9th Conference of the International Sports Engineering Association (ISEA)

Using image processing for biomechanics measures in swimming

Robin P. Dubois[c], David V. Thiel[a], Daniel A. James[1,b]*

[a] *Centre for Wireless Monitoring and Applications, Griffith University, Brisbane Qld, Australia*
[b] *Centre of Excellence for Applied Sport Science Research, Queensland Academy of Sport, Brisbane Qld, Australia*
[c] *Department of Information and Electrical Engineering, Hamburg University of Applied Sciences, Hamburg, Germany*

Abstract

Underwater video analysis is common in elite level swimming for stroke correction, technique analysis and as a visual aid to support coaches and athletes. Because of the challenges of the underwater visual environment and ambulatory camera work, video can be degraded and considered non ideal, additionally the athletes movement can lead to blurred vision and bubbles from cavitation effects.

In this paper we use image analysis techniques to enhance images for improved clarity and automated detection of limb segments related to metrics of interest at the elite level. Thus we address the problems of detection, segmentation and estimation of body configuration using several image processing algorithms. These classic image-processing problems are even more complicated with video footage of poor quality as described above. Our approach is to use adaptive algorithms like the global probability for boundary detection to detect the swimmer's boundaries and deformable models arranged in a pictorial structure to recognize body and limbs plus their configuration.

Our results show that it is possible to detect the athlete and relevant body segments. Identification and estimation of the body configuration in every frame is demonstrated to be feasible. From this analysis, common metrics can also be extracted such as stroke counts, stroke rate and primary phases of an arm stroke. The goal of this work is to support the training of athletes and help them improve their technique.

Keywords: swimming, biomechanics, video analysis, segmentation, boundary detection, deformable model

* Corresponding author. Tel.: +61-73735-5036; fax: +61-73735-5384.
E-mail address: www.qsportstechnology.com

1. Introduction

Underwater video analysis is one of the most common techniques used in elite level swimming as an aid to support the coaches and athletes, though new technologies like wearable sensors are emerging [1]. Video is a frequently used instrument for biomechanical monitoring and technique analysis and can benefit from technological advances in processing to enhance it. One important reason might be that the visual sense is very important but a live observation is often not sufficient for the coach and/or the athlete to evaluate performance without a recorded video.

1.1. Swimming performance measures

Commonly used assessment techniques can be separated into three broad areas: performance, biomechanics and physiology although there is considerable overlap between them. Performance monitoring contains measurable movements of the swimmer during the monitoring period and typically times related to their movement such as splits and lap times. Biomechanical monitoring, a detailed part of performance analysis, uses direct and indirect measurements to quantify the movement of the swimmer, often to map them to theoretical models and norms. Physiological investigations mainly look at the energy systems of the athlete during training, competition and recovery.

Swimming assessment is traditionally a labour intensive process where stroke phase, stroke rates, stroke counts, and lap times are often manually recorded or extracted from video data. This manual process is dependent on high staffing levels and is generally unavailable for routine training activities or in remote areas. Beyond the basic measures, the coordination of key body segments in swimming is of growing interest for swimmers and coaches [2,3] though it is possible to obtain with sensors [4]. Understanding these movements can identify whether the action is enhancing swimming performance [5,6], or is potentially detrimental [7].

1.2. Analysing freestyle swimming

When performing technique analysis the extraction of common metrics such as stroke counts and stroke rate is but the beginning. As up to 90% of a freestyle swimmer's velocity has been attributed to the effect of arm stroke [8, 9] we want to detect the primary phases of arm stroke. Definition of arm stroke phases for freestyle swimming cover both propulsive and recovery motion. Generally, the propulsive phase can be further broken down into the following: the point at which the hand enters the water; *entry*. This is followed by small lateral movement away from the body; *outsweep*. Simultaneously, there is a downward movement; *downsweep*, to where the hand is below the elbow: *catch* is where propulsion of the armstroke is considered to begin [3, 8]. After the catch, the hand movement tends to be inwards, towards the midline of the body; *insweep*. The final portion of propulsion; upsweep, takes the hand to the point of leaving the water; *exit*. Exit of the hand from the water is typically at the waist/hip region. *Recovery* is any movement of the hand back to in front of the swimmer for the next hand entry. Because of reflections on the bottom side of the water surface the recovery phase is difficult to capture by underwater video recordings. The identification of armstroke components into stages is important for assessment.

1.3. Underwater environment

The underwater environment is challenging for video recording. To capture the swimming performance, ambulatory camera work is necessary, often this is a submerged camera mounted on a pole

and p[ropelled by human locomotion or more expensively, track mounted. The camera must be drawn underwater with a variable speed accordingly to the velocity of the swimmer. Without professional equipment and cameramen, the video can be degraded (such as motion artefact from the cameraman) and considered not ideal. Moreover the intricate environment underwater can lead to blurred vision and the swimmer's movement produces visual noise in the form of cavitation artefacts.

2. Image Analysis

In this study, the development of automated classification and enhancement techniques for analysing underwater swimming video footage is considered. With the ongoing evolution of computer technology, many detection and segmentation algorithms requiring high processing performance can be employed and computed within a reasonable time. Our approach is based on the ideas of Li et al. [10] and Mori et al. [11]. Both used two current algorithms for segmentation and boundary detection. The algorithm they use for segmentation is Normalized Cuts [12] and a second algorithms for boundary detection, are called Probability of Boundary [13] and a Global Probability of Boundary [14].

An important part of our approach is a combination of these two image-processing algorithms. They enable us to enhance videos and complete automated detections. During the segmentation, the image is cut into super pixels. The quality of the segmentation is considered optimal if one limb segment consists of one superpixel. The boundary detection algorithm is looks for object boundaries and its result is optimal if it draws a closed boundary around the whole body of the swimmer. The superpixels are needed to find body part candidates and the boundaries are used to calculate the segmentation quality of them.

2.1. Segmentation

Segmentation can be the first step in an object detection application; that is to segment images into foreground and background regions. An algorithm segments the image into a defined number of parts, which are supposed to be of different appearance, or separated by edges.

To achieve high quality segments, the Normalized Cuts [12] algorithm is used. The algorithm generates a superpixel map of the image with the number of super pixels k = 100. As is evident in figure 2, the Normalized Cuts algorithm generates super pixels of reasonable size. To achieve that it computes the optimal partition of a graph (in this case an image) into two sets (superpixel) and then recursively computes the optimal partition of the subsets until the desired number of subsets is reached. After the generation of super pixels, the amount of data is decreased and the image analysis can be done more efficiently.

2.2. Boundary Detection

The boundary detection (also referred to as contour detection), is an enhanced method of edge detection. The algorithm used in our approach is called Global Probability of Boundary [14] and was presented in 2008. It delivers the best performance [9, 14] in contour detection without post-processing like in [15]. Boundary detectors like [13, 14] detect edges using localised cues in brightness, colour and texture and then combine the results after weighting them to a boundary. The optimized detector in [14] also computes a global cue based on spectral partitioning.

The boundaries found in the image are needed to score the body parts detected by the torso and limb detectors.

2.3. Methodology

Image files are exported from existing archives of swimming video data [1] and processing routines are developed within an athlete data toolbox of Matlab [17], to enable current and future compatibility with other data sets. The process starts with the pre-processing where we convolve [10] the image with a 5x5 Gaussian Filter to reduce noise and add the binary image to the RGB image. At this point we first apply a contrast enhancement filter. This image is then used as input to each of the two algorithms mentioned above.

We considered using face detection for an estimation of the torso region, however because we use a lateral view and the head is above the water surface, this is difficult. This is a main issue because we need a robust detection of the human body in the video images. To handle this we adapt the method of [11] to find torso and limbs by using Normalized Cuts Segmentation to generate limb and torso candidates and use low-level cues to score these candidates. The low-level cues are contour, shape, shading and focus. The contour cue was found to be the most useful. It is evaluated using the result of the Global Probability of Boundary. To evaluate the shape cue, a deformable rectangle model is matched to the shape of one or more segments/superpixels. The overlap between the best matching rectangle and the segment(s) is measured. The shading cue is evaluated by constructing a shading descriptor, from a set of manually labelled limbs. The idea of a focus cue is based on the lack of higher frequency information in the background if the camera focuses on an object in the foreground. This cue is not adaptable in our particular case since the underwater video footage is often blurry and noisy at the same time.

After the finished torso and limb detection, we want to find possible body configurations through enforcing global anthropometric and kinematic constraints to minimise the number of possible configurations. This is desirable because of the number of possible configurations since 5-7 candidate half-limbs and about 50 candidate torsos lead to 2-3 million possible partial configurations. Mori et. al. [11] discard the physically impossible limbs regarding their size, position or rotation to each other and to each torso candidate and pay attention to the symmetry of clothing with the evaluation of a colour histogram.

With the obtained sets of partial body configurations we try to complete the configurations by extending all detected partial limbs into possible directions and evaluating the nearby superpixels. After doing this with all possible partial configurations, the resultant full configurations are scored. The score of a full body configuration is the linear combination of the limb and torso scores. This step creates a shortlist with the scores of full body configurations.

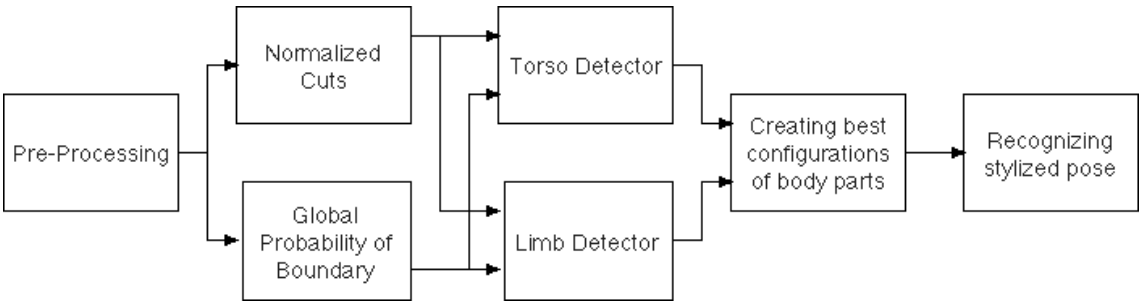

Figure 1: Dataflow of approach. The first step is the Pre-Processing to increase image contrast and reduce noise. Then the Normalized Cuts and the Global Probability of Boundary algorithms are applied to the pre-processed data. The next step is the torso and limb detection. After that the best possible body configurations are created and finally we try to recognize a stylized pose in the set of configurations

The configuration with the highest score is not always the true configuration, since a more likely pose could have a higher score. Mori et al. do not present a solution for the problem of picking the right body configuration out of the shortlist.

Ramanan et al. [16] present an approach for tracking people by finding stylized poses using a pictorial structure framework. They look for a certain canonical pose in a video sequence and then start tracking the person. That is suitable for our project since swimmers have poses, known as the stroke phases. We use this idea to recognize stylized swimming stroke phases in the video. Figure 1 shows the developed process we apply for our images.

3. Results

We have tried our method with example of underwater video footage for a freestyle swimmer. The segmentation with the Normalized Cuts algorithm provides reasonable superpixel maps. The limbs segments are exactly represented by one super pixel in almost every frame. The boundary detector is able to precisely detect the body of the athlete and draws an almost closed contour around the swimmer. Other strong boundaries are only lines on the bottom of the pool and the water surface (figure 2).

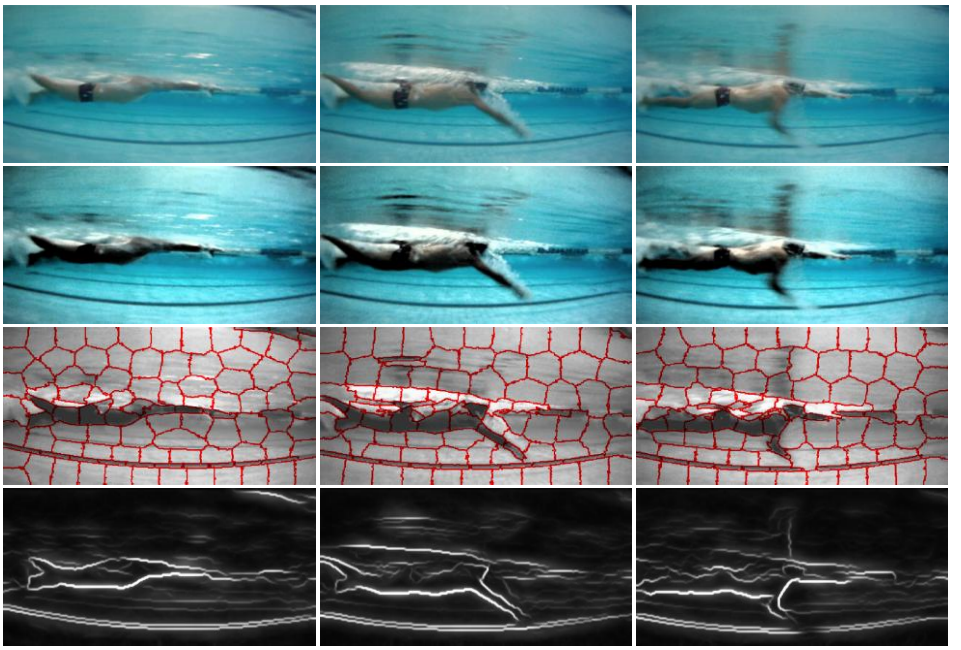

Figure 2: Results for three frames of a under water swimming video. On Top are the original input images with an athlete in entry (left), catch (middle) and insweep (right.) pose. The second row shows the same images after pre-processing. The third row contains the results of the Normalized Cuts algorithm and images on the bottom row are the output of the Global Probability of Boundary.

4. Discussion and Conclusions

In this paper we address a significant challenge in the sport of swimming, that of analysis of underwater video. The problems associated with video data collection such as variability due to operator,

artefacts from swimming strokes that occlude swimming segments are highlighted. A review of classic image detection methods reveals several techniques suitable for improving the view ability of underwater video, the enhancement of video as well as the automated detection of limb segments for future swimming phase analysis.

The results are satisfying and we demonstrate that is feasible to identify and estimate the body configuration of the swimmer. Our way of using unsupervised detection for swimming analysis is a novel approach to further enhance the quality and efficiency of coaching and training

Acknowledgments

This work is jointly funded by the Australian Research Council and the Queensland Academy of Sport.

References

[1] Davey NP, Andersen ME, James DA, An accelerometer-based system for elite athlete swimming performance analysis, *Proc. SPIE 5649, 409 (2005); doi:10.1117/12.582264*

[2] Lerda R, Cardelli C, Chollet D. Analysis of the interactions between breathing and arm actions in the front crawl. *J Human Movement Studies* 2001;**40**:129-144.

[3] Seifert L, Chollet D, Allard P. Arm Coordination symmetry and breathing effect in front crawl. *Human Movement Science* 2005;**24**: 234-56.

[4] James DA, Burkett BJ, Thiel DV, An unobstrusive swimming monitoring system for recreational and elite performance monitoring. Proc. Engineering 2011;**13**:113-9

[5] Hay JG, Liu Q, Andrews JG. Body roll and hand path in freestyle swimming: A computer simulations study. *J Applied Biomechanics* 1993;**9**:227-37.

[6] Payton CJ, Bartlett RM. Estimating propulsive forces in swimming from three-dimensional kinematic data. *J Sports Sciences* 1995;**13**:447-454.

[7] Yanai T, Hay JG. Shoulder impingement in front-crawl swimming: II. Analysis of stroking technique. *Medicine and Science in Sports & Exercise*. 200;**32**, 30-40.

[8] Maglischo EW. *Swimming Fastest*. Illinois: Human Kinetics; 2003

[9] Zamparo P, Lazzer S, Antoniazzi C, Cedolin S, Avon R, Lesa C. The interplay between propelling efficiency, hydrodynamic position and energy cost of front crawl in 8 to 19-year-old swimmers. *Eur J Appl Physiol* 2008;**104**:689-99

[10] Li S, Lu H, Xiang R, Chen Y. Human body segmentation based on deformable models and two-scale superpixel. *Pattern Analysis & Applications* 2011(5 May 2011), pp. 1-15

[11] Mori G, Xiaofeng R, Efros AA, Malik, J. Recovering human body configurations: combining segmentation and recognition. *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on* , 2004;2:II-326 - 333

[12] Shi J, Malik J. Normalized cuts and image segmentation. *IEEE Trans PAMI* 2000;**22**(8):888-905

[13] Martin D, Fowlkes C, Malik J. Learning to find brightness and texture boundaries in natural images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2002;**26(5)**:530-49

[14] Maire, M., Arbelaez, P., Fowlkes, C., Malik, J. Using contours to detect and localize junctions in natural images. *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, 2008;**1**:1-8

[15] Arbelaez, P., Maire, M., Fowlkes, C., and Malik, J. From contours to regions: An empirical evaluation. *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on* , 2009;**1**:2294-2301

[16] Ramanan, D., Forsyth, D.A. and Zisserman, A. Strike a pose: tracking people by finding stylized poses. *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on* 2005;**1**:271- 278

[17] James, D.A., ADAT: A Matlab toolbox for handling time series athlete performance data, *Procedia Engineering* 2011;**13**: 451-456