The Use of Change Detection Algorithms for Motion Detection in Sports

Jonathon M. Freer

Miami University

ABSTRACT

The ability to detect change between images has received increased interest due to the rise of important applications such as driver assistance systems, video surveillance, and medical diagnosis and treatment. Change detection algorithms used while processing video of sports can provide an opportunity to increase the accuracy of the gathering and analysis of sports statistics by tracking motion of different objects and players of the sport. Algorithms bound by the application of a particular sport can be combined with the general concept of change detection to accurately and efficiently recognize important "actions" and "events" that occur over the duration of the sporting event. Several systems have been developed that can track the motion and paths of players along with moving objects for a particular sport by using general change detection concepts accompanied by sport-specific algorithms. This paper focuses on two systems as examples that use change detection algorithms for analyzing video of the sport of tennis.

Keywords: change detection, change mask, binary image, binary map, simple differencing, motion tracking, segmentation.

1 INTRODUCTION

The interest of being able to detect change in images has significantly increased due to the rise of important applications such as driver assistance systems, video surveillance, and medical diagnosis and treatment ([3], Section I). However, the use of change detection algorithms in the domain of sports provides a unique set of benefits. It is common practice to keep track of sports statistics by hand, which not only can quickly become tedious and redundant, but is also limited to human vision and attentiveness. With the use of change detection algorithms, along with other domain (or sport) specific algorithm additions, valuable data can be derived via semantic analysis of the change detected in input images. In the domain of sports, statistics could be tracked using systems that use change detection algorithms to identify and record important "actions" or "events" such as the paths and motions of the players, the path and motion of a puck in ice hockey, and whether or not a goal has been scored in soccer.

The most fundamental aspect of change detection algorithms involves identifying the differences between two images. The algorithm then generates a change mask. In a particular change detection algorithm referenced as "simple differencing", a threshold is often used when generating a change mask in order to increase its accuracy of classifying "important" or "relevant" changes in the compared images ([3], Section IV). The process of segmentation is classifying the changes detected by semantic type ([3], Section I). The algorithms and methods involved with this classification process are domain (application) specific, as each application is composed of its own unique purpose. For example, an algorithm that analyzes a change detected in a surveillance video to see if it matches a picture of a missing person's face would not be relevant to an algorithm with a purpose of distinguishing whether or not changes detected are a moving football or a moving football player.

2 BACKGROUND

The main set up of these systems that use change detection algorithms for sports are composed of cameras capable of recording video. Recall that video is composed of **frames**, or images, that are sequentially shown at a particular frame rate. The images are composed of pixels, which depending on the type of image can be a variety of colors and intensities. Although it may appear obvious, the purpose of change detection algorithms is to identify the differences (or changes) between two images. However, what may not appear obvious is the complexity of this task. Changes in images can be attributed to the change in location and/or intensity of direct and/or ambient light sources, objects changing location or orientation, or even that the camera has shifted ([3], Section II). Fig. 1 demonstrates this complexity. Notice how specular reflections (labeled as "S"), the movement of objects (labeled as "M"), and then variations of appearances (labeled as "A") all qualify as changes.

Fig. 1. An example demonstrating the various causes of changes in images ([3], Fig. 1).

Change detection algorithms create what is known as a **change mask**, also referred to as a **binary image** or **binary map**. These images typically are composed of two colors or intensities (which warrants the term "binary" in binary image and binary map). The change mask can then be used to show the changes between the two images being compared. The semantic classification of these identified changes is known as **segmentation**.

3 ALGORITHM DESCRIPTION

The following sections 3.1 and 3.2 describe the general aspects of change detection algorithms. The mathematical definitions displayed in these sections can be found in [3], Section II. Section 3.3 explains the "simple differencing" change detection algorithm ([3], Section IV). Section 3.4 briefly touches on algorithms used in addition to change detection algorithms to aid in the detection of "important" changes ([3], Section III).

3.1 The General Concept

The most fundamental aspect of change detection is centered on the difference between images. Let the sequence of images being analyzed be denoted as the set $\{I_1, I_2, \ldots, I_M\}$ where M is the total number of images within the sequence. Each I is a single image of the image sequence and is composed of pixels. Denote a pixel coordinate as x and let the function I(x) represent the pixel coordinate of an image I. Note that each pixel coordinate x has a color or intensity that can be described using a set of real numbers and where the size of the set can be described by *l*. For example, a value of *l* equal to 3 represents RGB color images (because each pixel's color or intensity is a set of 3 real numbers, where the first is the amount of red color in the pixel, and the remaining numbers represent the amount of green and blue color, respectively. Put mathematically, a pixel coordinate $x \in \mathbb{R}^l$ is then mapped to a color or intensity such that $I(x) \in \mathbb{R}^k$. In most instances, k is 1 (grey scale) or some other relatively small number. For the domain of sports, M is very large, as video is needed for accurate and useful change detection in sports.

3.2 Generating the Change Mask

Assuming that we are using a value $k = 1$, a pixel coordinate in the change mask is equal to 1 if that pixel coordinate has been determined to contain a change at pixel coordinate x of I_M, or 0 if no change is detected at pixel coordinate x of I_M. Mathematically speaking:

$$B(\mathbf{x}) = \begin{cases} 1, & \text{if there is a significant change at pixel } \mathbf{x} \text{ of } I_M \\ 0, & \text{otherwise.} \end{cases}$$

3.3 Simple Differencing

The algorithm of simple differencing applies a threshold to this general concept. If a pixel coordinate exceeds that threshold, it is then deemed an "important" change and is recognized in the change mask. In mathematical terms:

$$B(\mathbf{x}) = \begin{cases} 1, & \text{if } |D(\mathbf{x})| > \tau \\ 0, & \text{otherwise.} \end{cases}$$

where the function D signifies the differences in color/intensities of a given pixel coordinate between the two images and the threshold is a value typically chosen empirically. The threshold can also be chosen to achieve results desired by a particular application or domain.

3.4 Alternative Algorithms

There are many instances when just a simple threshold as used in simple differencing will not suffice. Many real world applications of change detection algorithms have "alternative", or more complex methods of determining which changes are "important" and which are "unimportant". It is often that these algorithms are used prior to the differencing of the images and are referred to as the "preprocessing" steps ([3], Section III). Since these algorithms are application specific and are separate algorithms themselves, the next sections will give a high level overview of the more common preprocessing algorithms. As illustrated in section 4, many sport-specific systems employ their own algorithms to aid the change detection algorithm in its effort to classify "important" changes. All preprocessing algorithms are discussed in much higher detail in resource [3], Section III.

3.4.1 Common Preprocessing Algorithms

There are two main types of preprocessing procedures that are used to filter out common "unimportant" changes: geometric and radiometric adjustments ([3], Section III).

3.4.1.1 Geometric Adjustments

It is a common occurrence that the camera being used moves slightly, causing pixel intensities to change and be recognized as changes. However, these changes are barely ever desired and are seen as "noise" or interference. To eliminate the issue, it is common to use image registration algorithms. These algorithms essentially align several images into the same coordinate frame by performing low-dimensional-spatial transformations ([3], Section III).

3.4.1.2 Radiometric (Intensity) Adjustments

There exist scenarios when light sources cause intensity variations due to their change in position or intensity. Several algorithms are used to compensate for these variations before performing change detection. One of the algorithms used is known as intensity normalization. This algorithm normalizes the pixel intensity of an image to the pixel intensity of another image. Another algorithm, known as "sudden changes in illumination", categorizes changes that are caused by a sudden change in illumination as "unimportant" ([3], Section III). Other algorithms that deal with Lambertian surfaces can be found in [3], Section III and won't be discussed here.

4 CHANGE DETECTION IN SPORTS

All sports have different environment set ups, rules, and standards that warrant the use of specific algorithms in conjunction with the change detection algorithm. Despite the need for these sport-specific algorithms, the change detection algorithm represents the core process that makes motion tracking possible. It is common in sports to use an image of the empty playing field as a "background image". The introduction of players and other moving objects will be displayed as changes in reference to the image of the empty playing field by the change detection algorithm. However, image "noise" is often an issue within the change mask. There are many changes that the change detection algorithm may detect that are undesired, and section 3.4 shed light on some alternative algorithms that attempt to filter out this noise. The following sections 4.1 and 4.2 describes how the change detection algorithm, along with other sport specific algorithms have been used to track player motion and ball motion in the sport of tennis.

4.1 Tennis System #1

The following system found in [2] heavily relies on change detection algorithms to track the motion of tennis players and the tennis ball during a tennis match. The main goal of this system was to track the player's trajectory during gameplay. There are several algorithms associated with dynamically clustering paths, analyzing curvatures of contours of changes detected, merging clusters, and analyzing trajectories that were used in addition to a change detection algorithm to complete this task ([2], Section 3). Even though these algorithms are critical to the success of this system's goal, they require extensive explanation that would stray from the purpose of this paper. However, since most of these algorithms are used during player motion tracking, the next section (4.1.1) can provide a clearer example of how change detection, in addition to simpler, more sport-specific algorithms, are used to track tennis ball motion.

4.1.1 Tracking Ball Motion

In the sport of tennis, there is typically a standardized yellow color of the tennis ball. This provides a unique opportunity to use the color of the ball as a method of detecting its motion throughout the course of the game. Fig 2 shows the hue of a tennis ball in images that were captured with different lighting conditions ([2], Section 4). Fig. 3 shows the saturation of the tennis ball in these images. This system uses this information about the hue and saturation of the ball to identify its motion.

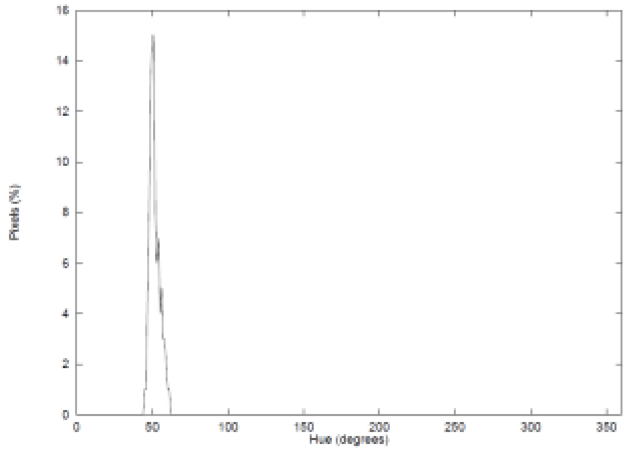

Fig. 2. The hue distribution of the tennis ball ([2], Figure 4).

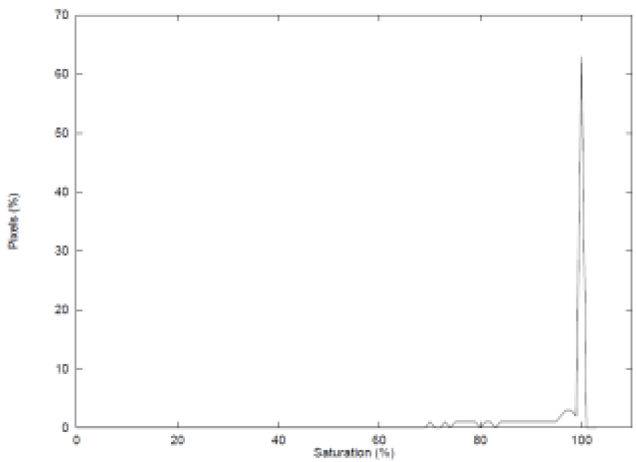

Fig. 3. The saturation distribution of the tennis ball ([2], Figure 5).

This system uses simple differencing to locate areas of change between the current frame and the frame that preceded it using a threshold. Morphological closing is then used to close the gaps in the change mask (for more information on morphological closing, consult [2], Section 3). Then, in order to distinguish the changes detected involving the players from the changes detected involving the tennis ball, the regions of change are converted to HSV (Hue Saturation Value) space and are compared to the hue saturation model of the tennis ball. The change region that has the most similarity to the hue saturation model of the tennis ball is then segmented as the tennis ball ([2], Section 4). In order to track the path that the tennis ball has traveled throughout gameplay, this system calculated the center of the region segmented as the tennis ball and matched it to the location of the tennis ball in the previous frame. Fig. 4 and Fig. 5 illustrate this segmentation and path construction, respectively.

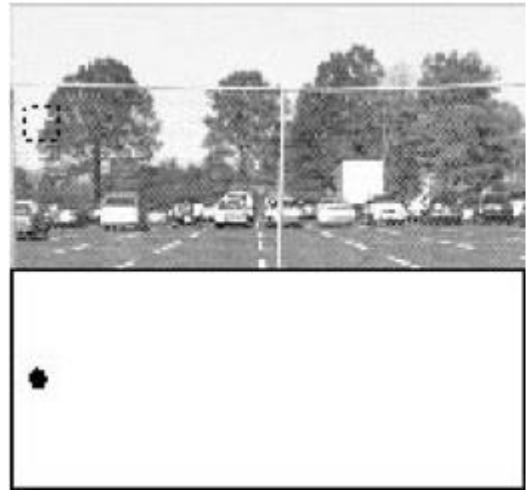

Fig. 4. The segmentation of the ball using a sport-specific algorithm in combination with change detection ([2], Figure 8).

Fig. 5. The path of the tennis ball ([2], Figure 9).

4.2 Tennis System #2

The system found in [1] also was used to analyze the sport of tennis as in the previous section (4.1). However, this system did not choose to track the ball as the previously mentioned system did.

In contrast to the previously mentioned system, a significant amount of preprocessing was performed on the image sequence before using a change detection algorithm to identify changes and to segment the changes. This system begins by identifying where in the images the court is located by using mean values for each color space for the four types of court classes in tennis ([1], Section 3.1)(Fig. 6). This increases the efficiency of this system by breaking down the image into “areas of interest” that can be searched for changes, rather than searching the entire image.

Fig. 6. Demonstration of the identification of the playing field. The region between the two horizontal black lines has been identified as the playing field ([1], Figure 2).

Since court lines are white by standard in the game of tennis, this system extracts these pixels (Fig. 7). This system then compares the number of white pixels extracted to the number of white pixels extracted for the court in the previous frame. If the difference between these values is less than a chosen threshold, then the selected area is confirmed as the court ([1], Section 3.1).

Fig. 7. The tennis court with black lines representing the search area for these court lines ([1], Figure 2).

4.2.1 Tracking Player Motion

Once this system has constructed a more useful and efficient “background” image, it uses the simple differencing change detection algorithm to produce a change mask. In this system, the pixels in the change mask are first initialized to a value of 255 (white) and are then changed to a value of 0 (black) if the difference of values between a pixel coordinate from the background model and the same pixel coordinate from the current frame is larger than a particular threshold ([1], Section 3.3). Fig. 8 demonstrates an example of a change mask produced by this system.

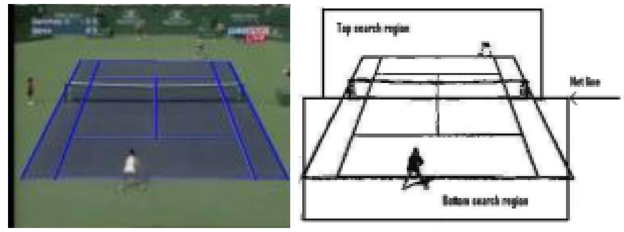

Fig. 8. Example change mask created by the system described in section 4.2 ([1], Figure 3).

This system then divides the change mask into two sections, each representing a region to search for each player. Then a window with a predetermined size representing the player bodies is used to search each section (top and bottom) for the player. Each pixel in each section is set as the center of the player body window, and then the count of the number of 0 values within the player body window is recorded. Once all of the pixels have been checked (all pixels have been the center of the player body window), the player body window is centered on the pixel that had the maximum 0 count. This algorithm is applied to both the top and bottom search regions. The player body window for the top search region is smaller due to the perspective of the camera used ([1], Section 3.3). Once the player positions are determined for the first time, a small local search area encompassing the player body windows is used to search for the updated player positions (positions of the player body windows). The same algorithm that finds the maximum 0 count is applied to these local search areas and the position of the player body window within this local search area is updated ([1], Section 3.3).

5 EVALUATION

The performance of a particular system using change detection algorithms is subject to the application or domain in which it is being used. For example, a system using a change detection algorithm to track the location of a football would be evaluated visually based on its ability to track the football throughout the duration of its use.

An approach that has been deemed reliable is to display a flicker animation. A flicker animation is video that is composed of a pair of images that are played rapidly at intervals of about a second each. Changes that are detected are then displayed and appear to “flicker”, hence the name of this evaluation approach ([3], Section X).

Another common approach of assessing the performance of change detection algorithms is to use an expert observer. For example, changes detected between two X-ray images could be evaluated by a radiologist, at which can then express the “correctness” of the result of the change detection algorithm ([3], Section X). For the systems shown in previous sections, an observer of the game could confirm the accuracy of the tracking of players and moving tennis ball during gameplay.

The performance of the system discussed in section 4.1 when tracking the motion of a tennis ball was assessed with a ball speed of roughly 60 miles per hour and reported to be “very encouraging” ([2], Section 5). The performance of the system discussed in section 4.2 was assessed using three tennis video sequences, consisting of more than 20 minutes of video and 10,000 frames. The algorithm used to detect the playing field was reported to have a 99% detection rate. The algorithm for detecting player motion was assessed on the criteria of whether or not 70% of the player’s body was encapsulated by the player body window. It was reported that the accuracy of this algorithm was about 98% ([1], Section 4).

Simple differencing was used in both systems mentioned in sections 4.1 and 4.2, and is still used in many applications ([3], Section IV). However, simple differencing is unlikely to produce the best results by itself. Simple differencing alone often is too sensitive and captures many changes that are not “important”. Many algorithms used in conjunction with change detection algorithms are more likely to outperform simple differencing ([3], Section IV). However, the system mentioned in 4.1 (from resource [2]) claimed that simple differencing was chosen because it is a “fast operation and works across varying light conditions” ([2], Section 3). The systems mentioned in 4.1 and 4.2 used several other algorithms as well as simple differencing to achieve their overarching goals.

REFERENCES

[1] Han, Jungong, Dirk Farin, Peter H.N. De With, and Weilun Lao. "Automatic tracking method for sports video analysis." Eindhoven University of Technology, 2005. Web. 8 Mar. 2014.

[2] Pingali, Gopal Sarma; Jean, Y.; Carlbom, I., "Real time tracking for enhanced tennis broadcasts," *Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference on* , vol., no., pp.260,265, 23-25 Jun 1998.

[3] Radke, R.J.; Andra, S.; Al-Kofahi, O.; Roysam, B., "Image change detection algorithms: a systematic survey," *Image Processing, IEEE Transactions on* , vol.14, no.3, pp.294,307, March 2005.