Image Analysis and Development of Graphical User Interface for Pole Vault Action

Rajeev Kanth, Jukka-Pekka Skön, and Kari Lehtomäki
Savonia University of Applied Sciences, Kuopio, Finland
Email: {rajeev.kanth, jukka-pekkka.skon, kari.lehtomaki}@savonia.fi

Paavo Nevalainen, Petra Virjonen, Mikko-Jussi Laakso, and Jukka Heikkonen
University of Turku, Turku, Finland
Email: {ptneva, pekavir, milaak, jukhei}@utu.fi

Abstract—In recent years, motion estimation analysis has become one of the vital research areas in sport and has attracted the interest of many researchers toward events such as swimming, pole vaulting, and hurdling. In this paper, we present a novel method for determining the step length, speed, and the feet-contact-time on the running track of a pole vault athlete using a mono-camera arrangement. The step length and step frequency are essential descriptors of the approach run in pole vaulting. The approach along a linear trajectory is familiar to many throwing and jumping events. The measurement setting and image processing, as well as the step registration stages such as the block matching and optimal flow algorithm are presented and compared to alternative methods. The validation of the step size and step frequency accuracy is provided, using manually digitized step sizes as the baseline. The proposed methodology is efficient and straightforward, providing immediate feedback to the athlete and coaches. We were also successful in building a basic Graphical User Interface (GUI) to illustrate pole-vaulting actions during a performance. This research could be used as an initial step for developing a fully interactive platform that is capable of yielding supportive instructions to the athletes and the coaches on a real-time basis for self-assessment and further improvement.

Index Terms—motion estimation, gait recognition, graphical user interface, computer vision, pole vaulter action

I. INTRODUCTION

Each athlete dreams of winning an Olympic medal. With this end, the gradual enhancement in computer vision and artificial intelligence technologies are becoming supportive by providing useful information for athletes to improve their efforts in their sports and realize this dream. In the future, computer vision technology will certainly assist athletes and coaches to deliver better performance. The development of techniques for the identification of individuals and especially athletes has driven the interest and curiosity of many researchers. One identifying feature of a person is their gait. In this paper, we have made a motion estimation analysis of a pole vault athlete during a performance. Pole-vaulting is one of the four major jumping events of track and field. The aim is to jump over a horizontal bar using a long flexible pole. The competitor who clears the greatest height is declared the winner. To help us in our research, a raw video of one athlete’s performance was provided by Turku Sports Academy of Finland. Various techniques from the literature are considered in this paper, and a simple and novel method is developed to extract feature points and create visualizations that may provide insight for coaches and the athletes themselves.

The existing literature reveals that several investigations have been carried out within the sector of computer vision systems for tracking and modeling the human body. Since the early 2000s, the computer vision research community has begun to investigate silhouette-based human identification from the body and gait as a biometric [1]. Gait recognition has been substantially researched and has given a new impetus toward surveillance and security motives. The interest is driven by the need for automated person identification for security-sensitive environments such as banks, parking lots, and airports. The focus has been on dynamic face recognition and recognition of body dynamics, including gait [2]. Emmanuel Ramasso et al. [3] have introduced a novel tool called the temporal credal filter with a conflict-based model change (TCF-CMC) to smooth belief functions online in a Transferable Belief Model (TBM) for human action recognition in athletics videos. Toshihiko Fukushima [4] has discussed the active bending motion of a pole vault robot to improve reachable height, explicitly analyzed the Transitional Buckling Model, and has concluded that the active actuation of the athlete in the pole-support phase plays a vital role in strengthening vaulting performance. In a similar way, El-Sallam [5] has presented a low-cost markerless system for the optimization of athlete performance in sports such as the pole vault, jumping, and javelin throw. The proposed method utilizes a number of calibrated cameras to capture a video of an athlete from different viewpoints; the body is then segmented from the background in each video frame, and silhouettes of the segmented body are then reprojected to reconstruct an estimate of the 3D body shape of the athlete.
An original investigation on a framework for high-level visual content interpretation and annotation for sports events has been carried out by Sutasinee Chimlek [6]. In this context, the article aims to generate meaningful descriptions for visual content based on aggregate information of the detected primitive objects, spatial relations, and specific relevant prior knowledge to aid the interpretation of visual content. Slobodan Ilic et al. [7] have demonstrated a physics-based non-linear plane beam model for tracking large deformations. These models not only contribute to robust and precise tracking in the presence of clutter and partial occlusions but also allow the computation of the forces that produce observed deformations. Similarly, Costas et al. [8] have presented an unsupervised, automatic human motion analysis and action recognition scheme tested on athletics videos of the pole vault, high jump, triple jump, and long jump.

The literature has illustrated that human gait recognition is becoming of particular interest for surveillance and security applications because it is a biometric that is available in low resolution and has a non-invasive nature [9]. Compared to the conventional biometric features, such as the face, iris, palm print, and fingerprint, gait has many unique advantages such as being non-contact and perceivable at a distance. However, gait features have a high intra-personal variation in shape and are also influenced by external conditions like footwear, clothing, load carrying, and the person’s mood. These issues have been well presented by George et al. [10], and the authors also realized that accidentally stubbing a toe on a dumbbell on the floor could markedly change one’s gait for a few days.

Detection and analysis of human motion are also a key component in video-based human-computer interaction systems. For the description and recognition of human movement, two approaches have been developed: first, texture-based methods that extract dynamic features for human movement description, and second, a framework that considers ballistic dynamics for human movement segmentation and recognition [11]. The above-explained research techniques are in good alignment with our work [12] where we have demonstrated that video assessment is one of the dominant modes of performance analysis, particularly in sports such as swimming. Likewise, D. Gouwanda et al. [13] note that human movement analysis has been widely used in sports training to identify problems in technique in various sports such as golf and running. In this context, several approaches such as cameras and magnetic tracking systems have been implemented. Video cameras have been used primarily with either active or passive reflective markers placed on the body to capture the subject’s movement. In addition, magnetic tracking systems have been utilized where sensors measure the strengths of magnetic pulses generated by a transmitter. The intensity of the pulses is related to the distance between the sensor and the transmitter. Usually, magnetic systems are used in animation rather than in sports and gait analysis because possible distortions are likely to be received that may cause unreliable results.

In most of the literature, we found that the majority of approaches explored the use of camera work with 2D images, since video material has been adequately accessible for analysis. In addition, a real surveillance scenario operates in varied environments and, because of this feature, many gait analysis approaches are oriented toward view-invariant gait recognition. 2D gait analysis uses a single camera positioned perpendicular to the subject’s direction of walking, while the 3D approaches with two cameras are considered better options for the needs of viewpoint invariant recognition. To relieve viewpoint dependence, the inherent periodicity of gait can be used to achieve a technique, which can deal with the effects of the viewpoint. Understanding the actual physical movement of a pole-vaulter is an important issue, not only for extracting a pattern but it is also useful in determining the exact step size and the center of mass of an athlete. In order to improve the performance of an athlete, the analysis of the physical state of the pole and biomechanics of the athlete in the pre and peri (before and during) pole-support phase is crucial. One study [14] shows that the athlete’s total energy when crossing the bar can exceed 120% of the initial energy at take-off from the ground.

The step length and frequency analysis is only a sub-problem of full gait analysis, and it can be implemented by a simple mono-camera arrangement with immediate feedback to the athlete. Sports analytics has a wide variety of measurement technologies, and even video-based analysis includes a sector for? expensive universal motion analysis systems and for more specific highly sophisticated solutions. The simplicity of the approach run provides an opportunity for the economical, easy-to-install solution presented here.

Based on the above findings, we present a fast-to-implement approach to capturing several pole-vaulting actions and to building an introductory Graphical User Interface (GUI). This paper is structured as follows. The following two sections express our views on the Experimental Setup and Camera Calibration, and State-of-the-art Algorithms respectively. Section IV describes the research methods for extracting feature data points using bitwise XOR operation and utilization of image difference functions on the video frames. The results achieved and their validation is also described in this section. Development of a software tool for motion estimation analysis is discussed in section V. At the end of the paper, conclusions and possible directions for future research are presented.

II. EXPERIMENTAL SETUP AND CALIBRATION

A mono camera arrangement was made for setting up the experimental environment. A single front camera at the transverse orientation was used to capture a video of the athlete’s performance as shown in Fig. 1. A similar projection plane approach was used. The centerline of the running track and the vertical direction define the projection plane, which was covered by 27 calibration images. This produced a calibration data set \( D = \{(p_i,x_i)\}_{i=1,n} \) where, \( p_i \) are the pixels and \( x_i \) is the
horizontal distance along the track. The calibration data set size \(n=27 \times 4 \times 17 = 1836\). The calibration squares were 11.5 cm wide.

An interpolation function \(f: p \rightarrow x\) from the pixel positions to the horizontal distance was created based on \(D\). A leave-one-out test as in [15] and [16] was used to estimate the standard deviation accuracy of the horizontal length \(\sigma_x = 1.6\text{mm}\). The athlete moves approx. 8 pixels (5.94 cm) between frames. This equals a speed of approximately \(v = 5.9\text{m/s}\), which can be measured by the relative 1.1% std. accuracy. The video frame rate is 99 sec⁻¹.

III. STATE-OF-THE-ART ALGORITHMS

In this section, we briefly discuss the approach utilized in the analysis of the raw video clips provided by Turku Sports Academy, based on block matching and optical flow algorithm for motion estimation. Several algorithms presented in [17]-[20] have also been studied. Conceptually, we have considered some stages/phases for analyzing the step sizes of the athlete theoretically during a performance. The approach shown in Fig. 2 demonstrates the techniques for determining the step size and ground contact time of the athlete.

A. Video to Image-Frames Conversion

An interactive MATLAB code was executed for raw video clips as the input and minimum square error leading to the displacement between the frames as the output. The technique utilized in the code will assist in determining the step size of a pole-vault athlete with less computational complexity. 492 image frames were extracted from the given video clip of 15 seconds and saved into separate image files and then a further rebuilding of the movie by recalling the frames from the disk was carried out. Out of the 492 image frames, the five frames in which the athlete’s foot touches the ground were selected for the block matching and optical flow algorithm.

B. Spotting the Feet of an Athlete

In determining the step size of the athlete, the second stage/phase is to recognize the foot spot on the running track. An efficient computational approach has been adopted. The location of the foot is traced out by first cropping the image size from 378×637 (height × width) pixels to 60×542 pixels and secondly by examining the color intensities in subsequent frames. The image shown in Fig. 1 is of 378×637 pixel size while the image in Fig. 2 is of 60×542 pixel size. The pixel intensities (0 to 255) are represented by an unsigned integer format (uint8) of one-byte length. In Fig. 2, the shoe color of the athlete is white, which means that intensities (255) have been tracked for each frame before sending them to the block matching or optical flow algorithm.
C. Block Matching and Optical Flow Algorithm

Fig. 3 shows the calibration at the projection plane. The intensity values of 90×428×3 (uint8) arrays are considered as the inputs for the block matching and optical flow functions in MATLAB. Block matching estimates the motion between two images or two video frames using a block of pixels by examining the difference in location of pixels with the required intensity (in our case, white shoes of the athlete, i.e., detecting 255 color intensity pixels) in the Kth and (K+1)th frames. The mean square error (MSE) and mean absolute difference (MAD) can be computed as per the following two equations as shown in Fig. 4 [21]:

\[ \text{MSE}(d_1,d_2) = \sum \sum \left| s(n_1,n_2,k) - s(n_1+d_1,n_2+d_2,k+1) \right|^2 / N_1 \times N_2 \tag{1} \]

\[ \text{MAD}(d_1,d_2) = \sum \sum \left| s(n_1,n_2,k) - s(n_1+d_1,n_2+d_2,k+1) \right| / N_1 \times N_2 \tag{2} \]

In equations (1) and (2), B is an \( N_1 \times N_2 \) block of pixels, \( s(x,y,k) \) denotes a pixel location at (x, y) in frame k, and the block estimates the displacement of the center pixel of the block as \( (d_1, d_2) \).

The optical flow estimates the direction and speed of the object from one image to another or from one video frame to another using either the Horn-Schunck or the Lucas-Kanade method. To compute the optical flow between two images, the constraint equation needs to be solved [22].

IV. Proposed Research Methods and Results

The conceptual approach discussed in section III, i.e., the block matching and optical flow algorithm, is suitable for applications where the shape of the desired objects does not change. In the current pole-vault application, the form/position? Of the athlete’s feet and the other body mechanics are regularly changing. Therefore, a block matching and optical flow algorithm without any tweaking will not suit the pole-vaulting application. In a similar way, we tested another approach by applying HOG descriptors and modifying the code for tracking the athlete using a rectangle and finding the athlete’s step length, but this approach was also unable to track the body in different postures. Hence we have created a new approach for determining the silhouettes of an athlete using absolute image difference and bitwise XOR operation among the frames of the video. This approach is simple and efficient. A software package on the feature extractions using Open CV Python was developed. In addition, a software package including a GUI for feature extractions using Open CV Python and MATLAB-GUIDE was also developed.

A. Extracting Silhouettes & Location of the Athlete

First, the images were turned from RGB images into grayscale images. Each pixel only has an intensity value from 0 to 255. The background image was calculated by taking the median value of the images’ pixels throughout the video as recommended in [23]. However, based on the nature of the current video clip of a female athlete, video frame 265 was chosen as the background image. The approaches mentioned in [24]-[26] have been utilized in some other video clips of male athletes where the background variations were influential. The eight ground-touching video frames shown in Fig. 5 in the gray format were chosen because the full foot of the athletes was in contact with the running track. Difference image frames using python commands (cv2.absdiff (image, background)) and bitwise XOR operation (cv2.bitwise_xor (image1, image2)) were carried out to obtain the image shown in Fig. 5. The obtained image has a characteristic of 1920×500 pixels (width and height) having a TIFF format of 96 dpi horizontal and vertical resolution.

![Figure 4. Algorithm for step size of an athlete using MATLAB/Simulink](image-url)
B. Finding the Step Size, Speed, and Contact Time

Determining the distances between consecutive steps requires a coefficient of step length. This coefficient in our computation was considered to be 115/2, i.e., 57.5 which came from the calibration 2 mm chessboard square, equivalent to 11.5 cm. In this specific calibration mode, the length of the first two consecutive steps was found to be 2.45 cm and 2.75 cm respectively. Hence all the other step distances could be found. Step size, speed, and the ground contact time are the essential parameters that greatly affect the pole-vaulter’s performance. During this analysis, we found that the athlete’s speed distribution along the track was not uniform; hence it was important to determine the speeds during each successive step. The concept of determining the speed of the athlete is to first find the number of frames between the two steps in question. The total number of frames involved in the phenomena was 266. The time required to run 266 frames was (4 sec×434)/266 = 2.45 sec. The coefficient for determining the time required to cross step distances is (2.45/266) × difference of frames. The ground contact time is computed by the difference between the initial and final video frames that show the athlete’s feet in contact with the track. For example, the contact time in milliseconds during the 3rd step is the product of the coefficient, frame difference, and 1000.

C. Obtained Results and Validation

In this work, to date we have determined three necessary parameters, i.e., horizontal step distances, the speed of the athlete, and ground contact time as shown in Fig. 6, Fig. 7, and Fig. 8, respectively. The detected step sizes of the athlete were compared to the manually calibrated data and were very consistent. The plan is to perform result validation regarding the speed of the athlete using a rotating device whose exact speed is known. However, this research work is ongoing to build a robust and efficient platform for the precise measurement of pole-vaulting actions.
We have developed an introductory software tool (version 1.0) for determining the step length, speed, and feet-ground contact time of an athlete’s performance. The backend code has been generated using Open CV Python. The GUIDE features of MATLAB have been utilized to develop the front end of this tool, as shown in Fig 9. Both the numeric values and distinct plots for different parameters of pole-vaulting actions are embedded within the tool. Several push buttons, axes, and text forms have been inserted to visualize the actions interactively.

VI. DISCUSSION AND FUTURE DIRECTION

Our research work makes a novel and significant contribution to the image and video analysis of pole-vaulting performances. The obvious practical implications of the results are that the developed platform is not only applicable for pole-vaulting but can also be useful for other sports such as hurdles and swimming. We carried out two different approaches for image and video analysis of the performance of a pole-vault athlete. In the first approach, we performed modeling based on a block matching and optical flow algorithm. In the second approach, an analysis was carried out to determine step sizes, speed, and contact time. A number of video clips for several athletes (both male and female) were also analyzed.

There are many avenues for further future research. The first crucial future task is to reflect on the optimization of the athlete’s biomechanics phenomena during his/her performance in order to achieve the best possible height in the pole-vaulting action. Second, it would probably constitute a breakthrough in the sports video analysis domain if comprehensive investigations into an athlete’s action in all the pre and peri (before and during) pole support phases during the performance could be conducted and implemented interactively within a GUI. A fully interactive platform where an action could be visualized on a real-time basis, which supports mobile devices, would be the most desirable application in the future. A comparative analysis for several athletes on a single screen (GUI) could also be a challenging topic for future research work.

ACKNOWLEDGMENT

The authors would like to thank Turku Sports Academy for providing the video material. The venue used for the pole-vaulting performance is the Kupittaa Sports Hall, Turku. This work is supported by the AUMUK Project, University of Turku, Finland.

REFERENCES


Petra Virjonen is a project researcher at the University of Turku in the Department of Future Technologies. She works in a project for developing autonomous ships and forest machines. She also works as a research scientist in Turku University of Applied Sciences. Besides motion analysis for athletes, her research interests include ship path prediction with machine learning, and building acoustics as well as psycho-acoustics. She holds a Master’s degree in physics, and is currently a PhD student focusing on optimizing physical parameters using psycho-acoustical data. At her spare time she coaches juniors in athletics.

Mikko-Jussi Laakso is heading the Centre for Learning Analytics at University of Turku, Finland. His main research interest are automated assessment, programming education, immediate feedback, learning analytics, gamification, effect of collaboration in computer science education. The team develops ViLLE- the collaborative education tool platform. He has published more than 80 international peer-reviewed articles and he has collected more than 2M€ in R&D projects.

Jukka Heikkonen has been a professor of computer science of University of Turku, Finland, since 2009. His current research as the head of the Algorithms and Computational Intelligent (ACI) research group at the University of Turku is related to intelligent and learning systems specially including machine learning, probablistics, and information theoretical modelling issues applied in wide varying application domains. He has worked at top level research laboratories and Center of Excellences in Finland and international organizations (European Commission, Japan) and has led many international and national research projects. He has authored more than 150 scientific articles.