of such objects in motion.

Stimulus patterns representing living animals in motion have rarely been studied. Michotte's (1963) study of perception of larva motion may be pointed to as an important exception. This relative lack of studies concerning visual perception of animal and human motion patterns has no relation to the biological importance of this type of visual perception. It seems evident that, throughout animal evolution, valid information about other animals' motion has achieved a very high survival value.

Our everyday experience also tells us that human vision not only detects motion directions in man and animals, but also distinguishes different standard types of limb motion patterns. We immediately see whether a person is walking, running, or dancing, and also if he is moving forward with identical speed in these three cases. It is also a common experience that our visual apparatus is very sensitive to small deviations from such standard patterns. We immediately recognize, for instance, a slight limp in walking, we distinguish between a tired and an elastic gait, etc. Furthermore, we think we sometimes can recognize a person exclusively from his style of walking, his gestures, etc.

The geometric structures of body motion patterns in man and higher animals (e.g., the patterns of walking) are determined by the construction of their skeletons. Human walking, for instance, as well as the same types of motion in most domestic animals can readily be described as combinations of several pendulum-like motions of the extremities relative to a joint. These motions are often combined with an elastic bending of

the spinal column. The time relations in the series of iterated pendulum motions as a rule are highly regular. This interaction brings about a body displacement with a rather constant speed.

Different types of human displacement, i.e., walking, running, dancing, skating, etc., are all built up from such combinations of pendulum-like motions, which are highly specific for the different types of motions. The typical character of the pendulum motions about the hip and the knee joints during walking of a young man and their phase relations are shown in Fig. 1.

The present study is the first in a planned series of investigations on perception of such rather complex patterns of live motion and their outcome in body displacement. It will provide an empirical and theoretical basis for forthcoming studies in this field. The aim here will be to study the visual information from some typical motion patterns of the human body when the pictorial form aspect for these patterns has been abstracted. For this purpose, it was important to develop methods for giving the Ss an adequate motion stimulation without any interference from figural perception.

METHODS

A Method for Isolating Information in Motion Patterns from Information in Form Patterns

When analyzing motion in physics, it has, as is well known, been found necessary to introduce the concept of *particle*, i.e., abstract mathematical points as carriers of the motion. In an analogous way, it has, in the experimental (and/or theoretical) study of motion perception, sometimes been found advantageous to work with artificially isolated stimulus elements. Bright or dark spots moving against a homogeneous, contrasting background are typical as elements in such experiments.

Allowing such elements to represent the motion structure to be studied experimentally makes it possible to load the structure with well-controlled motion information. A classical example is found in Wertheimer's (1923) demonstration of the "law of common fate." Wertheimer demonstrated that when in a static

*This research was supported by a grant from the Swedish Council for Social Science Research. Experimental assistance was provided by Ragnar Johansson.

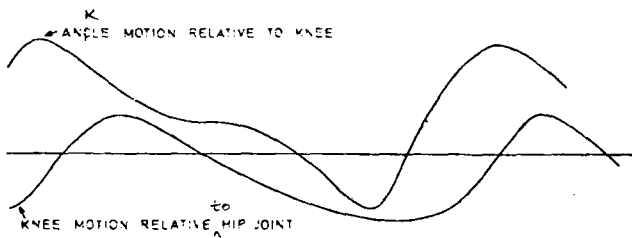


Fig. 1. The relative pendulum motion components of knee and ankle of a walking person. The two motions are plotted against their common time axis, and therefore the diagram also illustrates typical phase relations between the two motions.

"Gestalt" built up of a pattern of dots, some dots begin to move in a unitary way, the static form is broken up, and the moving dots form a new unit. Johansson (1950) has applied the same method to the analysis of motion perception in a rather programmatic way. Another example is Green's excellent study of coherence in motion patterns (Green, 1961).

This method of using bright spots as carriers of element motions in complex motion patterns has been applied also in the present study. It will allow us to analyze human motion patterns without interference from the pictorial information inherent in perception of the moving pattern.

It goes without saying that the method described has nothing to do with the old sensation-perception controversy recently reactivated by Gibson (1966). The point motions will not be thought of as some kind of motion sensations. The model here is chosen for methodological reasons and has no theoretical implications. In fact, all experiments applying this method in a highly convincing way demonstrate how a theory of sensation type is extremely inadequate for motion perception.

Methods for Producing the Proximal Patterns of Stimulus Motion

Two bright spots seen in motion against a structureless background are perceived as end points on a moving, invisible, rigid line or rod. In the same way, four points representing the corners of a square bring about visual information about motion of a rigid square (Johansson, 1950, 1964; Johansson & Jansson, 1968; Börjesson & von Hofsten, 1972, 1973).

From a mechanical point of view, the joints of the human body are end points of bones with constant length and at the same time the points of connection between such motion units. Consequently, the motion tracks of the main joints were chosen as representative motion elements.

So far, two similar methods for generating the proximal pattern of spot motion have been developed. The original one makes use of a film technique and, in a later alternative, a video recording technique is applied. Both of these methods are constructed for producing proximal stimuli representing perfect projections of the live motions under study. In a forthcoming study, systematic variations of such patterns will be treated, in which a computer technique will be applied.

Method 1

This is the originally developed method. Flashlight bulbs (6 V, .5 A) were fitted into small dull black funnel-shaped holders and attached to the main joints of an assistant dressed in tight-fitting dark clothes (see Fig. 2). The lamps were powered by a variable transformer. When the assistant walked over the studio floor along a rectilinear track, the motion tracks of these joints were recorded by a 16-mm film camera. There was no lighting in the studio except the faint light from the flashlight bulbs.

This method gives a very good recording of the motion tracks of the different joints without revealing any traces of the

background or the body contours. However, it has also some limitations and drawbacks. The arrangement with lamps on the body and their wire connections to a transformer is rather inconvenient and clumsy, and the problems were serious when other than rectilinear motion directions were introduced. Therefore, Method 2 was devised.

Method 2

Here a video technique is applied. It has many advantages when compared with the film technique. The principle, with bright spots representing the joints, is the same, but these spots are generated in another way with this technique. The possibilities of manipulating the signal relations in a TV record are taken advantage of. In this way, light-reflecting spots rather than light-emitting spots can be used.

Small patches of tape (15 mm ϕ), which have a surface of glass-bead retroreflective material ("reflex patches"), were attached over the joints of the assistant actor. In one variant, instead of using reflex patches, ribbons of this material were attached around the joints. (This is to be preferred in some studies where the actor is moving in curvilinear tracks.)

The actor is now flooded by the light from one or two searchlights (1,000-4,000 W) mounted very close to the lens of a TV camera. In this way, practically all the light hitting the reflex patches is reflected back into the camera. The result is an extremely high brightness contrast between the reflex patches and the background (and consequently a large difference in signal amplitude).

The movements of the actor are recorded by a video tape recorder. In the experiment, this record is displayed on the screen of a TV monitor, adjusted in a special way. Its contrast control is turned to its maximum setting, thus yielding a "supercontrast," which amplifies the brightness difference between the reflex patches and the actor's skin. The brightness control is set near its minimum. These control settings result in a blocking of all signals except the high-intensity signals from the

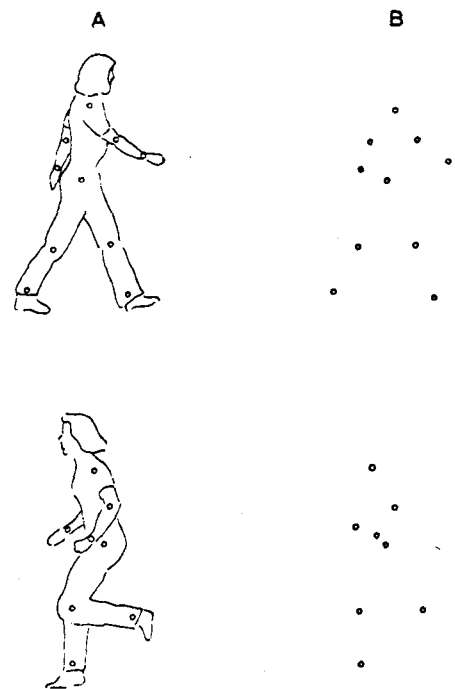


Fig. 2. Outline contours of a walking and a running subject (A) and the corresponding dot configurations (B).

reflex patches. These stand out as bright spots with sharp contours against a totally dark screen.

PERCEPTION OF WALKING AND OTHER RELATED HUMAN MOTION PATTERNS

Demonstration 1

Stimulus Material

The stimulus material will be described here as produced by Method 2.

Reflex patches were fastened to each side of the main limb joints of the "actor," as shown in Fig. 2A. The actor walked in a normal style over the scene from the left to the right and back, in a direction fronto-parallel to the camera lens. About eight steps in each direction were within the angle covered by the camera lens chosen. The start and turning were outside the angle of recording, and therefore the actor was walking into and out of the field of view. Figure 2B shows the isolated dot pattern in the same moment of motion, and Fig. 3 is an example of motion tracks of the elements during walking.

This motion pattern is displayed on a TV monitor and shown to the Ss. The visual angle subtended by the screen in the experiments was about 15 deg. Thus, the height of the swarm of bright spots representing the walker was about 3 deg of visual angle. (Ordinarily, in motion-and-space research, it is not recommended that the proximal stimulus be displayed on a visible screen. For discussion, see Johansson, 1964. In the present demonstrations, however, it was found to be acceptable.)

The Perception from Demonstration 1

The combination of spot motion in Demonstration 1 when looked at on a TV screen (or, in the film version, on a movie screen) always gives an O the most vivid impression of a person walking in fronto-parallel direction.

This motion pattern has been shown in many class demonstrations as well as under more strict experimental conditions. It always evokes the same spontaneous response after the first one or two steps: this is a walking human being! This perceptual effect has been observed without exception. Perceiving a walking person in the motions of these 10 spots seems to be equally as self-evident and natural as seeing a real man walking in a motion picture. The O has the freedom neither to combine the moving points in other groupings by an act of concentration nor to see these elements just as a series of unrelated points in motion. It might be added that when the motion is stopped, the set of elements is never interpreted as representing a human body (cf. Fig. 2B).

The moving dot pattern from Demonstration 1, together with a number of similar ones representing various types of human bodily activities, is included in a film in progress. This film illustrates some main results from the Uppsala research on visual motion perception.¹

Because of the manifest character of the perceptual response to this pattern of element motions, it is reported here as a demonstration rather than in terms of an experiment on identification. Before starting an analysis concerning the rationale for the perhaps

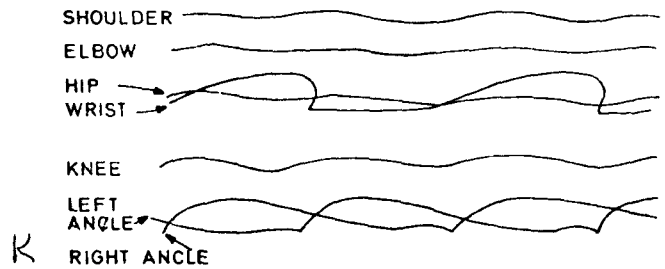


Fig. 3. Typical motion paths of seven elements representing the motions of the right side joints plus the ankle joint of the left leg of a walking person.

somewhat unexpected outcome of this stimulus configuration, some important generalizations will be reported.

Demonstration 2

In Demonstration 1, the distal motion path of the actor had a fronto-parallel direction. In this case, the analysis of the proximal pattern is especially simple, because just in this projection the constant lengths of the skeletal bones result in constant distances between their proximally projected end points.

Far more complex geometrically is the general case where the distal track can represent any angle between 0 and 90 deg to an O's fronto-parallel plane. Under these conditions of projection, the distally constant distances between couples of joints will be represented in the proximal stimulus by continuously changing distances between the elements representing these joints.

From the point of view of theory, this general case is of course more interesting than the fronto-parallel case, but also far more complex. Demonstration 2 shows that, obviously, there do not exist any extra problems for the visual apparatus in its spontaneous "treatment" of such patterns.

Stimulus Material

Either one of the two methods described above can be used for producing the stimulus material. The flashbulbs or the patches of reflex tape are now attached to the frontal side of the actor's main joints. It may be regarded as preferable in this case to attach stimulus patches (or flashbulbs) to the shoulder joints and two hip joints (12 elements instead of 10). However, this is not essential for the result.

The actor is instructed to walk toward the camera in directions between 80 and 45 deg to the fronto-parallel plane. The scenes are recorded on video tape or film. A lens with rather short focal distance is preferred because of the more drastic change in pattern size during an approach which is obtained in this way.

Result

This pattern has also been shown to a large number of Os both under well-controlled experimental conditions and on motion picture screens in class demonstrations. The outcome is wholly analogous to the results from

Demonstration 1. All Os immediately reported seeing a person walking toward them. Furthermore, the perceived direction of motion track roughly corresponds to the recorded one.

These results were found without any exception. Every O immediately reported the correct motion pattern. Also, in this case, the degree of perceptual vividness appears to be very high for all Os.

Demonstration 3

Human walking is also readily recognized when the number of recorded elements in the stimulus pattern is reduced. Five elements representing the hip-and-legs part of the pattern in Demonstrations 1 and 2 were used in Demonstration 3.

This element combination has been found to give correct responses irrespective of motion direction. Os always report seeing a walking human being. In most cases, the Ss describe this pattern as two walking legs. A few Ss, however, have described the upper point (points) as shoulders of a walking man.

Demonstrations of Other Types of Human Motion Patterns

The walking pattern was chosen as some kind of standard pattern in our studies of human motion. However, a number of other demonstration patterns have also been produced and studied. These patterns include running in different directions and also in a circular track, cycling, climbing, dancing in couples, various types of gymnastic motion, etc. In all these cases, spontaneous and correct identification of the types of activity has been made without exception.

TWO MAIN LINES OF ANALYSIS

How can 10 points moving simultaneously on a screen in a rather irregular way give such a vivid and definite impression of human walking? This question forms the main theoretical problem of this paper. We will now take the first steps toward its solution.

The first spontaneous answer to the above question perhaps would be a reference to previous experience. We would guess that there exists a heavy overlearning in seeing human walking, which makes it natural for us to see this motion pattern of the joints' motion as representing a walking man.

Surely, experience plays a certain role, perhaps also a rather important role. We can, for instance, state that the identification of the perceived motion constellations as representing human walking must be an effect of previous experience. But saying that the perceptual grouping of elements is made meaningful by previous visual experience of walking is not the same as saying that the grouping is *determined* by experience. Furthermore, in our case, it was most probably the first

time in their lives that our Ss had seen a walking pattern built up from moving spots. Why did these spots evoke the same impression of motion as the outlines of a walking man usually does?

Let us make the problem a little more explicit by raising the following question: Is the perceptual grouping to a human Gestalt determined by a recognition of the walking pattern, or is this recognition dependent on a spontaneous grouping in consequence of some general principles for grouping in visual motion perception?

I think that the best a priori guess would be to suppose that the definite grouping is determined by general perceptual principles, but that the vividness of the percept is a consequence of prior learning. If the grouping were mainly due to learning, we would expect that it might be possible to see, at will, the points as moving independently of each other or to see other structures.

As a matter of fact, my study of biological motion started from the theoretical position that the principles found in studies of perception of mechanical motions should also be revealed in the perceptual outcome from the complicated systems of biological motion. The vividness and strength of the visual motion configurations from biological motion patterns, however, were even more definite than expected. Perhaps the perceptual learning factor is responsible for this effect. In experimental studies to be reported here, as well as in some forthcoming studies, I will try to throw some light on this problem complex. In the next section, I will give an outline of the theoretical framework for a model able to include both mechanical and biological motion patterns.

The geometrical frame finally chosen for this analysis is not the traditional one, based on the parallel axiom, but the geometry of central perspective where convergent lines are also treated as parallel.

VISUAL VECTOR ANALYSIS

For many years, the present writer has been engaged, together with a number of co-workers, in research on visual motion perception. Data from this research have successively been condensed to a model for motion and space perception, i.e., the vector analysis model. The presently chosen geometrical frame for this analysis, however, is not the ordinary Euclidean one represented by the Cartesian coordinate system. Instead, it is the geometry of central projection from 3-space to 2-space (or 1-space) which has given the reference frame (cf. Johansson, 1971).

This model is not yet fully structured, but its main outline is established. It consists of a set of simple rules for a mathematical vector analysis of the proximal motion patterns, within the geometrical framework mentioned, the results of which form a near analogue to the corresponding percept.

Among the studies that yield the basic material for this model of perceptual vector analysis, or that are concerned with its theoretical structure, the following may be mentioned: Rubin (1927), Duncker (1929), Johansson (1950, 1958, 1964, 1966, 1971), Hochberg (1957), Marmolin & Ulfberg (1967), Johansson & Jansson (1968), Börjesson & von Hofsten (1972, 1973). The most complete review to date is found in Johansson (1971).

This model, which initially was worked out from studies of simple mechanical motion perception, has so far been found also to make theoretically understandable the perceptual outcome of biological motion patterns. Therefore, it will be tentatively applied here. Before proceeding to such an analysis, it is advantageous to repeat the main characteristics of this model. This will be done by (1) giving short explicit formulation of the basic principles, and (2) describing a few empirically found examples which have a special bearing for our present analysis of biological motion patterns.

Basic Principles of Visual Vector Analysis

Principle 1

Elements (= inhomogeneities, in the projected light distribution, see p. 202) in motion on the picture plane of the eye are always perceptually related to each other.

This principle was investigated in a systematic way in Johansson (1950) (cf. Wertheimer, 1923; Cohen, 1964; Johansson, 1971). As a matter of fact, it can be deduced from most experimental studies of motion perception.

Principle 2

Equal and simultaneous motions in a series of proximal elements automatically connect these elements to rigid perceptual units.

This is, strictly speaking, hardly more than a consequence of Principle 1 above. Far from trivial, however, is the following third principle, which in itself encompasses Principle 2 as a simple special case.

Principle 3

When, in the motions of a set of proximal elements, equal simultaneous motion vectors can be mathematically abstracted (according to some simple rules), these components are perceptually isolated and perceived as one unitary motion. More or less complex hierarchies of equally diverging perceptual motions often turn out to be the consequence of this visual stimulus analysis.

This third principle is, of course, the essential principle of perceptual vector analysis.

The term "equal" introduced in Principles 2 and 3 above has a very special definition. It is, as mentioned

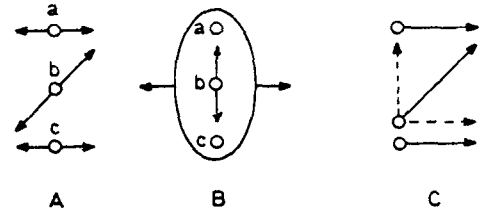


Fig. 4. Examples of the main principle in visual vector analysis. (A) Proximal pattern; (B) diagram of the percept from this stimulus combination; (C) vector analysis of the motion of the middle point corresponding to the percept. For further description see text.

above, defined in the framework of central perspective. Equal motion directions and velocities for translatory motion are therefore not only Euclidean parallel motions with the same velocity. (This latter description is valid only for projections from fronto-parallel motion.) The term, "equal," also includes all motions (1) that follow tracks that converge to a common point (a point at infinity) on the picture plane, and (2) whose velocities are mutually proportional relative to this point. This will be further developed in connection with the examples. See also, especially, Johansson (1971).

The perceptual motion analysis forms the basic mechanism in visual motion and space perception. It has the consequence that the ever-changing stimulus pattern on the retina is analyzed to maximal rigidity in coherent structures. The finding that it follows in a mechanical way the rules of central perspective means that we automatically obtain perceived size constancy as well as form constancy from projections of rigid objects in motion. However, what is critical for perceiving rigidity is not rigidity in distal objects, but rather, the occurrence of equal motion components in the proximal stimulus. This has been demonstrated in experiments using artificially generated changing light distributions. These experiments have also made clear how this process is fully automatic and independent of cognitive control (see Johansson, 1964, 1971).

Clearly, the vector model for motion perception represents a proximal approach to the stimulus-percept problem and is in this way contrasted with the traditional distally anchored object constancy model.

Experimental Examples on Visual Vector Analysis

Example 1

This example, illustrated in Fig. 4, shows how a translatory motion in one element in the proximal pattern is perceptually split up into two translatory components in accordance with the principles of perceptual vector analysis. The stimulus pattern, consisting of three elements in translatory motion, is shown in Fig. 4A. Two elements, a and c, follow horizontal tracks, and the third, b, moves in a diagonal direction. All the three elements move to and fro in

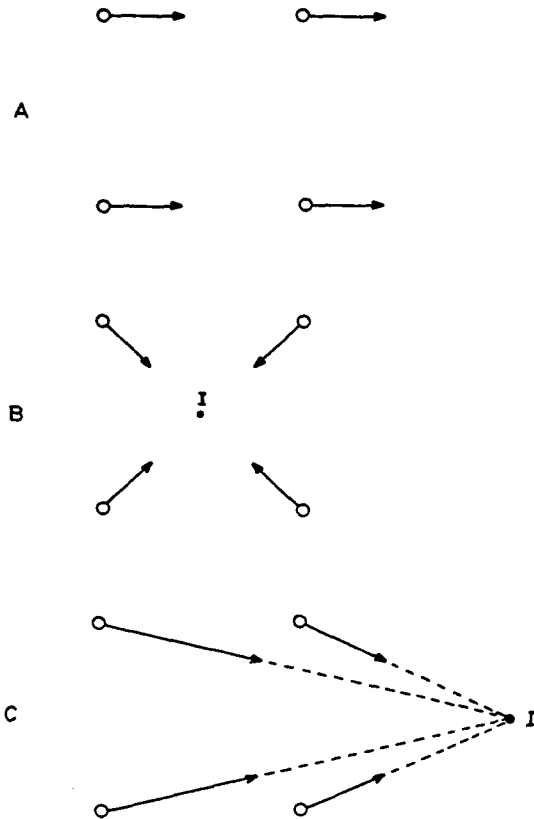


Fig. 5. Three examples of element motion which are equal in the geometry of the model. Element motions or components in element motion of this character are seen as a unit motion. The dots with common motion are seen as forming a rigid unit.

accordance with a triangular function. They have the same frequency and are in phase. Optimal frequency is .5-1.5 Hz.

An O perceives this pattern as follows: Three dots, forming a vertical line, move horizontally to and fro. During this motion, Point B moves vertically up and down along the line. This percept is given in diagram form in Fig. 4B.

In Fig. 4C, a vector analysis of the motion of Element b is given in accordance with the percept description. The diagram demonstrates the correctness of the analysis from the mathematical point of view (cf. Johansson, 1950, 1958; Hochberg, 1957).

Example 2

Figure 5 presents some proximal combinations of motions of four dots arranged as a square. The motions diagrammed are all examples of Principle 2, above. It is easily understood that Fig. 5A gives the impression of a fronto-parallel motion of a rigid square. The four dots are seen as connected by some invisible rigid structure. Figures 5B and 5C show proportionally equal motions of the elements relative to Point I. These patterns always result in perception of motions in depth of a rigid square. For a systematic analysis, see Johansson (1964).

Example 3

Figure 6 shows a more complex motion pattern. Geometrically, there exists no common element motion here. We can abstract from the element motion, however, a common component of convergence toward Point I (point at infinity). Experiments with such patterns have also made it clear that all Ss perceive a motion in depth corresponding to this equal component. The residual components are seen as a form change or sometimes as a rotation in depth (see Johansson, 1964).

Example 4

In Example 3, a rotary motion is sometimes seen. This percept is not stable, and it is, to a certain degree, inadequate from the geometrical point of view. (We have here some kind of ambiguity between polar and parallel projection.) Example 4 illustrates a case of vector analysis with an enforcing and mathematically correct perception of rotation.

Let the two points, C and P, in Fig. 7a represent the axle point and the periphery point on a wheel rolling along a straight line in a fronto-parallel direction relative to an O.

From geometry, we know that the center point in such a configuration will describe a translatory motion and that the peripheral point will describe so-called cycloids, as indicated in the figure. A cycloid (which is a special type of trochoid) can be described mathematically as composed of a circular-motion component and translatory-motion component equal to the motion of the center point of the circle.

If only Point P is visible to the O, he will always describe (or draw) it as moving in a curve, as in Fig. 7b. Adding Point C, however, immediately changes the

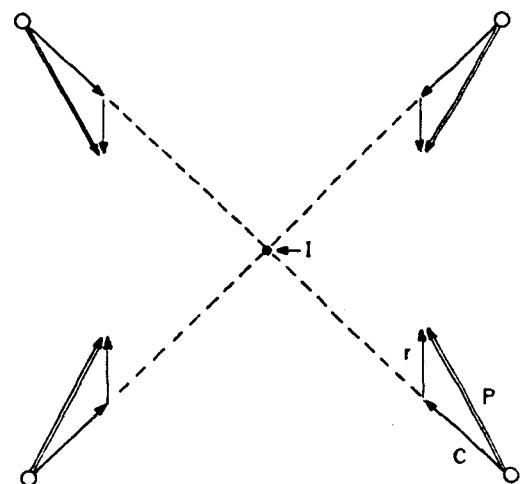


Fig. 6. Motion analysis in a nonperspective shrinking configuration represented by four elements. A common concurrent component (C) in the motion pattern is perceived as a translatory motion in depth. The residual (r) is seen as a shrinking or sometimes as a rotation in depth; p is the physical motion of the element.

perceived motion of P. The following description is typical: "I see one of the points rotating around the other one. Both move forward together. It is like a rolling wheel." This percept is diagrammed in Fig. 7c.

It is evident that this percept description represents a correct mathematical vector analysis. The motion of P is split up in one component common with the motion of the center point, C, and a residual rotary component. This perceptual vector analysis gives no option. Perception of combinations of trochoidal motions has been the subject of a recent study (Johansson, 1973).

Example 5

Our last example concerns perception of pendulum motions and therefore has a direct bearing on perception of human motion patterns. Figure 8 shows three combinations of motions in two elements, perceptually representing the end points of a rod in pendulum motion (cf. Principle 1 above). Experiments have shown that change of angle, but at a constant distance between the elements, results in perception of a fronto-parallel pendulum motion (Fig. 8A), and that change of both angle and distance produce perception of a pendulum motion in depth (Fig. 8B).

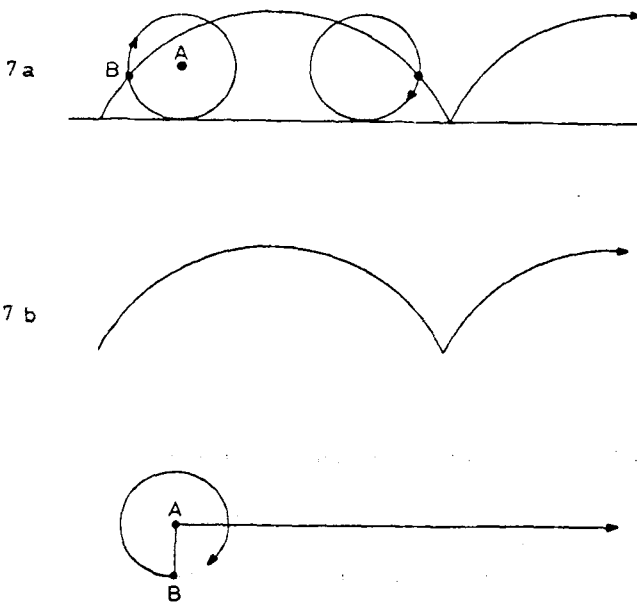


Fig. 7. (a) A point on the periphery of a circle which rolls without slip along a line is generating a cycloidal curve. The center point, A, generates a straight line. All points between A and B describe prolate cycloids. (b) The percept from exposure of Point B alone; the point is seen as moving in a cycloidal track. (c) Percept from exposure of both center point and periphery point; Point A is seen to rotate about Point B while B is moving in a linear track. The two points are seen as rigidly connected and, therefore, AB is perceived as, for example, a spoke in a rolling wheel. The perceived motion of (a) represents a correct analysis of its motion in a circular and a translatory component.

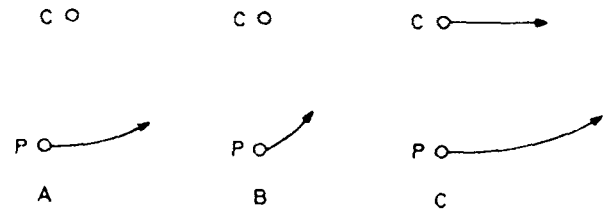


Fig. 8. Three pendulum motion combinations of two stimuli. (A) This combination is perceived as a fronto-parallel pendulum motion of Point P about C (constant distance between the elements, but change of angle). (B) This is always reported as a pendulum motion in plane in a certain angle (30-60 deg) to the fronto-parallel plane (change of distance and angle). (C) is finally seen as the same pendulum motion as in (B), but the axis is now moving during the cycle.

In Fig. 8C, we finally meet again a case of perceptual vector analysis. This motion pattern is built up from the pattern in Fig. 8B plus an equal component of translation added to both elements. This common component is effectively separated perceptually and seen as a reference frame for the pendulum motion.

ANALYSIS OF THE WALKING PATTERN

The experimental examples, 1-5, given above have one characteristic in common: in the proximal pattern, a component can be mathematically subtracted that is common for all elements in the configuration. We have also stated that such a perceptual separation of common motion components from the rest of the motion pattern has always been found in experiments critical in this respect. This perceptual separation of a common component is a typical example of perceptual vector analysis.

When the common component was subtracted in the examples given, the residual motion formed a translatory motion, a rotary motion, or a pendulum motion. These motions were also perceived in accordance with the mathematical analysis. The main conclusion exemplified with these examples is that the common motion is just a reference frame for the deviating component. Changes in this component do not change the perceived primary motion: the rotation, pendulum motion, etc. (In this sense, also, the static background forms a common motion state: the null motion.)

The motion patterns of the element combination forming the demonstrations of biological motion as described above can be analyzed in the same way by subtracting common motion components. In this analysis, the semitranslatory motion of the hip and shoulder elements (= the trunk) will be found in the motions of all elements. The motions of the knees and the elbows are rigid pendulum motions relative to this moving reference frame. Thus, the motions of these elements ought to be perceived in exactly the same way as the corresponding elements in Examples 4 and 5 above (cf. Figs. 7 and 8).

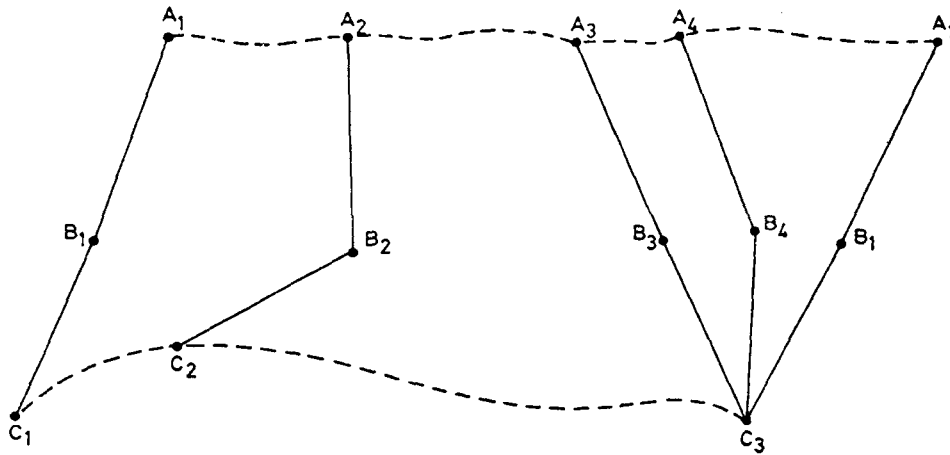


Fig. 9 (A).

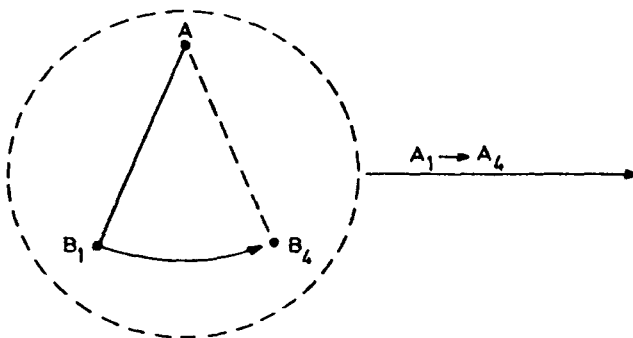


Fig. 9 (B).

Fig. 9. Graphical analysis of leg motion during walking in accordance with the model of perceptual motion analysis. (A) The motions relative to the static background of the hip (A), the knee (B), and the ankle (C) during one step. (B) The common semitranslatory component in A and B and the residual pendulum motion of B relative to the common component. (C) The common component in B and C and the residual pendulum motion in C.

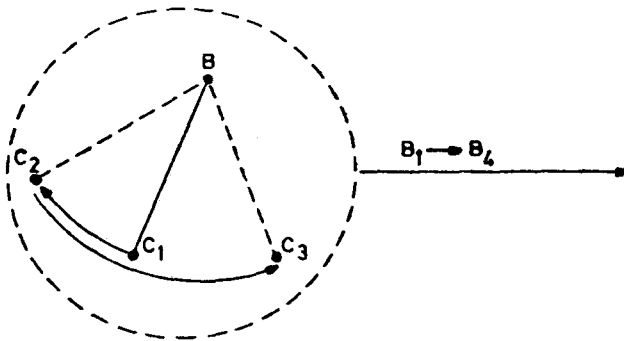


Fig. 9 (C).

The principle, however, can also be applied to the ankle and wrist elements. We are in this case concerned with the next step in a hierarchy of reference frames. The ankle describes a pendulum motion relative to the knee motion, which, in its turn, is a pendulum motion relative to the hip, while the hip describes a translatory motion relative to the ground. This is diagrammed in Fig. 9.

The motion of the hip (and shoulder) forms a common component in this analysis, inherent in all the other element motions relative to the ground. This fact affords the possibility of testing the validity of the motion analysis model when it is a question of biological motion. Let us raise the question: is the recognition of walking from the motion of our 10 points independent

of the course of the common component?

It is evident that, when accepting the motion analysis model as hypothetically valid for live motion perception also, we have to give an affirmative answer to this question. This fact was used for a partial test of the general applicability of the motion analysis model. Accepting the model enforces us to the following prediction: The semitranslatory component in walking (cf. Fig. 3), which is inherent in the motions of all our 10 elements, will play no decisive role for the identification of this pattern. Therefore, introducing drastic experimental changes in the common translatory component can be used as a test of the validity of the model.

It is also evident that, accepting a hypothesis that

identification of the walking patterns were due primarily to extended experience and overlearning would result in another prediction about the outcome from the experiment proposed. From this starting point, we have to assume that drastic changes of the most important component in walking will more or less destroy the identification. In this way, the experiment under discussion to a certain extent is critical also for the choice between a theory based on the experience factor and a model of the information analysis type.

The hypothesis testing discussed was carried out in the following three experiments.

Experiment 1

This experiment was designed to study the effect of subtracting a common component from the element motions in the walking pattern.

A typical diagram of the elements' motion tracks in walking patterns and their time relations is given in Fig. 3. The slightly wavy motion track of the hip (and the shoulder) in this diagram represents the maximum of equal motion in this combination of element motions.

This wavy motion can be looked upon as composed of two components: a major, horizontally directed, one and a minor, vertical, component.

In the experiment, we will subtract only this horizontal translatory component, leaving the up-and-down motion as a small common motion residual. (The reasons for this are technical.)

Method

The video technique described above was applied. About 10 steps of fronto-parallel walking over the scene were recorded. With one exception, the recording arrangements were exactly the same as in the recording of Demonstration 1, described above. The exception was that by careful panoraming, the actor's image during the recording was "locked" in a hair cross marking on the monitor screen. In this way, the horizontal translatory motion of the dot pattern was subtracted from the record.

The stimulus pattern was presented to the Ss against a background of total darkness. This was accomplished in the following way: The TV monitor was placed in a dark tunnel, 5 m in length. In the front gable of this tunnel was a window measuring 15 x 8 cm. The tunnel was placed in a normally lit room. Therefore, a visual-contrast black was seen through the window. The Ss were told to look binocularly into the tunnel.

Owing to this arrangement, the Ss could see the room and the outside of the tunnel but not any traces of the inside of the tunnel through the window or the TV screen. The pattern of bright spots was seen against a homogeneous black background. (A consequence of this was that some Ss reported seeing the walking pattern as a real, little manikin walking in the back part of the tunnel. The strange miniature size of the walker was, of course, due to the conditions with perceived tunnel size, the angular size of the pattern, and the total darkness in the tunnel.)

Procedure and subjects. Ten students from another faculty took part as Ss and were tested for spontaneous recognition of the pattern. None had seen our biological motion patterns before or knew that they would be shown such a pattern. They were given a neutral instruction to describe what they saw through the

tunnel window. After seeing a full scene (about 10 steps), they had to describe what they had perceived.

In addition to these naive Ss, a number of Ss who had seen human walking patterns before were tested.

Result

All Ss immediately reported seeing a walking person. The perceived miniature size was sometimes commented on. Some of the Ss also spontaneously described the event as a walking on some kind of a moving belt. The (invisible) ground then was experienced as moving backward.

Experiment 2

The 100% correct recognition responses in Experiment 1 made it necessary to manipulate the amount of information given to the Ss. The consecutive steps in the sequence mentioned mean a tenfold iteration of the step cycle and thus a considerable redundancy. Therefore, in Experiment 2, a tachistoscopic presentation was introduced.

Procedure and Subjects

A presentation time of 1.0 sec was chosen. This means that the S was shown a little less than half a step cycle (one "step"). This interval was chosen randomly during the walking period. Technically, the exposure time was controlled by an electronic timer switching the electron beam of the picture tube on and off. In other respects, the stimulus material and the arrangements were identical with those described under Experiment 1. The same holds true for the choice of Ss. Ten naive Ss took part.

Result

After the first exposure, all Ss, without any hesitation, reported seeing a walking human being.

Experiment 3

In Experiments 1 and 2, a common motion component was subtracted. From the point of view of vector analysis, this can be regarded as representing a simplification of the original stimulus pattern. In Experiment 3, the opposite principle was applied. Here, instead, an extra component was added to the primary motion of each element. This extra motion was a circular motion with a diameter of the circle covering about one-third of the distance between the foot and the shoulder elements.

This extra component was produced by an arrangement with a slowly rotating mirror in front of the camera lens, oriented at a 45-deg angle to the optical axis of the lens. The mirror was mounted on its axle at an angle which deviated somewhat from 90 deg. Therefore, when the mirror rotated, the image was not stable but described a circular motion. The mirror

reflected into the TV camera the scene to be recorded, and in this way the extra component was generated.

The scene recorded in this way was the same as in Demonstration 1 and in Experiment 1.

This stimulus pattern contains a rather complex common motion. It is the result of an addition to the ordinary semitranslatory common component in the walking pattern of a superposed rotary motion. Mathematically, it is an approximation to a special form of trochoid (a curtate cycloid). See also Example 4 above, where the element on the periphery of the rolling wheel describes pure cycloids.

Procedure and Subjects

A stimulus sequence of three to four steps was shown to 10 Ss under the same experimental conditions as in Experiment 1. Furthermore, the same sequence was shown a little more informally to a group of Os, who had seen Demonstration 1.

Result

All Ss immediately reported seeing a walking man. The walking pattern seems to be spontaneously seen in the same way as in Demonstration 1 and in Experiment 1. All Ss also said that the walker moved in a highly strange, "wavy" way.

Conclusions from Experiments 1-3

The experiments were designed for limited hypothesis testing. The results are highly conclusive. The general result is that subtracting or adding common components to the element motions do not have a disturbing effect on the identification of the walking pattern.

This result means that, in the essential respects covered by the hypothesis underlying the experiments, the vector analysis type of model has also proved valid for perception of complex motion patterns representing biological motion.

DISCUSSION

The demonstrations described have in a conclusive way made clear that 5-10 elements in adequate combinations of proximal motion give the visual system highly efficient information about human motion. Five points are enough for identification of human legs in motion.

From the theoretical reasons given above, a more or less positive outcome was expected from the experimentation with such point patterns. The overwhelmingly positive result reported here, however, was a surprise. In short, the results mean that the proximal motion patterns presented have been found to carry all the essential information needed for immediate visual identification of such human motions.

In the formulation of the problem, the question was raised about the role of experience in the identification of biological motion patterns. It seems to me that the

experimental answer is rather decisive. We can, with a high degree of confidence, conclude that it was not previous learning of motion patterns which determined the perception of walking in our Experiments 1-3. Instead, we have found that it seems to be a highly mechanical, automatic type of visual data treatment that is most important. Mathematically lawful spatio-temporal relations in the proximal stimulus pattern (complex or impoverished) determine the perceptual response. Strongly in favor of such a position is also the unexpected result from Experiment 2, namely the finding that an unfamiliar dot pattern becomes perceptually organized in less than 1 sec.

At the same time, we ought to remember that what the experiments have shown is the validity of some main principles determining the *possibility* of recognizing a biologically meaningful motion type in spite of more or less drastic changes in the original proximal pattern. It seems very probable that the vividness in the perception and the 100% correct responses which have been so typical for identifying the patterns as human motions can be derived from prior perceptual learning.

The model for stimulus-percept analysis which has been applied here is in a programmatic way anchored in a geometric-kinetic analysis of proximal stimulus. This anchorage in the proximal stimulus is a consequence of the insight that all the information available is given in the continuously changing light distribution entering the pupil.

Biological motion represents motion combinations which are far more complex than the mechanical motions from the studies of which the model originally was developed. So far, however, this model of proximal motion analysis has proved valid for biological motion too.

The model describes visual motion perception as being built up from structuring the changing proximal pattern in hierarchies of equal and deviating motion components. Our experimental tests of the validity of the model for biological motion have studied only the effects of manipulation of equal components of the first degree, that is, motion components common for all motion elements. Thus, the tests, while perfectly positive for the applicability of the model, must be regarded as more or less partial. At the same time, however, these experiments may be regarded as a highly informative first step in a more conclusive series of experimental checks aiming at verification/falsification (or modification) of the model. The rationale for this is easy to give. When the model has been found to predict adequately the basic effects of motion analysis, we have good reasons for expecting it to also cover (perhaps after some modification) some remaining, more limited aspects of visual information from the proximal stimulus.

REFERENCES

Börjesson, E., & von Hofsten, C. Spatial determinants of depth

- perception in two-dot motion patterns. *Perception & Psychophysics*, 1972, 11, 263-268.
- Börjesson, E., & von Hofsten, C. Visual perception of motion in depth: Application of a vector model to three-dot motion patterns. *Perception & Psychophysics*, 1973, 13, 169-179.
- Cohen, R. *Problems in motion perception*. Uppsala: Almqvist & Wiksell, 1964.
- Duncker, K. Ueber induzierte Bewegung. *Psychologische Forschung*, 1929, 12.
- Gibson, J. J. *The senses considered as perceptual systems*. New York: Houghton & Mifflin, 1966.
- Green, B., Jr. Figure coherence in the kinetic depth effect. *Journal of Experimental Psychology*, 1961, 62, 272-282.
- Hochberg, J. E. Effects of the gestalt revolution: The Cornell Symposium on Perception. *Psychological Review*, 1957, 69, 2.
- Johansson, G. *Configurations in event perception*. Uppsala: Almqvist & Wiksell, 1950.
- Johansson, G. Rigidity, stability and motion in perceptual space. *Acta Psychologica*, 1958, 14, 359-370.
- Johansson, G. Perception of motion and changing form. *Scandinavian Journal of Psychology*, 1964, 5, 181-208.
- Johansson, G. Geschehenswahrnehmung. *Handbuch der Psychologie*, 1966, Band 1, 745-775.
- Johansson, G. Motion perception. A model for visual motion and space perception from changing proximal stimulation. Report No. 98, Department of Psychology, University of Uppsala, 1973. (Also *Studia Psychologica*, Bratislava, in press.)
- Johansson, G., & Jansson, G. J. Perceived rotary motion from changes in a straight line. *Perception & Psychophysics*, 1968, 4, 165-170.
- Marmolin, H., & Ulfberg, S. Motion perception and form change. Unpublished report, Uppsala University, 1967.
- Michotte, A. *The perception of causality*. (Translation) London: Methuen, 1963.
- Rubin, E. Visuelle wahrgenommenen wirkliche Bewegungen. *Zeitschrift für Psychologie*, 1927, 1, 103.
- Wertheimer, M. *Untersuchungen zur Lehre von der Gestalt*. *Psychologische Forschung*, 1923, 4, 301-350.

NOTE

1. Produced by James B. Maas, Cornell University, and distributed by Houghton Mifflin Co., Boston, Massachusetts. It is also possible to receive, from the present writer, at cost, a film copy of some demonstrations described here.

(Received for publication February 28, 1972;
revision accepted February 28, 1973.)