Human Movement Science 34 (2014) 63–80

Contents lists available at ScienceDirect

Human Movement Science

journal homepage: www.elsevier.com/locate/humov

Automatic recognition and scoring of olympic rhythmic gymnastic movements

M. Pino Díaz-Pereira [a], Iván Gómez-Conde [b], Merly Escalona [b], David N. Olivieri [b,*]

[a] Department of Special Teaching, Area of Physical Education and Sports, University of Vigo, Ourense 32004, Spain
[b] Department of Computer Science, University of Vigo, Ourense 32004, Spain

ARTICLE INFO

Article history:
Available online 4 February 2014

PsycINFO Classification:
4120
2221
3720

Keywords:
Human motion recognition
Sports scoring
Quality of movement
Spatio-temporal analysis
PCA

ABSTRACT

We describe a conceptually simple algorithm for assigning judgement scores to rhythmic gymnastic movements, which could improve scoring objectivity and reduce judgemental bias during competitions. Our method, implemented as a real-time computer vision software, takes a video shot or a live performance video stream as input and extracts detailed velocity field information from body movements, transforming them into specialized spatio-temporal image templates. The collection of such images over time, when projected into a velocity covariance eigenspace, trace out unique but similar trajectories for a particular gymnastic movement type. By comparing separate executions of the same atomic gymnastic routine, our method assigns a quality judgement score that is related to the distance between the respective spatio-temporal trajectories. For several standard gymnastic movements, the method accurately assigns scores that are comparable to those assigned by expert judges. We also describe our rhythmic gymnastic video shot database, which we have made freely available to the human movement research community. The database can be obtained at http://www.milegroup.net/apps/gymdb/.

1. Introduction

In artistic sports, expert judges rank performance by evaluating a constellation of pre-established parameters. The scores assigned to a particular exercise are based upon relative comparisons to a set

* Corresponding author. Tel.: +34 988387026.
E-mail addresses: pinod@uvigo.es (M.P. Díaz-Pereira), olivieri@ei.uvigo.es (D.N. Olivieri).

0167-9457/$ - see front matter
http://dx.doi.org/10.1016/j.humov.2014.01.001

of movements considered to be excellent with respect to form and execution. During a competition in such sports, the key task of a judge is to assign a quantitative score to the entire exercise that is based upon a detailed analysis and comparison of each movement to a pre-established standard. Because of the large amount of variability in assessing such criteria, the degree of objectivity that a human judge can achieve is controversial and one of the most contentious aspects of artistic sports competitions.

There are many factors that affect the quality and accuracy of the judgement process in gymnastic sports. Some examples of such variability are the lack of precision in the evaluation criteria, the speed and rapid transitory duration of gymnastic routines, the perceptual limitations of the judges, and factors related to the context, such as the order of performance or pressure from other judges. An effective method capable of reducing the inherent subjectivity within such sports would be of considerable interest (Bucar et al., 2012; Gambarelli, 2008, 2012; Looney, 2004).

We describe a new method that automatically assigns a judgement score to a single (or atomic) rhythmic gymnastic movement by comparing it to a set of stored movements, which are of the same type and have been assigned scores by expert judges. The principal contribution of our algorithm is that it efficiently determines a quantitative score by comparing the velocity flow fields of an input movement to those of standard movements, represented as trajectories in a reduced dimensional space. The key conceptual insight of our method is the realization that all the necessary information for identifying quality differences of a movement is contained in these *spatio-temporal trajectories*. By comparing these unique spatio-temporal velocity trajectory representations of human motion, obtained from a video shot or a live stream, our algorithm is analogous to the way an expert judge compares a movement against a standardized code book. The method we present has the advantage that it objectively compares quantitative details of each performance during the entire motion in an unbiased manner not easily assimilated by a human.

These spatio-temporal trajectories are derived from the velocity fields produced by a human movement through a specialized mapping procedure: (1) the difference vectors, obtained by comparing pairwise consecutive image frames of a video shot, are encoded into specialized image templates, one for each pair of frames, (2) the median velocity is factored out, and each residual velocity is compared with all the rest, represented as a covariance matrix, (3) a reduced dimensional eigenspace is found by diagonalizing this covariance matrix, and finally, (4) all the image frames in the video shot are projected into the resulting eigenspace, where the bases vectors correspond to the directions of maximum variance of the relative velocities, thereby accentuating differences between movements. The set of sequential video frames, projected into this eigenspace, trace out a curve, that we refer to as a *velocity covariance trajectory*.

The performance, or execution, of an atomic gymnastic movement traces out a *unique* velocity covariance trajectory. Similar performances of the same gymnastic movement will produce similar, but unique trajectories, since the velocity fields are similar. Our algorithm uses this fact: given two performances of the same movement type, v_1 and v_2, a deviation in the execution of a movement v_1, will result in perturbations along its trajectory with respect to that of v_2. By measuring the relative distance between the perturbed trajectory, v_1, to that of v_2, we obtain an information measure that is proportional to the relative quality difference.

We implemented our algorithm as a computer vision software application, thereby making it unnecessary for a gymnasts to wear markers that would interfere with her body movements and render the method invalid for live competitions. A block diagram of the major steps in our method are shown in Fig. 1. Briefly, an input video or live video stream of the gymnastic movement is acquired. The video stream is analyzed frame by frame with low-level computer vision algorithms to construct a spatio-temporal trajectory containing details of the human motion. This trajectory is used in a pattern recognition step in order to determine the type of gymnastic movement. Once the type of movement is known, the spatio-temporal trajectory is compared against those stored in the database pertaining to the same movement, which have previously assigned judgement scores. Time synchronized trajectories can be used to produce a relative difference metric between the set of movements, given by Δd. This difference metric is used as input to a linear model, trained from a large set of videos scored by expert judges, for obtaining an absolute score for the input movement.

Because our system analyzes atomic gymnastic movements individually, the scoring process does not take into account global considerations that judges use to assign the final scores in live rhythmic

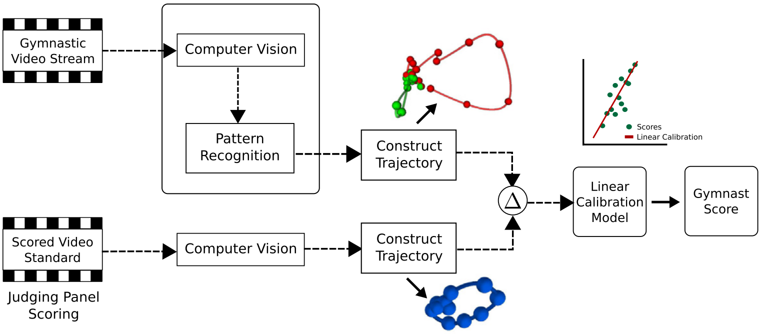

Fig. 1. Block diagram of the major components of the system. Video shot data is processed with computer vision algorithms that produce a reduced dimensional space for classification. A distance metric (Δd) and a calibration procedure is used to assign gymnastic scores that are consistent with those assigned by expert judges.

gymnastic competitions. However, our system could be used to aid the judgement process by providing a real-time *ticker-tape* like feed of atomic action scoring that would be valuable for arriving at the global evaluation. In this study, we limited our analysis to representative types of movements: *Passé-rotation* in 360°, *stag-leap* in ring, and *walkover-forward* with hand support. Representative video stills of these movements are shown in Fig. 2 and more information can be found in Supplementary Video Online Resource.1. While we demonstrate our algorithms and software system using a single camera at a fixed angle, our method is easily extended to include multiple cameras that can capture motion from different angles, thereby including more velocity field information in the feature vector.

2. Background

In general, there are two procedures for evaluating athletes in sports: by directly measuring their performance or by judging their execution (Blazquez-Sanchez, 2010). For sports that directly measure performance, quantitative techniques and instruments are used for obtaining indicators of an athlete's achievements. Those sports that require judgmental determinations, however, must consider qualitative aspects of movements that are difficult to quantify.

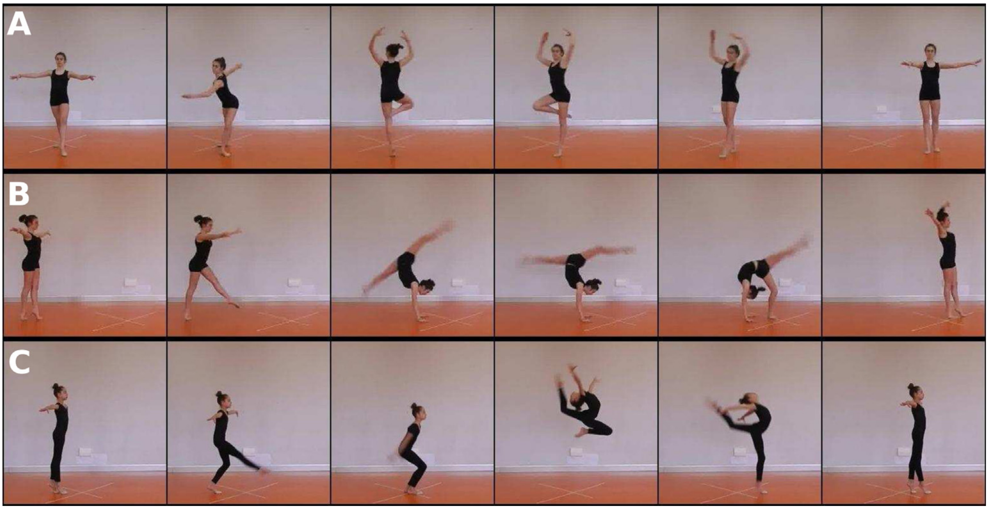

Fig. 2. The different movements. (A) the Passé Rotation in 360°. (B) The Walkover with hand support (C) the Stag Leap in ring.

There are several sports that require evaluation based upon judgments. These sports can be classified from the way that evaluations are made. For example, sports such as ski jumping, cross-country skiing, and boxing use mixed formulas, consisting of absolute and judged metrics, in order to obtain point-based scores. In gymnastic disciplines, synchronized swimming, and figure skating, the outcome depends primarily on performance judgments. For these artistic sports, the very nature and complexity of the judgment process is controversial and represents a limiting factor towards general acceptance as a mature sport.

In rhythmic gymnastics, as with other artistic sports, the judge's evaluation process considers many aspects of continuous spatial and temporal body movement. Because judges must evaluate performances in real time and from their particular vantage point, the evaluation process is not exact, but based upon an educated approximation. A large body of academic literature describes how the wide variability of evaluations is dependent upon the judge's mental capacity, personal experiences, and expertise (Heinen et al., 2013; Pizzera, 2012; Pizzera and Raab, 2012; Ste-Marie, 2003) for assessing the rapid succession of movements. The purpose of these studies is to better understand and identify those aspects of the judgment process where bias and errors can occur. The ultimate goals of such studies are to obtain better objectivity and reproducibility during the scoring process.

2.1. Judgement in rhythmic gymnastics

In rhythmic gymnastics, the judge must quantify the athlete's performance based upon two simultaneously occurring parameters: the quantity and the quality of the movements, both occurring during a short time. Bobo-Arce (2002) has described this rhythmic gymnastics judgment process as a series of sequential mental processes that must be quickly converted into points. In this description, the evaluation begins from the moment that a gymnast begins the execution of her routine, which for this sport consists of a combination of standard body movements, the use of various hand objects (e.g., hoop, ball, clubs, ribbon and rope), and choreographed with musical accompaniment. The judge must perceive each individual movement as such, classify these movements, assess their degree of difficulty and quality based upon standardized criteria, and finally provide a quantitative record of the entire performance.

The routines are evaluated by two judging panels, referred to as the *execution* and *difficulty* judges, and in both cases, the score is based upon a 0–10 point scale. The Code of Points (CoP) (FIG, 2013) is the regulatory compendium that provides the evaluation criteria by which the judgement panels will utilize to record and analyze the movements of the gymnasts during competition.

The judging panel for *execution* assesses whether the movements of a routine are executed respecting the quality standards for each type of movement. This evaluation is based on analyzing movements during the exercise and its comparison with an optimal execution model (Ste-Marie, 1999). When the movement deviates from the perfect prototype, the judge deducts a penalty from the starting score of 10 points. Penalties range from .10 to .50 according to the severity of the error. For example, if a jump routine is executed with a *lack of amplitude in its form*, the judge would deduce .10 points from the execution score, in accordance with the CoP criteria.

As the name implies, the *difficulty* judging panel is responsible for assigning a score to the technical challenge of the whole routine that corresponds to movements cataloged in the CoP. This jury calculates the final score for the entire set of exercises by summing the scores of the individual cataloged movements executed during the performance. In particular, the CoP catalog provides the assigned numeric rating to technical body movements that are classified into three fundamental types: jumps/leaps, balances, and rotations. In each of these categories, there are several actions possessing differing levels of difficulty, which are assigned a point value that can either increase or decrease the final score. Each difficulty has a value ranging between .10 and .50, depending upon the specific criteria for each group of movements.

2.2. Bias associated with judged rhythmic gymnastics

Because the evaluation process in rhythmic gymnastics is complex and consists of many different steps, subjectivity and inaccuracies are inevitable, producing scores that can differ greatly among judges. Several researchers have reported on the factors that influenced judges' evaluations of gymnastic elements (Ansorge and Scheer, 1984, 1988; Dallas et al., 2011; Plessner and Haar, 2006; Puhl, 1980; Scheer et al., 1983; Ste-Marie, 1999) in order to better understand the issues involved and mitigate errors in the future. These studies have described factors such as the judges' ability to identify errors, their experience, and their particular manner and observation speed for identifying specific aspects of movements in live performances.

Some authors have suggested that basing judgements primarily upon the CoP (FIG, 2013) as the primary assessment tool is flawed (Bobo-Arce, 2002), since it lacks precision in the manner of assigning scores and there is a large amount of information to be memorized. These problems are compounded during the observation process, because the judge must simultaneously analyze and record the movements during the short, rapid, and continuous execution of a gymnast's performance. This need to rapidly collect and evaluate information about a specific movement greatly influences a judge's evaluations (Abernethy and Russell, 1987). Another factor affecting a judge's decisions may be the order of routines in the overall exercise; gymnasts competing early in the competition may be scored different from those at the end (Ansorge and Scheer, 1988).

Still other studies have concentrated on several psychological factors that can affect a judge's ability to objectively evaluate quality. These factors include motivation, self-confidence, and cognitive style, locus of control, expectation, emotional stability or anxiety (Cabrera, 1998; Fasold et al., 2012; Feigley, 1980; Weinberg and Richardson, 1990). Other more physiological factors that contribute to judgmental error, variability and subjectivity in assigning evaluations are the judge's mental abilities, perceptivity, memory skills, and attention skills/span (Ste-Marie, 2000, 2003).

External factors can also impact a judge's ability to provide unbiased scores and introduce variability in judgemental scoring. Aigner et al. (2004), Dallas et al. (2011), Plessner and Schallies (2005), and Plessner and Haar (2006) have recently demonstrated how the physical viewpoint of the judge can affect the final evaluation. Boen et al. (2008) has tested whether open feedback (i.e., the ability to hear or see the scores of colleague judges after each performance) could influence judges, so as to conform with the opinions of their colleagues. Damisch et al. (2006) also investigated the consequences of sequential judgments on the final evaluation provided by judges. In the data from their study, the authors suggest that individual evaluations in rhythmic gymnastics may incrementally coalesce towards a similar score based upon judgements made in preceding performances, thereby creating a group effect that introduces scoring bias.

Taking all these considerations into account, it is evident that there is an inherent difficulty in scoring seemingly immeasurable aspects of a gymnastics routine, including the expressive character, technical expertise, originality of the movement, and the quality and variety of the choreography. The judge must internally use higher level performance scales (such as satisfactory, good, excellent) for arriving at a numerical score. Thus, the evaluation is an educated approximation, conditioned by the particular mental process and evaluation capacity of the judge. As a result, there is sufficient motivation (Bucar et al., 2012; Gambarelli, 2008, 2012) for developing more objective methods for judging gymnastic competitions.

2.3. Judgement software systems for sports

There are many computational techniques reported in the scientific literature for automatically analyzing human motion in videos. A broad review of such algorithms is provided by Poppe (2010). For extracting detailed kinematic movements, many tracking approaches use body markers and multiple cameras. One example is from Coffey et al. (2011), where human movement data was obtained from markers and eight cameras in order to reconstruct three-dimensional information, which was then processed with a principal component analysis (PCA). Wu et al. (2007) describe another marker-based method that obtains data from infrared light emitting diode (IREDs) to classify gait patterns with a Kernel-PCA technique. Other tracking approaches include Kalman filters (Cerveri et al., 2003) or Particle filters (Deutscher and Reid, 2005) with the goal of recovering full biomechanical human motion from video, with or without markers.

Apart from the general problem of studying human motion, several specific methods have been developed for use in sports in order to automatically evaluate the quality of the body motion during execution. An example is provided in work by Pers et al. (2002), where a technique is described for studying the overall human motion of an athlete during an entire exercise or game. In another example, Figueiredo et al. (2012) analyzed video data obtained from swimmers who wore 21 markers, located along critical anatomical landmarks used to obtain the absolute position of their limbs in three-dimensions. In gymnastics, as in other judgment based sports, there is a growing interest in automatic computational methods that can score athletes based upon their movements. In a recent example, Sevrez et al. (2009) analyzed gymnasts performing backward giant circles on the high bar under different loading conditions. In this study, they used the SIMI motion software (www.simi.com) and markers attached to the athlete's body to obtained detailed motion patterns associated with these exercises.

For methods that rely upon the use of markers to analyze human motion in sports, there are two fundamental disadvantages: (1) the markers introduce an inherent bias since they interfere with the athlete's freedom of motion and, more importantly, (2) they cannot be used during competitions. Thus, a robust and efficient markerless scoring method for analyzing an athlete's movements is of great value for live competitions, as well as for providing reliable feedback to the athlete and coaches during training.

In this paper, we describe a method and software system that represents a significant step towards the goal of producing markerless scoring. In particular, we build upon a previous algorithm that we developed (Olivieri et al., 2012) for analyzing human motion with a spatio-temporal paradigm. Here, we show that by computing the difference between the spatio-temporal velocity trajectories of an input gymnastic routine to that of a previously scored routine, we are able to assign an accurate judgement score to this input routine.

3. Method

Videos of several different representative rhythmic gymnastic movements were recorded from eight athletes. The gymnasts were from the "Pavilion Gymnastics Club" in Spain and are between the ages 12 and 14. We recorded 10 different actions that include variants of jumps, rotations, balances, and pre-acrobatic elements. The video shots were filmed with a Sony Handycam DCR-SR78 with normal room lighting. For each subject and for each action, approximately five to seven executions were recorded. In total, there are approximately $8 \times 7 \times 10 = 560$ video shots. We have made our rhythmic gymnastic video shot database freely available to the human movement research community to aid future development and provide a way to compare different analysis methods. The database may be requested at http://www.milegroup.net/apps/gymdb/, where more information about the recorded movements may be found.

The Code of Points (CoP) (FIG, 2013) provides a detailed specification of different gymnastic exercises, classified into three groups of fundamental body movements: jumps, balances and rotations. Jumps are movements in which the body projects in space through the impulse of one or two feet; it should be both well defined and sufficiently high to demonstrate the corresponding form during the entire flight. Rotations are defined by symmetric turns about the longitudinal axis. To be considered well executed, a rotation should consist of a minimum of 360° of angular displacement, and should be well defined and fixed during the execution. A balance is defined by maintaining a fixed position during the entire execution, and can be performed using different body positions. Another group of body movements are *pre-acrobatic elements*, which do not have a fundamental character, but confer to the exercise an expressive quality and originality. The *pre-acrobatic elements* include movements that consist of a rotations of the body about the transverse or anteroposterior axis (that which connects the front and back in humans) using different body supports (e.g., hands, chest, etc.) without stops in the longitudinal axis.

In our study, we focused upon a representative movement from three different groups, for the rotation, we selected a *Passé rotation* in 360° (Fig. 2A), for the jump, we used a *stag-leap in ring* (Fig. 2B), and for pre-acrobatic elements group we selected the *walkover-forward with hand support* (Fig. 2C).

3.1. Overview of algorithm and software system

Our algorithm and software system require a combination of computer vision techniques (Szeliski, 2011) and supervised machine learning (Bishop, 2006). In computer vision, algorithms are used for extracting low level information from images or videos streams, and often converting such information into high-level representations. We use standard computer vision algorithms to extract spatio-temporal velocity fields of human movement from videos. In supervised machine learning, appropriate models are selected and parameters are adjusted to maximally conform to a set of data so that future unknown trials may be predicted or classified. In the context of machine learning, the fitting of model parameters to labeled data is referred to as training, while assigning a label to an unknown trial is referred to as prediction. We use machine learning to automatically classify an input video into a specific movement class, as well as calibrate models for assigning scores to gymnastic routines that are consistent with those from expert judges.

Fig. 3 shows an overview and simplified workflow of our system, in which the training block is distinguished from prediction block. During the training phase, each frame of the video is transformed to a corresponding velocity template, referred to as the Motion Vector Flow Instance (MVFI) template (Olivieri et al., 2012), which encodes the velocity field. The high dimensional image representation of these templates is subsequently reduced, through Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) transformations. In this way, each of the image templates are represented as points along a trajectory in this low-dimensional eigenspace.

The trajectories of similar gymnastic movement types will form well defined clusters in this new eigenspace. Given the clustering of distinct types of movements, the prediction of an unknown input video shot is found by performing the same operations: projecting the resulting vector into the above space and determining the most likely gymnastic movement class, based upon nearest proximity to one of the existing clusters. These operations are shown in Fig. 3 in the prediction block.

3.2. Motion Vector Flow Instance templates

Our method uses the spatio-temporal information of a movement within a video to analyze human motion in gymnastics. This approach has its origins from work described in (Huang et al., 1999; Huang et al., 1999) for the study of human gait. As compared with full 3D kinematic studies, this technique reduces the complexity involved in human movement recognition by comparing the difference of the

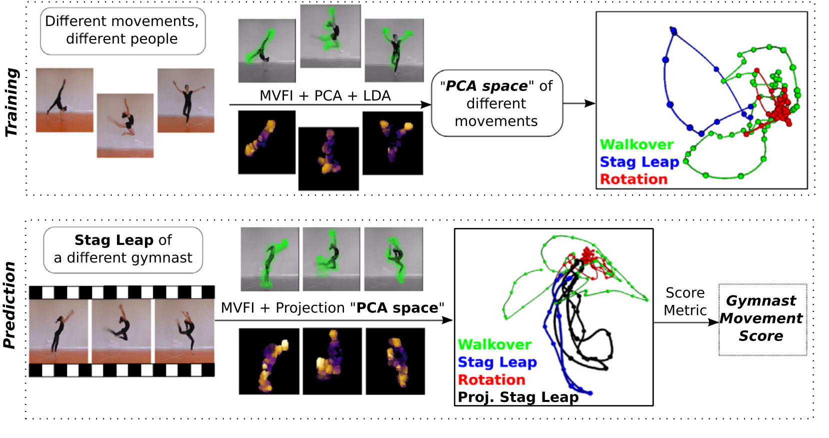

Fig. 3. Workflow of the system. The training phase (top box) utilizes a set of video shots, labeled with their corresponding movements and judgement scores. Each of these video shots is converted to MVFI template sequence, and the PCA eigenspace is found by combining all such template sequences of the entire training set. In the score prediction phase (bottom box), a new video shot is transformed into a sequence of templates, which are projected into the PCA eigenspace found in the training phase. By comparison with previously scored video shots, an absolute judgement score can be assigned to the input query.

motion fields traced out over time between the distinct types of movements. The main advantage is that this method has a concomitant reduction in computational costs that allows for real-time usage. As described, the MVFI template encodes the human motion velocity field for each frame of the video, so that different human actions can be distinguished from their corresponding velocity field. Thus, in order to recognize one action from another, or a particular person from another, the fundamental feature vector should encode essential velocity components of this motion field during the entire duration of the movement.

For each frame in a video shot, a corresponding dense velocity (or optical) flow field can be obtained, which is a quantized representation of the motion vectors on a regularly spaced two-dimensional grid. The MVFI template uses this dense optical flow map to encode the velocity vectors as rectangles, where the width and height encode the vector direction, while the pixel color (or grayscale value) encodes the magnitude. Fig. 4 illustrates the optical flow field and MVFI templates for one frame of a video shot taken from a gymnastic movement. A demonstration of the MVFI templates for the gymnastic motions from representative video shots from this study are provided in Supplementary Video Online Resource.1. A more complete description of the construction of these templates is given by Olivieri et al. (2012).

3.3. Constructing the velocity covariance trajectories from image frames

As described above, once each frame has been replaced by its MVFI template representation, a dimensionality reduction transformation is found giving rise to the *velocity covariance trajectory*. While details of spatio-temporal transformations can be found elsewhere (Huang et al., 1999; Olivieri et al., 2012), we outline the essential definitions since these trajectories are central to our method for producing gymnastic scores.

We consider a set of m gymnastic movement types, or classes, so that all such classes are given by $C = \{c_1, c_2, \ldots, c_m\}$. Within each class, there can be a set of video shots, such that in the kth class, c_k, there are a total of $\eta_k \equiv \mathrm{card}(c_k)$ (the cardinality, or size of c_k) video shots: $\left\{\mathbf{s}_i^{(c^k)}\right\}_{i=1,\ldots,\eta_k}$. Thus, the collection of all video shots to be analyzed is given by the set:

$$\mathbf{S} = \left\{\{\mathbf{s}_1^{(c_1)}, \mathbf{s}_2^{(c_1)}, \ldots, \mathbf{s}_{\eta_1}^{(c_1)}\}, \{\mathbf{s}_{\eta_1+1}^{(c_2)}, \ldots, \mathbf{s}_{\eta_1+\eta_2}^{(c_2)}\}, \mathbf{s}_{(\eta_1+\eta_2)+1}^{(c_3)}, \ldots, \mathbf{s}_{\sum_{\tau=1}^{m}\eta_\tau}^{(c_m)}\right\}$$

where shots belonging to the same class have been grouped in brackets for clarity. The total number of video shots from all movement classes is given by the cardinality of $\mathbf{S} : N_s \equiv \mathrm{card}(\mathbf{S}) = \sum_{\tau=1}^{m}\eta_\tau$.

Each video shot contains a set of image frames (consisting of $w \times h$ pixels). The jth image frame corresponding to the ith shot, $\mathbf{s}_i$, is written $\mathbf{x}_{i,j}^{k}$; the total number of such image frames is written $n(s_i) \equiv \mathrm{card}(\mathbf{s}_i)$ and can, in general, be different for each shot $\mathbf{s}_i$.

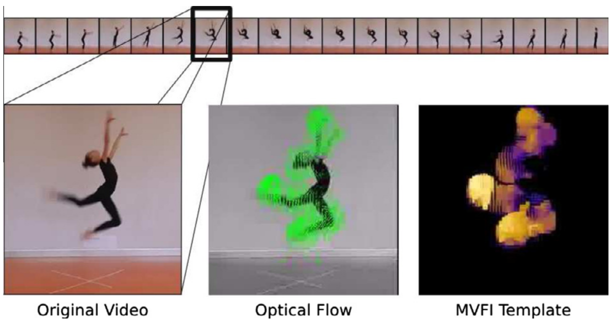

Fig. 4. Illustration of the optical flow and MVFI template for an image frame in a video shot sequence. This video shot contains a representative rhythmic gymnastic jump movement.

These definitions can be used to write the set of images for each shot. Thus, for shot $\mathbf{s}_1$, pertaining to class c_1, the associated set of image frames is written: $\mathbf{s}_1^{(c_1)} = \{\mathbf{x}_{1,1}^1, \ldots, \mathbf{x}_{1,n(s_1)}^1\}$. Generalizing to the ith shot pertaining to the kth class, the set of image frames is written:

$$\mathbf{s}_i^{(c_k)} = [\mathbf{x}_{(\eta_1+\cdots+\eta_{i-1})+1,1}^{c_k}, \ldots, \mathbf{x}_{(\eta_1+\cdots+\eta_{i-1})+1,n(s_i)}^{c_k}]$$

We form the column vector from the collection of all such video shots, $\mathbf{S}$, constructed from the corresponding sets of image frames $\mathbf{x}_{i,j}^k$, so that:

$$\begin{aligned}\mathbf{X} &= \left[\mathbf{s}_1^{(c_1)}, \ldots, \mathbf{s}_{\gamma_{j-1}}^{(c_j)}, \ldots, \mathbf{s}_F^{(c_m)}\right] \\ &= \left[\overbrace{\mathbf{x}_{1,1}^1, \ldots, \mathbf{x}_{1,n(s_1)}^1}, \ldots, \overbrace{\mathbf{x}_{1,n(s_{\gamma_{j-1}})}^j, \ldots, \mathbf{x}_{\gamma_{j-1},n(s_{\gamma_{j-1}})}^j}, \ldots, \overbrace{\mathbf{x}_{1,n(s_F)}^m, \ldots, \mathbf{x}_{F,n(s_F)}^m}\right]\end{aligned}$$

where the index $\gamma_{j-1} = \eta_1 + \cdots + \eta_{j-1} = \sum_{\tau=1}^{j-1} \eta_\tau$ is the partial sum up to the $j-1$ term, while the index $F = N_s$ (or equivalently, card($\mathbf{S}$) indicates the last shot, as defined previously.

Since we shall compare different video shots, we can factor out the common mean velocity component of the human movement from the entire vector $\mathbf{X}$, so that: $\mathbf{m}_\mathbf{x} = \frac{1}{N_T}\sum_{i=1}^{N_s}\sum_{j=1}^{n(s_i)}\mathbf{x}_{i,j}^{(c)}$, where the outer sum is over all shots, while the inner sum is over all the images within each shot; N_T is the total number of image frames from all the shots, or $N_T = \sum_{i=1}^{N_s} n(s_i)$. By factoring out the common mean velocity, we form $\bar{\mathbf{X}} = (\mathbf{X} - \mathbf{m}_\mathbf{x})$. The velocity difference between one image frame with the rest of the entire set, is found by forming the covariance matrix, $\mathbf{C}_\mathbf{x} = \frac{1}{N}\bar{\mathbf{X}}\bar{\mathbf{X}}^T$. Any non-zero element of the covariance matrix, indicates the difference of the velocity component between frame $\bar{\mathbf{x}}_{i,j}^k$ and $\bar{\mathbf{x}}_{i',j'}^{k'}$. Thus, there is minutely detailed information in this matrix about how one movement is different from another in time.

Nonetheless, as constructed, the covariance matrix $\mathbf{C}_\mathbf{x}$ is prohibitively large and sparse, making the eigenvalue problem intractable and diagonalization of the matrix impossible. An approximation, based upon a *Singular Value Decomposition* transformations (Fukunaga, 1990)), avoids a direct inversion of $\mathbf{C}_\mathbf{x}$ by diagonalizing an alternative matrix $\widetilde{\mathbf{C}_\mathbf{x}} = \frac{1}{N_T}\bar{\mathbf{X}}^T\bar{\mathbf{X}}$. In this approximation, the eigenvectors and eigenvalues, $(\tilde{\mathbf{u}}_i, \tilde{\lambda}_i)$, that diagonalize the approximate matrix $\widetilde{\mathbf{C}_\mathbf{x}}$ are related to those of $\mathbf{C}_\mathbf{x}$. These steps constitute the well known Principal Component Analysis (PCA) transformation. In practice, the span of the space is truncated to a K-dimensional eigenspace so that only the $k \leqslant K$ largest eigenvalues $|\lambda_1| \geqslant \lambda_2| \geqslant \cdots |\lambda_k|$ are retained, which is justified because $\lambda_j \approx 0$ for $j > k$. The partial set of eigenvectors span a space $\mathbf{y} = [\mathbf{y}_{1,1} \ldots \mathbf{y}_{F,n(s_F)}]$, and are calculated from the set of original images, through diagonalization of the approximate covariance matrix, $\mathbf{E}$, that defines the eigenspace:

$$\mathbf{y}_{i,j}^{(c_k)} = [\mathbf{u}_1 \ldots \mathbf{u}_k]^T \mathbf{x}_{i,j} = \mathbf{E}\mathbf{x}_{i,j}^{(c_k)}$$

Because the data is divided into different movement classes, a further operation can be applied whose effect will be to maximally discriminate separate classes. In particular, Linear Discriminant Analysis (LDA) is applied that maximizes the *between class* variance, while simultaneously minimizing the *within class* variance. The result is another linear transformation (details are described elsewhere (Olivieri et al., 2012)). In terms of the original image frames, the PCA+LDA transformation produces a new vector $\mathbf{z}_{i,j}$:

$$\mathbf{z}_{i,j} = [\mathbf{v}_1, \ldots, \mathbf{v}_{c-1}]^T \mathbf{y}_{i,j} = \mathbf{V}\mathbf{y}_{i,j} = \mathbf{V}\mathbf{E}\mathbf{x}_{i,j}$$

In this paper, we refer to the space spanned by this set of eigenvectors as the *velocity covariant eigenspace*, since it is based upon the encoding of the original velocity components of the motion. The collection of images $\mathbf{x}_{i,j}$ from a video shots $\mathbf{s}_i$, each projected into this space, constitute a *velocity covariance eigenspace trajectory*.

For classifying an unknown input video shot to a movement class, a simple distance metric for comparing the set of points along a trajectory to the different clusters in the space is with a k-nearest neighbor algorithm (KNN). The KNN method determines class membership of a trajectory by summing the distances from each of the projected points to all those points found in the separate clusters. We also used the KNN metric for comparing the distance between two trajectories in the scoring process, which we will describe in the next section.

Another more modern approach for constructing the covariant eigenspace is through the kernel-PCA (KPCA) method (Scholkopf et al., 1999). In this method, the transformation can be nonlinear and is defined by a kernel function $K(x, x')$, where typical choices are polynomial or gaussian functions. By exploiting the *kernel trick*, which obtains the inner product without the need to construct an orthogonal eigenvector basis, the KPCA is computationally efficient and can be more optimal in separating the velocity covariance trajectories than the linear PCA for discriminating different human movements.

3.4. Conceptual interpretation of the trajectories

Fig. 5 illustrates conceptually how image frames of a video shot and the corresponding MVFI templates are related to the *velocity covariance trajectory*. As can be seen, each point is an n-tuple coordinate of the MVFI velocity covariance vector at time t_n. We will recall that these n-tuple coordinate points are projections into the PCA eigenspace, described previously.

For the short segment shown in Fig. 5, the curve traces out a loop in the eigenspace over time. Since each point of the covariance trajectory consists of velocities that deviate from the mean velocity field, we expect large deviations from the mean would produce large sweeping curves in the eigenspace. This is indeed the case; the parts of the trajectory far from the origin represent movements that are furthest from the median velocity field. For the case of the jump in Fig. 5, the maximal deviation in the curve is when the body is extended and furthest from a static position. In this way, periodic movements, such as running or walking will create circular-like trajectories, similar to the dynamical phase space curves traced out by a periodic pendulum.

Because the covariance trajectories behave in a predictable manner, indeed in a manner reminiscent to a classical dynamical phase space (Thornton and Marion, 2003) (as to be expected, since in the templates we encode both the position and velocity from the optical flow field), we can postulate that similar motions, executed in a similar way should have very similar trajectories. If the motion is the same, but executed differently, then the deviation in the motion should show up in the trajectories within the covariance eigenspace.

Fig. 6 illustrates the concepts above. Here a gymnastic movements (a *walkover* with hand support) was executed by three different athletes in three separate trials; the video shots are represented as v_1, v_2, and v_3. Both v_1 and v_2 have a *high* score assigned by an expert judge, while the trial v_3, had a *low* score. Consistent with our hypothesis, the two high scoring trials have similar trajectories, while the low scoring trial v_3, has a radically different trajectory; the curve loops away from the origin just

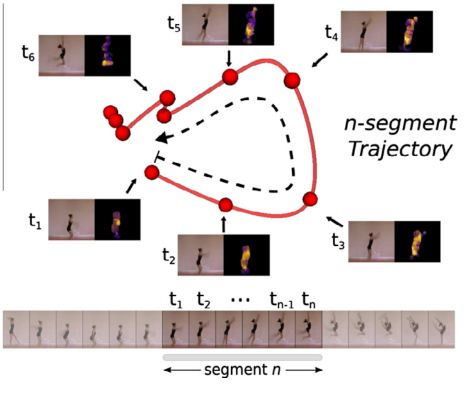

Fig. 5. Illustration of how consecutive time points along the velocity covariance trajectory are associated with frames from the video shot. Also shown for each video frame are the associated MVFI templates that encode the velocity field and are the basis for constructing the covariance trajectories.

as in the other cases, but is more closed, indicative of the lower velocity as confirmed from a visual inspection of the movement in the video shot.

Also referring to Fig. 6, if we wish to understand how different action v_1 is from v_2, we need only calculate some distance metric between them in this space. We can see, that if our metric consisted of the sum over all point to point differences, then the total distance between v_1 and v_2 is smaller than between v_1 and v_3. By using a calibration, we could transform this difference metric into one that is consistent with a judge's score. Later, we describe a linear model that carries out such a calibration procedure.

3.5. Software implementation

For this work, we developed a set of custom software tools that cover the large spectrum of tasks needed for the project; from low-level image analysis and machine learning, to the visualization of projected video frames in the orthogonal space. Our software is written in Python and uses several open-source libraries. The low-level image and video handling depends upon the OpenCV (Bradski, 2000) library. For obtaining the dense optical flow (a frame difference vector field defined on a fixed grid), we used the Farnbeck frame difference algorithm (Farnebäck, 2003) provided in OpenCV and optimized the input parameters for our problem. This is an effective frame differencing algorithm that uses parametric fitting across multiple frames to eliminate uncorrelated background movements which could negatively affect foreground object discrimination. Together with computer vision algorithms, we developed a graphical user interface, written in PyQt, with 3-dimensional visualization provided by the Visualization Toolkit (VTK) library for a representation of the MVFI trajectories. We implemented all the high level software in Python and have confirmed that it can operate in real time even on a modest laptop computer. A demonstration of the interface with some of its functionalities is provided in Supplementary Video Online Resource.3.

3.6. Judgement score for RG

In order to obtain a numerical score from a video shot of a gymnastics movement, we first binned the training data set into three categories, *low*, *medium*, or *high*, based upon the judge's score. From each of the categories, we carried out the transformation procedure described in the previous sections. Namely, we selected a representative video shot, transformed the frames to MVFI templates, and obtained the bases set from the PCA + LDA transformation, thereby projecting the videos into this space. To obtain a score that can be compared with a judge from a video shot of a gymnastic movement, we execute the same procedure in order to obtain the eigenspace trajectory for this shot and then we use a KNN distance metric to determine how close it is to each of the representative curves in the corresponding three categories.

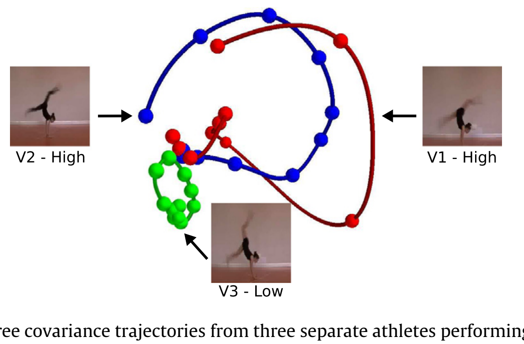

Fig. 6. Comparison between three covariance trajectories from three separate athletes performing the *walkover* movement. The movements were scored by an expert judge: V1 and V2 had *high* scores, while V3 had a low score. The curves illustrate that similar performances produce similar trajectories, while the execution of the same type of movement with a different execution (here a *low* score) produces large differences in the covariance trajectory.

More precisely, our KNN function produces a 3-tuple vector whose components are the probability of belonging to each of the intervals, defined by $\mathbf{k} = [k_{\text{low}}, k_{\text{med}}, k_{\text{hi}}]$. The vector $\mathbf{k}$ is used to calculate a single *weighted score* from a linear equation. For this, we define the central value of the score in each category are $C = \{\bar{c}_{\text{low}}, \bar{c}_{\text{med}}, \bar{c}_{\text{hi}}\}$. For an input shot $\mathbf{s}_i$, the weighted sum is given by:

$$w_i = k_i^{\text{low}}\bar{c}_{\text{low}} + k_i^{\text{med}}\bar{c}_{\text{med}} + k_i^{\text{hi}}\bar{c}_{\text{hi}}$$

However, this weighted score, w_i , is not calibrated to the scores assigned by an expert judge. For this, we establish a relationship between w_i and a *calibrated* score through a training procedure, where a linear least square fitting model is obtained for each interval ($\alpha = \{low, medium, high\}$) from several previously scored video shot data, $\{D_\alpha\}_{v=1,\dots n}$, where v represents the different data points used in each category. In this way, the linear model for the αth category is:

$$F^\alpha(p, w_i) = \min \left| D_i^{(\alpha)} - \left(\mathbf{p}_0^{(\alpha)} w_i + \mathbf{p}_1^{(\alpha)}\right)\right|^2$$

where $\widetilde{\mathbf{p}}^\alpha$ are the best coefficients found by minimizing the linear model within each category α. Using these coefficients, the calibrated, or *predicted* score consistent with a judge's evaluation is given by:

$$Y_i^*(\widetilde{\mathbf{p}}^{(\alpha)}, w_i) = \widetilde{\mathbf{p}}_0^{(\alpha)} w_i + \widetilde{\mathbf{p}}_1^{(\alpha)}$$

A summary of the procedure for obtaining the predicted judgement score from an input video shot is as follows: (1) we transform $\mathbf{s}_i$ to a velocity covariance eigenspace trajectory, (2) we calculate the KNN, to obtain the vector K, representing the probability of pertaining to each of the categories C, (3) we assign the input shot to one of the intervals based upon its value and the interval thresholds, and (4) we calculated the predicted score from the corresponding linear model Y_i^*. These steps are summarized in the functions shown in Algorithm 1.

Algorithm 1. Essential functions: create eigenspace transformation and score prediction.

function TRANSFORMATION(**Inputs: S**) ▷ **S**, set of shots
  for $i \rightarrow N_T$ **do** ▷ N_T, total number of shots
    $s_i' = \text{MVFI}(s_i)$ ▷ Create MVFI
  end for
  for $i, j = 1 \rightarrow N_T, n(s_i')$ **do** ▷ $n(s_i')$, number of frames of the shot s_i'
    $z_{i,j} = \mathbf{VE}x_{i,j}$ ▷ PCA+LDA transformation
  end for
  return Z,V,E ▷ **Z**, $z_{i,j}$ values. **V,E**, PCA and LDA matrix
end function
function SCOREPREDICTOR(**Inputs: Z,V,E,**s) ▷ s, new shot to score
  $s' = \text{MVFI}(s)$ ▷ Create MVFI
  for $j = 1 \rightarrow n(s')$ **do** ▷ $n(s')$, number of frames
    $z_j = \mathbf{VE}x_j$ ▷ Projection into the space
    $[k_j^{low}, k_j^{med}, k_j^{hi}] = KNN(z_j, \mathbf{Z})$ ▷ KNN to determine the category α
  end for
  $w = k^{low}\bar{c}_{low} + k^{med}\bar{c}_{med} + k^{hi}\bar{c}_{hi}$ ▷ The weighted sum
  $Y^* = \widetilde{p}_0^{(\alpha)} w + \widetilde{p}_1^{(\alpha)}$ ▷ Predicted score
  return Y^*
end function

4. Results

For all video shots, we applied the procedure described in the previous section for transforming each frame of the shot to a corresponding velocity covariance trajectory. Three gymnastic actions (A_1, A_2, and A_3) are shown in Fig. 7, , together with their projections into the eigenspace. All covariance eigenspaces, except for the bottom right, were constructed with two different action types, while the bottom right was constructed with three different actions. As can be seen, each action clusters into separate and distinguishable zones within the eigenspace.

To compare human movements from two separate video shots, we manually synchronized each gymnastic movement to the start of the desired frame. We also divided the video shots into small segments that capture the essential parts of the movement from each video. Fig. 8 shows an example of several synchronized segments from the different shots. We projected these video segments into the eigenspace and determined the KNN distance to each score category. Fig. 8 shows the training set projections (representative shots are from *low*, *medium*, and *high*) into the eigenspace and two projections from other shots (low and high). In particular, Fig. 8(a) shows the projection of a similar *low* scored movement (indicated by $P1$), demonstrating that its trajectory is very similar to the trained low scored shot (indicated by L). Similarly, Fig. 8(b) shows that another highly scored shot (indicated by $P2$) maps into the high score category (indicated by H) and has a very similar trajectory. Further demonstrations of the mapping of shots into the eigenspace are shown in Supplementary Video Online Resource.2.

For a newly projected video shot of a gymnastic movement (either rotations, jumps or pre-acrobatics), we used the procedure described in the previous section for extracting a calibrated judgment score. For our study, the scores that a judge could assign were in the range from 1 to 10 and were restricted to integer values.

Table 1 shows a subset of results from our data of the *walkover-forward* gymnast movement. The judge's raw global score is provided as well as categorical scores of fluency, form, and recovery that she uses to arrive at the final score. We also show the values of the KNN 3-tuple vector, which indicate the fraction of points that lie within each category. From the linear model described previously, these values are used to obtain the final calibrated score. As can be seen, our predicted scores are similar to the judge's global scores, and only one routine from the 10 videos included in this particular study, was erroneously classified by our system. The overall recognition performance for the *walkover-forward* movement is shown in the Fig. 9, which also shows the fitting procedure to each category (*low*, *medium*, *high*) (Fig. 9, upper plots), and a direct comparison (Fig. 9, lower plot) between the judge's scores and those of our system.

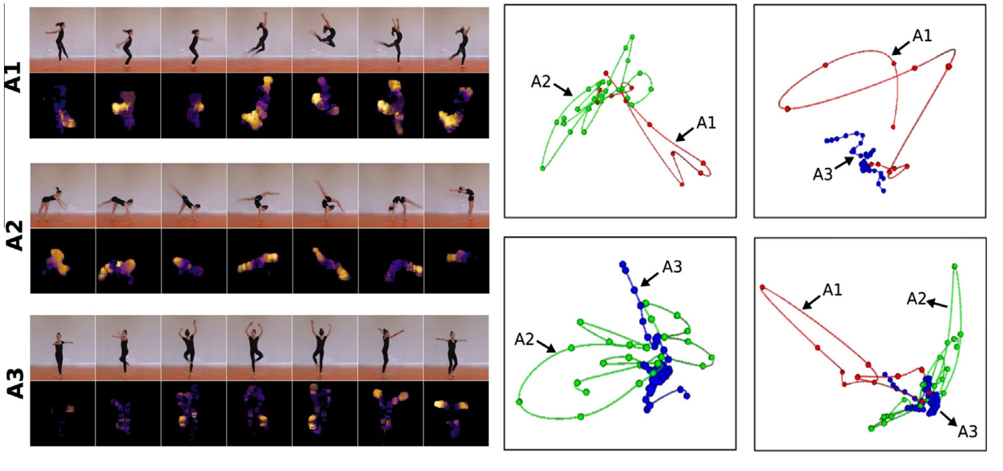

Fig. 7. Projections of the different actions into the space. Actions sequences (left) are shown with their corresponding MVFI template sequences. The projection of the shot sequences into the eigenspace (right) are shown in different configurations: either with two actions, or three actions.

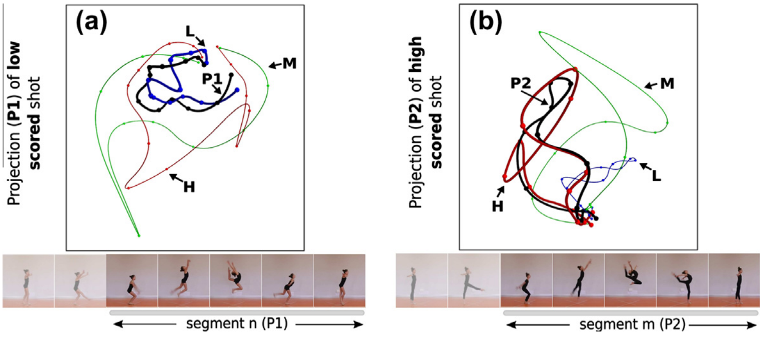

Fig. 8. Demonstration of trial trajectories: (a) a low scored shot *P*1 and (b) a high scored shot *P*2, which are projected into the eigenspace and map into their respective score categories (either low (*L*), medium (*M*), and high (*H*)).

4.1. Ground truth comparisons

We validated the accuracy of our algorithm by using a standard 10-fold cross validation. As a simplification, we defined two quantities that measure the true and false positive recognition rates for binary classification. To consider how our system performs while eliminating bias, we analyzed a hyperplane slice of the data where we divide it into two parts: those that are considered to be either medium/high score and those that are low scores. We repeated this procedure for another division of the data. An analysis for handling multi-class data, possibly based upon a confusion matrix or ROC analysis would be more complete, however we believe that this method is sufficient to understand the general recognition performance of our system.

We define the following standard performance metrics for binary classification. The quantity TP_{expert} is the number of gymnast shots that are classified by the expert judge as medium/high. TN_{expert} is the number of shots scored by the expert judge as low. TP_{system} is the number of shots classified by our system in the medium/high category and is coincident with the judge's classification, while TN_{system} is the number of shots classified by our system as low and is also consistent with the judge. On the other hand, the FP is the number of shots that are considered medium/high but are scored by the judge as low, while the FN is the number of shots classified as low, but are scored by the judge as medium/high.

Given these definitions, we can define the sensitivity R and specificity S. The first parameter used to measure accuracy is the sensitivity R, which is the ratio of true positives recognized by our algorithm $R = TP_{system}/(TP_{system} + FN)$. The other parameter we use to measure the accuracy of our algorithm is

Table 1
Results from executions of the *walkover-forward* exercise, showing a comparison between the judges score and those of our system.

	Video	Judge evaluation				System evaluation			
		Global	Specific			KNN values			Calification
			Fluency	Form	Recovery	Hi	Med	Low	
Hi	v1	9	7	7	7	1	0	0	**Hi** (8.9)
	v3	7,5	7	6	7	.64	.30	.06	**Hi** (7.6)
	v6	8	7	5	6	.42	.36	.22	**Hi** (7.7)
Med	v2	6	6	5	6	0.42	0.52	0.06	**Med** (6.9)
	v4	5	5	4	4	0	1	0	**Med** (5.7)
	v5	6	6	5	4	.62	.35	.03	*Hi* (7.6)
	v7	7	6	5	6	.32	.5	.18	**Med** (6.2)
	v8	6	6	4	4	.38	.54	.08	**Med** (6.7)
Low	v9	3	2	1	1	.05	.20	.75	**Low** (3.4)
	v10	2	1	1	0	0	0	1	**Low** (2.4)

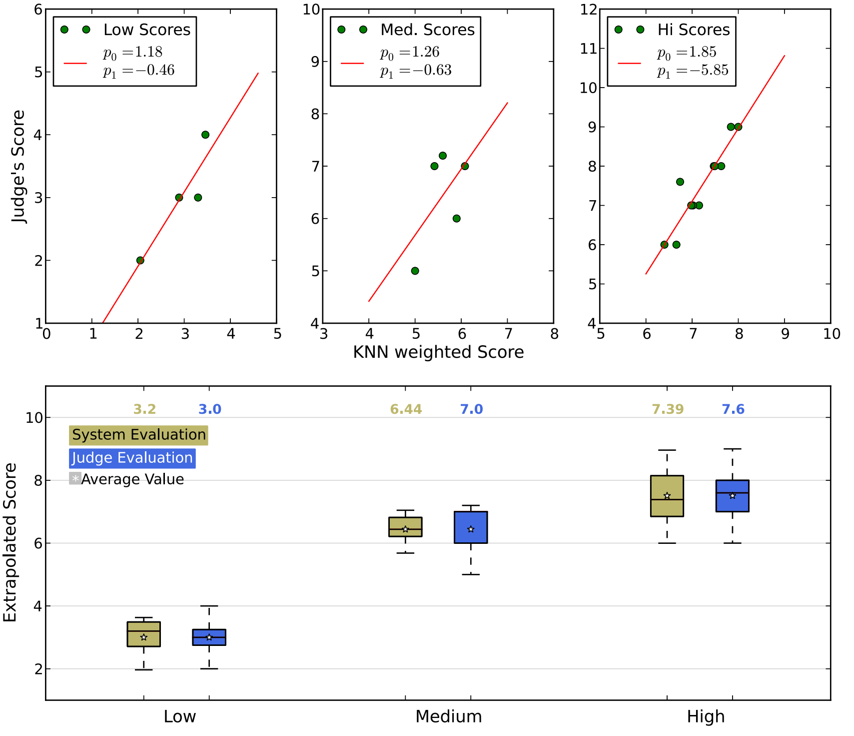

Fig. 9. Results of the score prediction for *walkover-forward* movement: (upper) fits of the KNN weighted score to the judge's score categories (low, medium, and high) used to obtain a model for determining a calibrated score for an input movement, (lower) statistical comparison of the judge's scores to our calibrated scores.

Table 2
Evaluation comparisons of our system to those of expert judges. The scores from expert judges provide a ground truth comparison for testing the scores obtained with our algorithm through the quantification of: the false positives (FP), the false negatives (FN), the specificity (S), and the sensitivity (R).

Action	TP_{expert}	TN_{expert}	TP_{system}	TN_{system}	FP	FN	S	R
Jump	18	12	17	10	1	2	.91	.89
Rotation	18	12	16	7	5	2	.58	.89
Walkover	20	10	18	9	1	2	.9	.9

the specificity S, defined as the ratio of true negatives normalized by the total number of actions identified by our algorithm: $S = TN_{system}/(TN_{system} + FP)$. These parameters are used to quantify the results of our algorithm against the evaluations of an expert judge. In Table 2, we show the performance of our system with respect to these quantities. As can be seen, the specificity of our method $\approx 85\%$, while the sensitivity is $\approx 90\%$ for all actions considered.

5. Discussion

Different authors have described the importance of introducing automated software systems for objectively evaluating aspects of gymnastic movements, or biomechanical issues such as technical quality. Such a system would free judges from the arduous task of recording movements and concentrate their attention on artistic and aesthetic aspects of the exercise, such as mastery, expressiveness, or the originality of the choreographed gymnastic performance. The present judgement process is rife with potential sources of errors, due to difficulty of memorizing the CoP catalog (Ste-Marie et al., 2001; Ste-Marie, 2003). Here, we have described a software tool that is capable of recognizing and scoring the quality of gymnastic movements.

Our results represent a significant contribution towards a fully automated system for aiding judges in their evaluations. With a machine learning procedure, our software is able to recognize different gymnastic movements and automatically assign performance quality scores to these movements. This study is a proof-of principle that should open the door to the possibility of being able to score all the actions in the CoP catalog, where there are a total of 150 movements.

In this study, we demonstrated our software system from three different movement categories (i.e., jumps, rotations and pre-acrobatics). These movements have a spatio-temporal structure that are quite different. Unlike the other two movements, for example, the jump has a flight phase where the athlete's body is not in contact with the ground. The Passé rotation in 360° has as a common characteristic body rotation around the longitudinal axis, where the toe is used as support. In the pre-acrobatic movements, the body uses the transverse axes of rotation or anteroposterior executed on different supports such as hands, back or chest, which is completely different from the other types of motions.

The differences between these actions are discernible as different velocity covariance trajectories. We showed that similar movements will have similar trajectories. Moreover, we showed that quality differences can be inferred from the relative distance between trajectories, and used to automatically assign scores. This is possible because the trajectories contain both the spatial and temporal structure that uniquely encode each movement. For example, during the Passé rotation, the spatial position of the gymnast does not vary, whereas the movement in the *walkover-forward* the motion starts and ends at a different position (approximately one meter in the forward direction). This is an important characteristic of the movement, and appears in the trajectory. Also the temporal development and fluidity of an exercise is different in the three types of movements that we compared. For example, the duration of the jump is quite different from the reverse rotation or pre-acrobatic movements. Or in another example, in the rotation and the walkover, the temporal development of the action is smooth and continuous, while the physical demands of the jump requires a rapid torque and variations in the intensity of the motion. All these temporal factors will be manifest in the distance between points along the trajectories of the projected video shots.

Our method produces predicted scores for atomic gymnastic movements with an accuracy greater than 85%. Scoring each individual gymnastic movement is analogous to the task of the *execution judge* in a real situation. However, as described, the final score in a competition is based on the sum of such scores from isolated movements. Our results allow us to distinguish movements of the same type, but with different levels of quality in execution assigning a score based upon three quality levels (i.e., low, medium and high), showing a high level of consistency and agreement with the assessment issued by the expert judges. For the system to be used in live competitions, an automatic alignment of isolated movements from a continuous video stream must be carried out in order to obtain the calibrated predicted score. We are presently developing this algorithm and it shall be included in future versions of our software. Also, our method can be extended to multiple cameras, in order to incorporate different shot angles that should improve the accuracy of the scores.

As a final point, the importance of our method for determining a quantitative difference between movements transcends the specific example of scoring rhythmic gymnastic performances. We have showed that the covariant velocity trajectories contain sufficient detailed information about the spatio-temporal aspects of a particular motion so as to extract quality differences between movements. Such trajectories are like fingerprints, and we can ask how similar one trajectory is to another in a very efficient manner by calculating a distance metric between them. This is very much a machine learning approach, where we allow the computer provide a score that is most convenient to a machine, as opposed to copying the exact way that a human would score.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.humov.2014.01.001.

References

Abernethy, B., & Russell, D. (1987). The relationship between expertise and visual search strategy in a racquet sport. *Human Movement Science, 6*, 283–319.

Aigner, W., Tomitsch, M., Stroe, M., & Rzepa, R. (2004). Be a judge! – Wearable wireless motion sensors for audience participation. In *Conference on Human Factors in Computing Systems, CHI EA 2004, Vienna* (pp. 1617–1621). doi:http://dx.doi.org/10.1145/985921.986173.

Ansorge, C., & Scheer, J. (1984). Biasing factors affecting judging of gymnastics. *Technique, 4*, 19–21.

Ansorge, C., & Scheer, J. (1988). International bias detected in judging gymnastic competition at the 1984 olympic games. *Research Quarterly for Exercise and Sport, 59*, 103–107.

Bishop, C. (2006). *Pattern recognition and machine learning (information science and statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc.

Blazquez-Sanchez, D. (2010). Evaluar en Educacion Fisica. *INDE*.

Bobo-Arce, M. (2002). El juicio deportivo en gimnasia ritmica : una propuesta de evaluacion basada en indicadores de rendimiento. Ph.D. thesis, University of Coruña.

Boen, F., van Hoye, K., Auweele, Y., Feys, J., & Smits, T. (2008). Open feedback in gymnastic judging causes conformity bias based on informational influencing. *Journal of Sports Sciences, 26*, 621–628. http://dx.doi.org/10.1080/02640410701670393.

Bradski, G. (2000). The openCV library. *Dr. Dobb's Journal of Software Tools*.

Bucar, M., Cuk, I., Pajek, J., Karacsony, I., & Leskosek, B. (2012). Reliability and validity of judging in women's artistic gymnastics at university games 2009. *European Journal of Sport Science, 12*, 207–215. http://dx.doi.org/10.1080/17461391.2010.551416.

Cabrera, D. (1998). El perfil de las jueces de gimnasia ritmica. *Apunts: Educación Física y Deportes, 53*, 103–104.

Cerveri, P., Pedotti, A., & Ferrigno, G. (2003). Robust recovery of human motion from video using kalman filters and virtual humans. *Human Movement Science, 22*, 377–404. http://dx.doi.org/10.1016/S0167-9457(03)00004-6.

Coffey, N., Harrison, A., Donoghue, O., & Hayes, K. (2011). Common functional principal components analysis: A new approach to analyzing human movement data. *Human Movement Science, 30*, 1144–1166. http://dx.doi.org/10.1016/j.humov.2010.11.005.

Dallas, G., Mavidis, A., & Chairopoulou (2011). Influence of angle of view on judges' evaluations of inverted cross in men's ring. *Perceptual and Motor Skills, 112*, 109–121.

Damisch, L., Mussweiler, T., & Plessner, H. (2006). Olympic medals as fruits of comparison? Assimilation and contrast in sequential performance judgments. *Journal of Experimental Psychology: Applied, 12*, 166–178. http://dx.doi.org/10.1037/1076-898X.12.3.166.

Deutscher, J., & Reid, I. (2005). Articulated body motion capture by stochastic search. *International Journal of Computer Vision, 61*(2), 185–205.

Farnebäck, G. (2003). Two-frame motion estimation based on polynomial expansion. In *Proceedings of the 13th Scandinavian Conf. on Image analysis* (pp. 363–370).

Fasold, F., Memmert, D., & Unkelbach, C. (2012). Extreme judgments depend on the expectation of following judgments: A calibration analysis. *Psychology of Sport and Exercise, 13*, 197–200. http://dx.doi.org/10.1016/j.psychsport.2011.11.004.

Feigley, D. (1980). Psychological factors in gymnastics judging: Personal feelings about locus of control and susceptibility to external influences. *NAWGT National Newsletter, 2*, 2–4.

FIG (2013). Code of points for rhytmic gymnastics. Fédération Interanational de Gymnastique. http://www.fig-gymnastics.com/.

Figueiredo, P., Seifert, L., Vilas-Boas, J., & Fernandes, R. (2012). Individual profiles of spatio-temporal coordination in high intensity swimming. *Human Movement Science, 31*, 1200–1212. http://dx.doi.org/10.1016/j.humov.2012.01.006.

Fukunaga, K. (1990). *Introduction to statistical pattern recognition* (2nd ed.). San Diego, CA, USA: Academic Press.

Gambarelli, G. (2008). The coherent majority average for juries' evaluation processes. *Journal of Sports Sciences, 26*, 1091–1095. http://dx.doi.org/10.1080/02640410801930135.

Gambarelli, G., Iaquinta, G., & Piazza, M. (2012). Anti-collusion indices and averages for the evaluation of performances and judges. *Journal of Sports Sciences, 30*, 411–417. http://dx.doi.org/10.1080/02640414.2011.651153.

Heinen, T., Mandry, S., Vinken, P., & Nicolaus, M. (2013). Motor skill acquisition influences learners' visual perception in gymnastics. *Science of Gymnastics Journal, 5*, 19–28.

Huang, P., Harris, C., & Nixon, M. (1999). Human gait recognition in canonical space using temporal templates. In *IEE Proceedings of Vision, Image and Signal Processing* (pp. 93–100). Vol. 146 (2).

Huang, P., Harris, C., & Nixon, M. (1999). Recognising humans by gait via parametric canonical space. *Artificial Intelligence in Engineering, 13*, 359–366.

Looney, M. (2004). Evaluating judge performance in sport. *Journal of Applied Measurement, 5*, 31–47.

Olivieri, D., Gómez-Conde, I., & Vila-Sobrino, X. (2012). Eigenspace-based fall detection and activity recognition from motion templates and machine learning. *Expert Systems with Applications, 39*, 5935–5945. http://dx.doi.org/10.1016/j.eswa.2011.11.109.

Pers, J., Bon, M., Kovacic, S., Sibila, M., & Dezman, B. (2002). Observation and analysis of large-scale human motion. *Human Movement Science, 21*, 295–311. http://dx.doi.org/10.1016/S0167-9457(02)00096-9.

Pizzera, A. (2012). Gymnastic judges benefit from their own motor experience as gymnasts. *Research Quarterly for Exercise and Sport, 83*, 603–607. http://dx.doi.org/10.5641/027013612804582597.

Pizzera, A., & Raab, M. (2012). Perceptual judgments of sports officials are influenced by their motor and visual experience. *Journal of Applied Sport Psychology, 24*, 59–72. http://dx.doi.org/10.1080/10413200.2011.608412.

Plessner, H., & Haar, T. (2006). Sports performance judgments from a social cognitive perspective. *Psychology of Sport and Exercise, 7*, 555–575. http://dx.doi.org/10.1016/j.psychsport.2006.03.007.

Plessner, H., & Schallies, E. (2005). Judging the cross on rings: A matter of achieving shape constancy. *Applied Cognitive Psychology, 19*, 1145–1156. http://dx.doi.org/10.1002/acp.1136.

Poppe, R. (2010). A survey on vision-based human action recognition. *Image & Vision Computing, 28*, 976–990.

Puhl, J. (1980). Use of video replay in judging gymnastics vaults. *Perceptual and Motor Skills, 51*, 51–54.

Scheer, J., Ansorge, C., & Howard, J. (1983). Judging bias induced by viewing contrived videotapes: function of selected psycological variables. *Journal of Sports Psychology, 5*, 427–437.
Scholkopf, B., Smola, A., & Muller, K. (1999). Kernel principal component analysis. In *Advances in Kernel methods – Support vector learning* (pp. 327–352). MIT Press.
Sevrez, V., Berton, E., Rao, G., & Bootsma, R. (2009). Regulation of pendulum length as a control mechanism in performing the backward giant circle in gymnastics. *Human Movement Science, 28*, 250–262. http://dx.doi.org/10.1016/j.humov.2008.10.002.
Ste-Marie, D. (1999). Expert-novice differences in gymnastic judging: An information-processing perspective. *Applied Cognitive Psychology, 13*, 269–281.
Ste-Marie, D. (2000). Expertise in women's gymnastic judging: An observational approach. *Perceptual and Motor Skills, 90*, 543–546.
Ste-Marie, D. (2003). Expertise in sport judges and referees: Circumventing information-processing limitations. In J. L. Starkes & A. Ericsson (Eds.) *Expert performance in sport: Advances in research on sport expertise* (pp. 169–190). Champaign IL, Human Kinetics.
Ste-Marie, D., Valiquette, S., & Taylor, G. (2001). Memory-influenced biases in gymnastic judging occur across different prior processing conditions. *Research Quarterly for Exercise and Sport, 72*, 420–426.
Szeliski, R. (2011). *Computer vision: Algorithms and applications*. *Texts in computer science*. New York, London: Springer.
Thornton, S., & Marion, J. (2003). *Classical dynamics of particles and systems* (5th ed.). Brooks Cole.
Weinberg, R., & Richardson, P. (1990). *Psychology of officiating*. Champaign: Human Kinetics Publishers.
Wu, J., Wang, J., & Liu, L. (2007). Feature extraction via KPCA for classification of gait patterns. *Human Movement Science, 26*, 393–411. http://dx.doi.org/10.1016/j.humov.2007.01.015.