Recognising the style of spatially exaggerated tennis serves

Frank E Pollick, Cali Fidopiastis*, Vic Braden#

Department of Psychology, University of Glasgow, 58 Hillhead Street, Glasgow G12 8QH, UK; e-mail: frank@psy.gla.ac.uk; *ATR Human Information Processing Research Laboratories, 2-2 Hirari-dai, Seika-cho, Soraku-gun, Kyoto 619-0288, Japan; # Vic Braden Tennis College, Mission Viejo, CA, USA

Received 22 March 2000, in revised form 28 September 2000

Abstract. A technique for the construction of exaggerated human movements was developed and its effectiveness tested for the case of categorising tennis serves as flat, slice, or topspin. The technique involves treating movements as points in a high-dimensional space and uses average movements as the basis for constructing exaggerated movements. Exaggerated movements of a particular style are defined as those points in the space of movements which lie on a line originating at the style average and in the direction defined by the difference between the style average and the grand average. In order to visualise the movements, computer animation techniques were employed to transform the three-dimensional coordinates of the movement into the motion of a solid-body figure. These solid-body models were used in perceptual experiments to assess the effectiveness of the exaggeration technique. After an initial training session on the exemplars from the original library, subjects viewed the synthetic tennis-serve motions and in two separate sessions either made three-alternative, categorisation judgments after viewing a single serve or rated dissimilarity after viewing a pair of serves. Results from both accuracy in the categorisation task and structure of a multidimensional scaling solution of the matrix of dissimilarities indicated that, as distance from the grand average increased, the service motion became more distinct and more accurately identified.

1 Introduction

One way to demonstrate a complicated movement to a group of novices is to exaggerate the movement. That is to say, inasmuch as it is biomechanically possible, to make the movement more distinctive by moving through postures that are extreme versions of the target motion. While common experience tells us that such exaggerations are useful in training specific aspects of a movement, it is not clear how to produce movement exaggerations automatically, or how effective such exaggerations might be in enhancing the recognition of a particular style of movement. In this paper we introduce a technique for the production of spatial exaggerations of human movement and examine whether recognition is enhanced for these exaggerated movements.

The topic how to create new movements that are more easily recognised as a particular style of movement involves several interrelated issues. One is how well we can recognise different styles of human movement. Another is how to represent a human movement appropriately and how to use this representation as a basis for the creation of new movements. In section 2 we review the literature on the recognition of movement style and discuss the possible information upon which style discrimination can be based. The different styles considered include such factors as personal identity, gender, and affective content. Next, in section 3, we consider the theoretical problem of movement representation in terms of proximal and distal specification of the movement stimulus. Finally, we introduce a technique for exaggeration (section 4) and present an experiment designed to test its effectiveness.
2 Recognition of movement style

Recognition of identity from point-light displays of gait has been studied by Cutting and Kozlowski (1977). They showed that, for a group of six friends, performance at recognising a friend from his or her gait, although significantly different from a chance level of 16.7% correct, was at only 38% correct (results ranged from 20% to 50% for individual subjects). Experiments involving the recognition of identity from low-quality videos (Burton et al 1999) have provided similar levels of accuracy. In their study they investigated the contributions of facial and body-movement information to identification. Results showed that, when presented with low-quality videos of fifteen different individuals, recognition rates were at approximately 25% with the face concealed, and rose to approximately 80% when the face was visible and the body or gait were concealed (chance was 6.7%). These results suggest that, although we seem capable of recognising identity from gait, we are not, on average, particularly good at it.

Research on identification of gender from point-light presentation of gait (Kozlowski and Cutting 1977) indicated that, averaged across all gaits, recognition rates were above chance at 63%, but that for individual gaits performance varied from 32% (statistically significant as misclassification) to 81%. It was further shown that gender of the walker could be related to biomechanical invariants such as size ratios of hips and shoulders and the centre of moment (Barclay et al 1978; Cutting et al 1978). These invariants could then be used in the production of synthetic gaits (Cutting 1978). In these synthetic gaits it was possible to vary the maleness of the gait. However, even in the maximally exaggerated cases, the percentage correct approached only 80%. These studies by Cutting and colleagues explored the perception of point-light walkers from a side view only. Research on the effect of viewpoint by Mather and Murdoch (1994), who used a different walker algorithm, showed that for a frontal view gender recognition was 79% accurate. They also provided evidence to support the claim that the discrimination of gender from gait is based on motion rather than structural information.

In other studies of gait, the relevant variables for distinguishing between different styles of gait such as running and walking have been examined (Hoenkamp 1978; Todd 1983). These studies both worked by finding a model which reduced the full complexity of the kinematic patterns into a moderately small set of parameters that could be examined. The results of Hoenkamp (1978) indicated that the most important parameter is the ratio between the time that the lower leg takes to swing forwards and its corresponding return motion. Todd (1983) showed that a walking gait could be transformed into running by adding or subtracting a constant value to the angle of the lower leg over the entire cycle.

The recognition of emotion from point-light displays of body movements has been explored by Walk and Homan (1984) with point-light displays of mimed movements, as well as by Dittrich et al (1996) with point-light and full-light videos of dancing movements. Of the six emotions examined (surprise, fear, anger, disgust, grief or sadness, joy or happiness) it was found in both studies that anger was the most reliably identified emotion. Other differences among the identifiability of the different emotions were noted, though they fell into no particular pattern between the two studies. The overall rate of recognition for the six emotions reported by Dittrich et al (1996) was 63% for point lights, as compared with 88% for full-light video displays.

The identification of different styles of movement has been studied in the context of the perception of tennis-serve movements. This research by Goulet et al (1989) used film clips to present service movements and examined eye movements and serve categorisation (flat, slice, and topspin) for novice and expert players. The eye-movement data suggested that the eye was directed primarily in the region around the serving shoulder. The categorisation data indicated that experts were better than novices in their identification (70% versus 52% correct responses) and that flat and topspin serves were more accurately identified than slice serves (flat — 70%, topspin — 70%, and slice — 40% correct responses).
Moreover, with clips of different duration, it was shown that the experts reached peak accuracy before the racket made impact with the ball, while novices improved up to the point where the ball crossed the net.

In summary, results on the ability to correctly and reliably categorise different types of movements provide a somewhat mixed set of results. It would appear that, although identity from movement is performed above chance, it is not typically done with high levels of accuracy. However, for gender, some emotions, and different styles of movement, fairly high levels of accuracy can be obtained.

3 Representing human movements
As described by Runeson (1994) in detail, there have been two distinct approaches to the study of biological-motion perception: the proximal and distal approaches. In the proximal approach a model of human-movement production is simulated in a manner that allows precise control of the proximal stimulus, while in the distal approach a natural motion is recorded and subsequently presented as a visual stimulus. Typically, the proximal approach is additive in the sense that parameters are added to the model to control certain variables, while the distal approach is subtractive in the sense that information about colour, facial expression, etc is subtracted from the movement recording. The most well known example of the proximal approach is the synthetic walkers of Cutting (Cutting 1978). Examples of the distal approach are the point-light demonstrations of Johansson (1973, 1975) where reflective markers were attached to actors and their motions were filmed. The resulting movies had only points of light at joints visible, thus removing most cues to identity, but were still spontaneously organised into the percept of human motion. Additional examples of the distal approach include the study of perceived lifted weight from point-light displays of unconstrained whole-body movements (Runeson and Frykholm 1981, 1983; Bingham 1993) and single-joint movements (Bingham 1987).

Comparison of the two approaches shows clearly that what is being traded off is naturalness of the movement for control of the proximal stimulus. In the case where one had a perfect model with appropriate independent variables then there would be negligible difference between the two approaches and the proximal approach would be superior since it would allow parametric manipulation of crucial factors to explore their effect on perception. Unfortunately, however, it is difficult to obtain models of human movement that accurately simulate movement and also have appropriate variables that can be manipulated to explore their effect on perception. For this reason we chose a distal approach to explore the exaggeration of human movement.

The present exaggeration approach is unique in the study of human-movement perception in that it uses a distal approach, but is additive. This technique has the advantage of being based on natural movements while also allowing information to be added to the distal stimulus. The technique works by first recording three-dimensional (3-D) movement data of different styles and then perturbing these data. The purpose of this perturbation is to make the different movement styles more recognisable. The principle behind the perturbation used in the present study was to amplify differences between the individual style averages and the grand average of all movements.

4 Exaggerated movements
At the heart of the exaggeration approach is the assumption that supernormal stimuli can be created from proper manipulation of the movement kinematics. By definition, a supernormal stimulus is one which produces a stronger response than the actual

(1) In the current work we record 3-D movements and perform the exaggeration in 3-D space, thus the control is over the distal stimulus. However, it is possible to perform the exaggeration on the 2-D projection of the movement and thus to provide precise control of the proximal stimulus.
stimulus it substitutes, and several examples exist from the animal world (Tinbergen 1953). The choice of exaggerating differences in average movements as a perturbation is motivated by the success of similar techniques for enhancing facial images. In the study of face processing this technique is known as caricaturing (Brennan 1985); and it has been used to study various kinds of face processing including the recognition of identity (Rhodes et al 1987), expression (Calder et al 1996; Young et al 1997), and attractiveness (Perrett et al 1994; Rhodes 1996; Rhodes and Tremewan 1996; Perrett et al 1998). At least two possible mechanisms for the effectiveness of facial exaggerations have been proposed, including that faces may actually be stored as exaggerations, or that exaggerations provide better access to the stored representation than their unexaggerated originals (Rhodes et al 1987).

In contrast to the work on static facial images, human movements occur in both space and time, and thus, when considering the information available to exaggerate, it is necessary to consider spatial and temporal factors separately as well as spatio-temporal interactions. For example, if we either change the duration of a movement while maintaining its spatial extent, or we change the spatial extent while maintaining its duration, then there is a change in the velocity of the movement. Although alteration of either spatial or temporal properties results in a spatiotemporal interaction, they can still be considered independently. Hill and Pollick (2000) have examined the exaggeration of temporal properties in the recognition of identity from drinking movements. They broke the drinking movement down into parts and exaggerated the duration of these parts relative to the average. After training on a set of natural movements, recognition of the natural movements trained upon and of the exaggerated movements was tested. It was found that the identity of the point-light drinkers was recognised from the exaggerated movements more readily than from the movements trained upon. In the present research we examine the effectiveness of spatial exaggeration in enhancing the recognition of human-movement style.

The movement we have chosen to exaggerate in the present study is the tennis serve. Tennis serves were chosen for various reasons. First, the different service styles of flat, slice, and topspin correspond to clearly different dynamic events. Second, the final execution phase of the movement leading to impact is constrained to occur in the same duration across the different styles and thus temporal variation would be of little use in making the discrimination. Third, studies from sports psychology (Goulet et al 1989) indicate that it is possible to use advance cues of body motion to predict the style of tennis serve. These advance cues are not well understood, but appear to involve subtle differences in the relationship between joints of the upper body.

A description of the procedure for obtaining movement exaggerations is best started by thinking of a space of movements (such as tennis serves), where each movement that has been recorded is a point in this space. One can assume that the variability within a style is less than the variability between styles, and thus the space would be inhabited by clusters of points belonging to the same style. The average of all the points, the grand average (GA), would form the centre of this space and the average of each cluster would form a style average. If one wanted to create new movements that were not in the original set of movement recordings, then it would be possible to do so by dynamics to refer to properties of movements such as forces, torques, mass, and inertia. (Kinematics refers to properties such as position, velocity, and acceleration.) We can consider the different serves to be based on different force interactions between the racket and the ball. In the flat serve the ball is struck so as to not induce any spin on the ball, whereas the slice serve induces sidespin and the topspin serve induces topspin (Groppel 1992). Spin is an important factor in determining the flight and bounce of the tennis ball and thus is strategically important. A basic description of the functions of these different serves is that the flat serve produces the fastest ball velocities; the topspin serve is pushed down by the spin and is thus more likely to be in; the slice serve bounces away from the opponent.
so by using these average serves as the basis. For example, given the GA and an individual style average, one can consider what happens to points along the line in the direction from the GA towards a style average. As one moves from the GA to a style average these interpolated movements become increasingly like the style average. As one passes the style average the points on this line are no longer in the volume of the original data set (of exemplars) and are thus extrapolated movements.

Our aim has been to investigate the recognition of average, interpolated, and extrapolated movements. We predict that, for each style of movement, accuracy in recognition will increase from interpolated to style average to extrapolated movements. Additionally, we predict that the GA movement will be seen as the origin for all the movements and that, as one moves further from the GA movement in the space of all movements, it will appear more distinct from the GA movement. Positive results for these two predictions will suggest that the extrapolated movements can be considered to be exaggerated movements.

In the present study we examined the effect of exaggeration on the perception of style using two separate techniques. One was a categorisation task where subjects were presented with a single serve and asked to categorise it as flat, slice, or topspin. If the exaggeration technique was effective, then we would expect that the percentage correct would increase as the serve was extrapolated further past the style average. The other task was to present a pair of different serves and to gather responses of the dissimilarity between the two. The resulting matrix of pairwise dissimilarity ratings was then analysed with multidimensional scaling (MDS) to obtain a psychological space of the serves. If the exaggeration technique was effective, we would expect the structure of the psychological space to contain sets of points, corresponding to the different styles, radiating out from the GA for increasing levels of exaggeration.

5 Methods
5.1 Subjects
The subjects were twelve recreational tennis players from the company (Advanced Telecommunications Research, ATR) tennis club who volunteered for participation in the study. All players were currently active in recreational tennis play and were of intermediate skill level. Visual acuity was normal or corrected to normal.

5.2 Collection of tennis-serve movement data
The 3-D movement data were recorded from the service movements of a professional tennis coach at the Vic Braden Tennis College in Mission Viejo, California. The technique used to obtain the data was first to use a pair of calibrated video cameras (Beta format) to videotape the motion of a serve. Next, for every frame of both video sequences, the two-dimensional (2-D) image coordinates were obtained by hand for a set of 23 anatomical locations plus the ball. These 24 points of interest are illustrated in figure 1. The 2-D position data as a function of time were then optimally filtered with a Fourier series (Jackson 1979) to smooth out noise in the individual time series. Finally, given the calibration data of the cameras and the two sets of image coordinates, the 3-D coordinates of these 24 points were reconstructed.

Tennis-serve data were collected from a single server who was instructed to perform flat, slice, and topspin serves. The service data comprised 15 serves which consisted of 5 repetitions each of flat, slice, and topspin styles of serve.

5.3 Construction of movement exaggerations
The tennis-serve exaggerations and interpolations were constructed from the 15 recorded tennis serves. In order to build these movements it was first necessary to normalise the time taken to perform a serve. The serve was defined to start when the elbow of the racket arm began to pull back in preparation for the ball toss. The end of the serve was
defined to be the frame before impact between the ball and the racket. The minimum number of frames needed to complete a serve was 68 and the maximum was 70, and no trend was evident between the style of serve and the number of frames needed to complete the service motion. All the serves lasting less than 70 frames were resampled so that they would contain 70 frames. This provided, for each serve, a time-normalised serve vector of dimension 5040 (24 points \times 3 spatial coordinates \times 70 time frames) with which to compute averages.

The GA serve was computed as the average of the 15 time-normalised serves. In a similar fashion, the flat, slice, and topspin style averages were calculated respectively from the 5 repetitions of time-normalised flat, slice, and topspin serves. The vector differences between the style averages and the GA were computed to find the vectors which defined the exaggeration directions. These difference vectors served as a standard unit between the style averages and the GA. The length of this difference unit was not identical for all three styles, but was 1.6, 1.0, and 1.6 cm for flat, slice, and topspin serves, respectively. New serves were constructed by moving out 0.5, 1.0, 1.5, and 2.0 difference units in the direction of the difference vector. The distance of 0.5 corresponds to a service motion interpolated between the GA and a style average, the distance of 1.0 corresponds to the style average, and the distances of 1.5 and 2.0 correspond to exaggeration serves extrapolated past the style average. All these new serves were constructed by using the average and difference vectors explained above. Since these new serves were combinations of vectors which contained all the spatial components of all the anatomical locations for all the frames of the motion, they can be thought of as a new sequence composed of transformed static frames.

The total set of constructed serves consisted of 13 service motions. These included 1 GA motion, and 4 types each of flat, slice, and topspin serves. These 4 types corresponded to the 4 style types constructed from difference units of 0.5, 1.0, 1.5, and 2.0 away from the GA motion.

5.4 Computer animation of movement data
Animations of a human form performing the tennis serves were produced by standard techniques of 3-D computer graphics. This involved creating a solid-body model of the tennis player and converting the 3-D position data into a format compatible to animate
this solid-body model. The solid-body model consisted of the following individual segments: the head; neck; upper-torso; middle-torso; lower-torso; left upper arm; left lower arm and hand; right upper arm; right lower arm; right hand and racket; left upper leg, left lower leg, and left foot; right upper leg, right lower leg, and right foot; and tennis ball. To create an animation, each of these segments needed to be correctly positioned and oriented for each frame of the movement sequence. This was achieved by converting the 3-D movement data into position and orientation data of the individual segments. Examples of animation sequences for the style averages for the three different types of serves are shown in figure 2a, and the different levels of exaggeration of the topspin serve are shown in figure 2b.

Figure 2. (a) Examples of the different (flat, slice, topspin) service style averages. The beginning, middle, and end of the service motion are shown. (b) Examples of different versions of the topspin serve at the beginning, middle, and end of the service motion. Different versions of the topspin serve are created by moving out from the grand average serve in the direction of the topspin serve. This difference vector is multiplied by a factor \( a \) to obtain new service motions (see text).

5.5 Presentation of movements
The orientation of the player relative to the viewer was arranged to be similar to that which would be obtained when the opponent returning the serve looked across the net. Consistent with this viewpoint, one can see from figure 2 that a standard initial position is obtained for the server, with the shoulders angled approximately in the direction of the flight of the ball (about 80° from the baseline). Movement animations were displayed on a Silicon Graphics Octane computer with MXI graphics. The frame rate achieved by the computer system was approximately 30 Hz. Subjects were seated approximately 1 m away from the computer monitor and viewed the display under
monocular conditions. When the tennis ball was at its apex, the distance from the ball to the toes of the player was 11 cm, and the distance from the head to the toes was 7.5 cm; the width of the player was approximately 2.5 cm. These distances correspond to visual angles of 6.3, 3.7, and 1.4 deg, respectively.

6 Procedure

The experimental procedure consisted of several components. The first component was a training phase where subjects received feedback on identification of the 15 service exemplars as flat, slice, or topspin. This was followed by obtaining pairwise dissimilarity judgments and three-alternative categorisation judgments on the set of interpolated and exaggerated serves. The order of these two different tests was counterbalanced across subjects and an additional training block inserted between the two different types of tests. The details of this procedure are explained below.

6.1 Training

Each subject began the experiment with a training session where he or she viewed the service exemplars that were used to construct the exaggerated serves. After viewing a single repetition of each serve, a subject identified the serve as either flat, slice, or topspin and received feedback on the response. The purpose of these training blocks was to ensure that subjects had acquired the ability to distinguish among the three different types of serves before entering into the testing phases of the experiment.

The initial training session consisted of 3 blocks of trials with each block containing 8 repetitions of each of the 15 different service exemplars (5 flat, 5 slice, 5 topspin). In a single trial, each serve was repeated once and then a dialogue box appeared in the corner of the screen and the subject used the mouse to click on one of the three possible options (flat, slice, or topspin). After entering a response the subject was immediately presented with a small window on the computer screen that informed him or her of the correct answer and his or her response. The subject then clicked on this feedback window to proceed to the next trial.

An additional intermediate training session was inserted between the testing phases of categorisation and dissimilarity judgments. This consisted of one block of trials with 8 repetitions of each of the 15 different types of serves. The purpose of this session was to assess the level of performance between testing phases and to minimise the potential for interference between the two types of tasks.

6.2 Categorisation judgments

The categorisation task involved a single block which consisted of identifying a tennis-serve motion as either flat, slice, or topspin. A total of 130 trials were presented in each block and these 130 trials consisted of each of the 13 different types of constructed serves (4 flat, 4 slice, 4 topspin, 1 GA) being shown 10 times. The constructed flat, slice, and topspin serves have a response which can be considered veridical, but the GA does not. The purpose of including it was to determine whether there was any bias in categorising the GA movement.

In a single trial, a subject was presented with a single repetition of a service motion. Immediately following this, a dialogue box appeared in the corner of the screen and the subject used the mouse to click on one of the three possible options (flat, slice, or topspin) to categorise the serve. Entering this mouse click prompted the next display to begin. No feedback was given to subjects for these judgments. Half of the subjects performed the categorisation judgments before the dissimilarity judgments and half of the subjects performed the dissimilarity judgments before the categorisation judgments.
6.3 Dissimilarity judgments
The dissimilarity task involved 3 blocks in each of which all possible pairwise dissimilarity judgments on the set of 13 serves (4 flat, 4 slice, 4 topspin, 1 GA) were made. Subjects viewed, for each possible pair except self-comparison, one service movement and then the other member of the pair. The timing of this presentation was first to show one serve (1167 ms), then to show a blank screen for 500 ms, and then to show the other serve (1167 ms). At the completion of the display of these two motions, the subject was asked to move a slider bar presented on the computer screen to indicate the level of dissimilarity. Locations on the slider bar were scored from 0 to 100 and these numbers were also visible to the subject when judging dissimilarity. The first block was treated as practice for a subject to learn a stable rating scale and thus was eliminated from consideration. The following two blocks were used in data analysis. Subjects were given no explicit instructions about what stimulus properties should be used to judge dissimilarity.

7 Results
7.1 Training
The results of training, averaged over all subjects, are shown in figure 3. It can be seen that there was a gradual improvement in performance as the number of training blocks increased. It also appeared that the recognition rate for topspin serves was higher than that for either flat or slice serves. There also did not appear to be a large improvement in performance between the initial and intermediate training sessions (the third and fourth training blocks). This point is important because the fourth training block came between the two testing phases and the small change in training performance here suggests that the results of the two testing phases are roughly comparable.

A two-way analysis of variance (ANOVA), with the repeated-measures factors of block and type of serve, was performed on the data of proportion correct. Results of this analysis showed significant effects of block ($F_{3,33} = 17.33, p < 0.001$) and serve ($F_{2,22} = 20.84, p < 0.001$), and no significant interaction ($p > 0.1$).

7.2 Categorisation judgments
The results of the categorisation judgments, averaged over all subjects, are shown in figure 4. It can be seen from the average of flat, slice, and topspin serves that there was an improvement in performance as the distance from GA increased. However, this improvement in performance did not appear to be uniform over the three types of serves. Instead, flat serves improved gradually, topspin serves quickly rose to nearly 100% correct, and slice serves showed little change. The minimal effect of exaggeration on the categorisation of slice serves is discussed in greater detail where it can be related to performance on the dissimilarity judgments.
A two-way ANOVA, with the repeated-measures factors of level of exaggeration and type of serve, was performed on the proportion-correct data from the categorisation judgments. Results of this analysis showed significant effects of exaggeration ($F_{3,33} = 54.28, p < 0.001$), and service ($F_{2,22} = 8.25, p < 0.05$), as well as their interaction ($F_{6,66} = 4.74, p < 0.001$). A posteriori analyses (Newman–Keuls, $p < 0.05$) of the means for the flat and topspin serves were performed. Results revealed that, for both the flat and topspin serves, the distance 0.5 from the GA was different from distances 1.0, 1.5, and 2.0. Furthermore, for the flat serve, distance 2.0 was different from distances 1.0 and 1.5.

The categorisation of the GA serve as flat, slice, or topspin revealed a bias to call the GA serve a slice serve. The proportion of responses for the GA serve were: flat 0.225 (SD 0.20), slice 0.65 (SD 0.25), and topspin 0.125 (SD 0.15).

7.3 Dissimilarity judgments

The dissimilarity judgments comprised two presentations of every possible pairwise comparison (excluding self-comparison). These two judgments were averaged to form a single matrix of dissimilarity judgments which was analysed by multidimensional scaling (MDS). The particular MDS algorithm used was a nonmetric individual scaling (INDSCAL) algorithm (Kruskal and Wish 1978) which found locations for the 13 tennis serves in a psychological space that best accounted for the matrix of dissimilarity judgments. A 3-D psychological space was chosen since lower-dimensional spaces displayed substantial increases in the stress of the solution while higher-dimensional spaces displayed minimal improvements in the stress. The 3-D INDSCAL solution had a stress of 0.19 and 76% of the variance of the scaled data (disparities) was accounted for by the corresponding distances.

A plot of the derived psychological space is shown in figure 5. Visual inspection indicates that the flat and topspin serves are nearly orthogonal to one another and the slice serves appear to be nearly aligned along dimension 3. Analysis of the angles between the centroids of the flat, slice, and topspin serves and with the cardinal directions bears out this observation. The angles between the centroids were: flat and topspin—$84^\circ$, flat and slice—$66^\circ$, and slice and topspin—$71^\circ$. The angles between the centroids and cardinal directions were: topspin and dimension 1—$150^\circ$, flat and dimension 2—$148^\circ$, and slice and dimension 3—$13^\circ$. The primary implication of this configuration is that the slice serves were not as distinct from flat and topspin serves as flat and topspin serves were from each other. Since the INDSCAL algorithm obtains a unique solution, which does not allow arbitrary rotations, it is not possible to rotate the solution to a more favourable orientation. Thus an interpretation of the dimensions, if it is to be done, must be performed on the current configuration.
The simplest but approximate interpretation of the dimensions would be that topspin, flat, and slice correspond with dimensions 1 to 3 accordingly.

If the exaggeration technique were effective, then it would be predicted that the psychological space of tennis serves would have the GA at its centre and the constructed serves radiating out from this origin with greater distance from the GA for greater exaggeration. The prediction of increasing distance from the GA holds, as we can see in figure 6 where distance from the GA in psychological space is plotted versus the distance from the GA in the stimulus space. A two-way ANOVA, with the factors of distance from GA in the stimulus space and service type, revealed a significant effect of distance from GA in the stimulus space ($F_{3,6} = 22.7, p < 0.01$). A posteriori analysis (Newman–Keuls, $p < 0.05$) revealed that distance 0.5 was different from distances 1, 1.5, and 2 and that distance 1.0 was different from distance 2.0.

![Figure 5. Results of the INDSCAL algorithm applied to the matrix of dissimilarity judgments. The letters F, S, T, and GA correspond to flat, slice, topspin, and grand average serves, respectively, and the subscripts to these letters indicate the distance from the GA. The resulting 3-D psychological space is seen as viewed from different directions. The radial structure of the space is most apparent when viewing the graph of dimension 1 (topspin) versus dimension 2 (flat). It can be seen that the GA serve is central to the strands of topspin and flat serves, which branch to the left and downward in the order of increasing distance from the GA.](image)

![Figure 6. The distance from the grand average (GA) movement in the psychological space plotted versus the corresponding distance from the GA movement in the stimulus space. The units in the stimulus space are difference units where 1 corresponds to the distance between the GA and the style average.](image)
7.4 Examination of slice serves

Both the categorisation and dissimilarity judgments indicated that the exaggeration technique was effective. However, in both cases performance on the slice serves appeared qualitatively different from the flat and topspin serves. In particular, there was no clear effect of exaggeration on the categorisation of slice serves. This lack of an effect was investigated further by examining the MDS data for evidence of differences between subjects and then applying this to a reanalysis of the categorisation data.

The INDSCAL model used in the MDS analysis of the dissimilarity matrix provided not only a unique psychological space to account for this matrix, but also gave individual subject weights which indicated how much each subject weighed the importance of the individual dimensions. This 3-D weight vector provides us with an idea of whether an individual subject is equally sensitive to all the dimensions of the psychological space or whether sensitivity is biased towards one or two particular dimensions. In as much as the three dimensions corresponded to flat, slice, and topspin, comparison of the weights would indicate whether a subject was equally sensitive to all three serves. Comparison of the weights was achieved through examination of the flattened subject weights (Norusis 1994). For flattened weights, distance from the origin corresponds to the equality of the weights for the three dimensions; a subject with a small distance from the origin has roughly equal weights for the three dimensions while a subject with a large distance has nonequal weights. Distance from the origin was used to partition the set of twelve subjects into two groups, one group of four subjects with a small distance from the origin (mean 0.25, SD 0.21), and a group of eight subjects with a large distance from the origin (mean 2.87, SD 1.55).

This division of the subjects into groups based on equal and nonequal weights can be applied to the results of the categorisation task to see how the two groups of subjects performed on categorising the serves. The results of categorisation for the two groups at each of the levels of distance from the GA are shown in table 1 in the form of confusion matrices. It can be seen that, for the group of subjects with nearly equal weights, there is an apparent improvement in recognising slice serves as distance from the GA increases. However, for the subjects with nonequal weights there is an increasing tendency to misidentify slice serves as flat when the stimulus distance from the GA increases.

Table 1. Confusion matrices of both subsets of subjects for the four different distances from the grand average (GA).

<table>
<thead>
<tr>
<th>Distance from GA</th>
<th>Service presented</th>
<th>Response rates for subjects with approximately equal weights</th>
<th>Response rates for subjects with nonequal weights</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>flat</td>
<td>slice</td>
</tr>
<tr>
<td>0.5</td>
<td>flat</td>
<td>0.45</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>slice</td>
<td>0.30</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>topspin</td>
<td>0.08</td>
<td>0.18</td>
</tr>
<tr>
<td>1.0</td>
<td>flat</td>
<td>0.68</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>slice</td>
<td>0.28</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>topspin</td>
<td>0.00</td>
<td>0.03</td>
</tr>
<tr>
<td>1.5</td>
<td>flat</td>
<td>0.73</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>slice</td>
<td>0.13</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>topspin</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>2.0</td>
<td>flat</td>
<td>0.75</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>slice</td>
<td>0.18</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>topspin</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
7.5 Discussion
An effect of the exaggeration procedure was seen in both the categorisation judgments and the structure of the psychological space. For the categorisation judgments the recognition rate increased as one moved from the interpolated to the extrapolated service motions. For the structure of the psychological space it was evident that the different service types radiated out from the locale of the GA serve. Thus, it can be concluded that extrapolating movements in the direction of the style average can serve as an effective means of producing exaggerated movements.

The recognition rates obtained during training the exemplars and testing the style averages are comparable to the values obtained by Goulet et al (1989) who used film clips. The similarity between recognition rates provides evidence that the procedure of transforming video images into 3-D computer-graphics movements preserved the essential information of the movements.

Examination of the results for the slice serve indicates that the exaggeration process was not as effective for the slice serves as for flat and topspin serves. Closer examination of individual subjects suggested that only a subgroup of the subjects, those with approximately equal weights for the three dimensions of the psychological space, showed enhanced recognition of the exaggerated slice serves. While these results with the slice serve qualify the success of the exaggeration technique, they are not unexpected given previous findings of individual differences in the perception of movement style. Substantial individual differences have been reported in the identification of gender or identity from gait (Cutting and Kozlowski 1977; Kozlowski and Cutting 1977). Thus, the division of subjects based on individual subject weights is reminiscent of these results. One possible reason for difficulty with slice serves might be that the style-average slice serve was closer to the GA serve than the flat and the topspin serves. This could make the slice serve intrinsically more difficult to recognise. Moreover, since the exaggeration procedure used the distance of the style average from the GA as a basis, the exaggerated slice serves would have moved less far from the GA serve as compared with the flat and topspin serves.

8 General discussion
We investigated the effectiveness of a technique to produce exaggerated tennis-serve movements. The technique worked by first collecting a library of 3-D tennis-serve movements containing several exemplars each of flat, slice, and topspin serves. From this library, exemplars of the same style were averaged together to create average movements for each style as well as a grand average (GA) movement. The vector difference between the style average and the GA provided a direction in which to move away from the style average to extrapolate new movements. It was predicted that these extrapolated movements would appear exaggerated in the sense that they appeared more distinct and recognisable as that particular style. Results of both pairwise dissimilarity ratings and categorisation judgments among a set of interpolated and extrapolated movements indicated that the extrapolated movements could be seen as more distinct versions of the particular style.

The current results are similar to those found in the caricaturing of facial images. With faces it has been found that interpolating between faces results in facial images which are more difficult to identify, while extrapolating in the direction between a certain type and the GA results in facial images which are more distinctive (Rhodes et al 1987; Perrett et al 1994; Calder et al 1996; Rhodes 1996; Young et al 1997; Perrett et al 1998). This similarity between results in caricaturing facial images and exaggerating movement styles is perhaps not surprising since the exaggerated movements obtained for the current study are, in effect, a frame by frame application of the caricature process applied to the 3-D positions of the body. Explanations for the effectiveness of the caricature technique with faces have included that faces are...
actually stored as exaggerations or that caricatures provide better access to the stored representation than their unexaggerated originals. Whether such explanations can be applied to the exaggeration effect found here remains a question for future research.

The theory of kinematic specification of dynamics (Runeson and Frykholm 1983) states that the richness of natural-movement kinematic patterns is what is essential in specifying the underlying dynamics of the action. Thus, it can be conjectured that, if a recorded natural kinematic pattern is manipulated, then this will result in an alteration of the informative details. While it would seem likely that most perturbations to the original kinematics would result in degraded perception of the movement, the exaggeration technique tested here indicates that it is possible to manipulate the spatial properties of a movement to enhance the perception of the movement. One possible explanation for this enhancement is that, within the range of exaggeration values used in the current study, the relevant dynamic properties covary with the kinematic properties. An example of such covariation can be found in the smoothness of planar arm movements, where it was found that kinematic and dynamic measures of motion smoothness show a strong covariation (Pollick and Kourtzi 1998). Whether the exaggeration technique, or any other systematic manipulation of movement kinematics, does indeed result in a covariation of movement dynamics is a difficult question to address owing to the great complexity in modeling dynamics for even a simple multiple-degree-of-freedom movement.

Although it is interesting to ask which specific properties of the tennis-serve movement were informative of the style distinctions, the exaggeration technique is not well suited to address this question. The reason for this is that the exaggeration technique is holistic in nature; it works to alter the kinematics of every joint. The holistic nature of the technique might indeed be an essential component to its success, since evidence exists to suggest that the perception of biological motion is mediated by holistic processes (Bertenthal and Pinto 1994).

An additional aspect to this holistic nature of the exaggeration technique lies in the existence of reactive impulses (Runeson and Frykholm 1981; 1983) and subsidiary movements in adjacent joints (Bingham 1987) that are intrinsic to every movement. One can conjecture that these secondary motions, typically too small to be detected with normal movements, become detectable at higher levels of exaggeration. Thus, enhanced recognition might be mediated by the creation of multiple, consistent cues to the identity of the movement.

The exaggeration technique has some practical advantages and limitations as a general technique for the manipulation of movement kinematics. One advantage is that a detailed model of the underlying movement and its dynamics is not necessary. This is an important point, since dynamical models for multiple degrees of freedom are generally quite complex and involve many assumptions and approximations which potentially limit their applicability for psychophysical experiments. The limitations of the exaggeration technique involve the space of movements used to set up the difference vectors. One such limitation is that exaggeration is limited to the styles for which average movement data are available. For instance, if the set of movements in a movement library does not span the entire space of possible styles, then there is currently no principled way to fill in such a gap. Also, the averaging of exemplars to create an average rests upon the assumption that the variability among the exemplars is limited to the type of variability associated with repetitions of the identical movement. If, for example, there was some systematic source of variation among the exemplars (such as personal style in the case of recording from multiple individuals), then there is no guarantee that the average would be prototypical. A final note to make about the exaggeration process is that the algorithm applied in the current paper calculated average movements in a space by using a representation of the joint locations in 3-D space; however, it is equally valid to find these averages in an intrinsic body space defined by joint angles.
The present results showed that exaggeration of spatial properties could modulate the recognisability of tennis-serve style. Previous results by Hill and Pollick (2000) similarly indicate that exaggeration of temporal properties can be used to modify the recognition of identity. Given that both spatial and temporal exaggeration processes result in spatiotemporal interactions, the results raise the issue that spatiotemporal properties, such as velocity, might be the important property to consider in the recognition of movement style. Such a result is consistent with the claims of Mather and Murdoch (1994) that motion cues, rather than the structural cues suggested by Cutting and colleagues (Cutting 1978; Cutting et al 1978), are important for the discrimination of gender from gait. However, even if perceivers are sensitive to the motion cues, it is likely that the structural form is still necessary to afford the motion differences. Further studies that combine both temporal and spatial exaggerations could be used to study this issue in more detail.

We described here a technique to produce exaggerated styles of movement from a library of pre-recorded movements. The technique relied on interpolation and extrapolation of the kinematics of the movement in the space of pre-recorded movements and did not require a detailed dynamical model. The results provide information about the role of spatial information in the recognition of human movement and indicate that spatial exaggeration can be effective in producing supernormal stimuli. As such, the exaggeration technique can be applied as a useful tool for many applications in the perception of movement style (identity, affect, gender, etc) where the recognisability of a movement needs to be manipulated in a systematic way. Moreover, the spatial-exaggeration technique is relevant to the development of tools for 3-D computer graphics that are concerned with producing new movements from libraries of pre-recorded movements (Bruderlin and Williams 1995; Witkin and Popovic 1995; Amaya et al 1996; Gleicher 1997; Rose et al 1998).

Acknowledgements. We wish to thank Shigeru Mukaida, Victor Zordan, and Jean-Christophe Marze for programming the graphics displays; Robert Hintermeister for his assistance in obtaining the 3-D coordinates of the tennis serves; and Chris Atkeson, Armin Bruderlin, Harold Hill, Pascal Mamassian, and Marcia Riley for many helpful suggestions. This work was supported by Engineering and Physical Science Research Council Grant GR/M36052 to Frank Pollick.

References

Bertenthal B I, Pinto J, 1994 “Global processing of biological motions” Psychological Science 5 221 – 225
Brennan S E, 1985 “The caricature generator” Leonardo 18 170 – 178


Groppel J L, 1992 High Tech Tennis (Champaign, IL: Leisure Press)

Hill H, Pollick F E, 2000 “Exaggerating temporal differences enhances recognition of individuals from point light displays” Psychological Science 11 223 – 228

Hoenkamp E, 1978 “Perceptual cues that determine the labeling of human gait” Journal of Human Movement Studies 4 59 – 69

Jackson K M, 1979 “Fitting of mathematical functions to biomechanical data” IEEE Transactions on Biomedical Engineering 26 122 – 124


Kozlowski L T, Cutting J E, 1977 “Recognizing the sex of a walker from a dynamic point-light display” Perception & Psychophysics 21 575 – 580


Tinbergen N, 1953 The Herring Gulls’ World (London: Collins)


Witkin A, Popovic ... , 1995 “Motion warping computer graphics”, in Proceedings of SIGGRAPH 1995 pp 105 – 108


© 2001 a Pion publication printed in Great Britain