Real-Time Analysis and Profiling of Coordinated Movements in Two-Dimensional Space using Footage from Multiple Cameras

Grace Amankwah
Zhejiang University of Science and Technology
School of Information and Electrical Engineering
Xiaoheshan, Zhejiang, China

William Morgan-Darko
Kwame Nkrumah University of Science and Technology
Department of Computer Science
Kumasi, Ghana

Yang Chunting
Zhejiang University of Science and Technology
School of Information and Electrical Engineering
Xiaoheshan, Zhejiang, China

ABSTRACT
Coordinated movements in teams involve strategic positioning and motions that are coordinated among a team of people. Since there are numerous coordinated movements that are in common use by military, criminal and sporting teams, experts in the fields of security and sports dedicate much of their time to understanding, devising and recognizing tactics that are in play in a particular scenario. In this study, a framework for automated analysis and profiling of coordinated movements using the machine vision techniques of visual object tracking and 3-dimensional reconstruction is discussed. The study also considers the feasibility of using real-time analysis of coordinated movements in assisted surveillance to buttress security. In this paper, a simplified description of the framework is proposed and its operation is explained.

General Terms
Machine Vision, Machine Learning, Pattern Recognition

Keywords
Tactical Movement, Team Coordination, 3D Reconstruction; Visual Object Tracking, Trajectory Mapping, Assisted Surveillance, Sports Analysis

1. INTRODUCTION
The term Coordinated Movements is used in this paper to describe planned, and possibly rehearsed, sets of movements carried out by two or more people who are acting in agreement with the intention of achieving a common goal [1]. Team sports involve coordinated movements among the teammates, according to what the team has trained for and been directed by their coach [2]. Many teams have distinct playing styles which they have refined to enhance their efficiency on the field of play and by which they can easily be identified.

Aside sports teams, various teams of professionals also make use of coordinated movements in their activities. Law enforcement agencies have periodic training sessions where they rehearse their movements in various hypothetical scenarios, in anticipation of events that they may encounter in the line of duty [3, 4]. Likewise, the Fire Department and the Armed Forces also practice coordinated movements to use in their operations. When it comes to security, there are some experts who are able to recognize which combat team is operating by assessing their movements and the coordination among the team. Some security experts are also able to identify criminal groups by likewise observing the coordination among the criminals as they move.

While human experts have been capable of analyzing, profiling and recognizing coordinated movements in real-time, an open discussion on implementing such a system in software is scarce in the Computer Vision society.

The discussion in this paper puts forward an approach to collecting data on coordinated movements using existing Computer Vision technologies that have been put together to achieve that specific goal. The technologies that are considered in this paper are Visual Object Tracking and 3D Reconstruction. The authors believe that these two technologies are a good starting point for the discussion and subsequent implementation of an analytical framework for coordinated movements.

2. 3D RECONSTRUCTION FROM MULTIPLE IMAGES
An aerial view of a delimited space is required for the method described in this discussion. Several platforms exist for taking aerial photographs such as helicopters, drones and kites but most are not suited for working indoors [5]. As such the established or common aerial photography platforms are unsuitable for this particular application. In order to arrive at the aerial view, the authors propose first obtaining a 3-dimensional reconstruction of the space using the footage from multiple cameras.

3D reconstruction refers to the creation of 3-dimensional models from a set of 2-dimensional images [6]. Many buildings, rooms and other walled or fenced enclosures are equipped with surveillance cameras for the purpose of monitoring the space. For the purpose of this particular discussion, frames from the camera footage are taken from the same instant and used to create a 3D model.

After the 3D model is created, the image is rotated so that it is being viewed from the top. This gives a 2-dimensional view of the space in which the coordinated movement is happening.

3. VISUAL OBJECT TRACKING
Visual Object Tracking refers to identifying the location and dynamic configuration of one or more moving objects in a video footage [7]. Visual object tracking is one of the challenging problems in the field of computer vision. It has several related applications such as automated video surveillance, traffic monitoring, security and robotics [8]. The objective of a visual object tracker is to estimate the location of a target in all the frames of a video sequence based on the given initial location (or a bounding rectangle) of the target.

Object tracking problem has been studied by the computer vision community for several decades. But still it remains a

challenging task to design an efficient and robust visual object tracking system for all the practical real-world applications [9]. Further, there are several factors that affect the performance of the object tracker such as illumination variations, scale variations, occlusions, deformations, motion blur, rotations, and low resolutions [10].

This paper proposes the use of visual object tracking in following the trajectory of the people engaged in a coordinated movement, in relation to other people or objects in the video feed. The particular application in this discussion has to do with tracking dots on a two-dimensional plane.

4. OVERVIEW OF AUTOMATED COORDINATED MOVEMENT ANALYSIS ALGORITHM

- Construct a two-dimensional aerial view of the action sequence using footage from the multiple cameras.

- Identify reference points from the footage, tag and track them in the aerial view.

- Identify the performers of the tactical movement from the footage, tag and track them as the subjects on the aerial view.

- Track the motions and positions of the subjects in relation to the reference points and use the tracking data to create a list of tactical movements.

5. TRAJECTORY AND COORDINATION MAPPING

Trajectory Mapping in this discussion refers to tracking the movements of the subjects along the floor of the grounds, relative to the subject's initial and previous positions as well as the positions of the reference points.

Coordination Mapping in this discussion refers to tracking the movements of the subjects along the floor of the grounds, relative to the positions and movements of the other subjects being tracked.

The record of the trajectories and coordination of the subjects becomes the digital profile of the coordinated movement.

5.1 Subjects and Reference Points

The main focus of the tracking algorithm is the movements of the performers of the coordinated movement, referred to as the subjects. Their motion (or lack thereof) forms the basis for the analysis of the coordinated movement.

For the purpose of observing the movements of the subjects in the coordinated movement, a number of reference points need to be set, against which the motion of the subjects can be measured. For instance, on a sports field there are often marking on the ground which can be used as reference points against which the positions and motions of the players can be measured. Another example is in the case of a tactical strike by a military group in a delimited space, in which case such things as the door, windows and stairs may be used as reference points.

Furthermore, some reference points may be moving objects or even people such as is the case with the ball in a football game or the victim in case of a robbery. In some cases, other performers in the tactical movement can be used as reference points against which the motions of another performer can be measured.

This gives rise to two kinds of reference points – static reference points and movable or moving reference points. In all cases the front (F), the back (B) and the sides (LS for left side and RS for right side) of the reference point are to be indicated in order to facilitate the correct placement of the positioning of the subjects. Figure 1 shows a visual representation of the structure of Subjects and Reference Points.
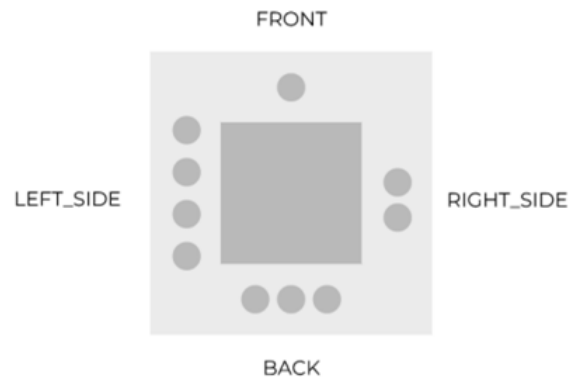

Fig 1: Visual Representation for Subjects and Reference Points with dots to indicate the Front (1 dot), Right Side (2 dots), Back (3 dots) and Left Side (4 dots).

Listing 1 describes the class structure for Subjects and Reference Points.

```
class TrackableObject {

    String name;

    Side side;

}


enum Side {

  FRONT, BACK, LEFT_SIDE, RIGHT_SIDE

}


class Subject extends TrackableObject {}


class Reference extends TrackableObject {}
```

5.2 Relative Movements

The motions of the subjects are considered relative to one or more reference points. While distance is considered, the parameter of greater importance in a coordinated movement is the direction of motion for a moving subject and the position of rest for a static subject. A number of directions and positions are presented.

5.2.1 Towards (TW)

In this motion, the subject closes the distance between themselves and a reference point. Figure 2 shows a visual representation of the "Towards" relative movement. The subject is tagged with an "S" and the reference is tagged with an "R".
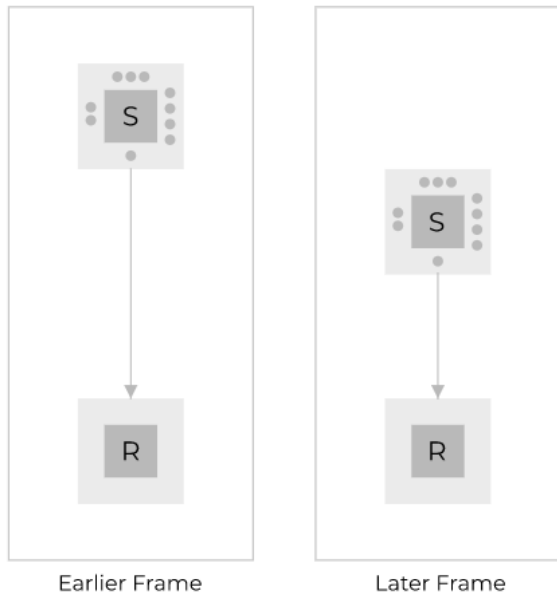
Fig 2: Towards (TW) movement

5.2.2 Away (AW)

In this motion, the subject widens the distance between themselves and a reference point. Figure 3 shows a visual representation of the "Away" relative movement.
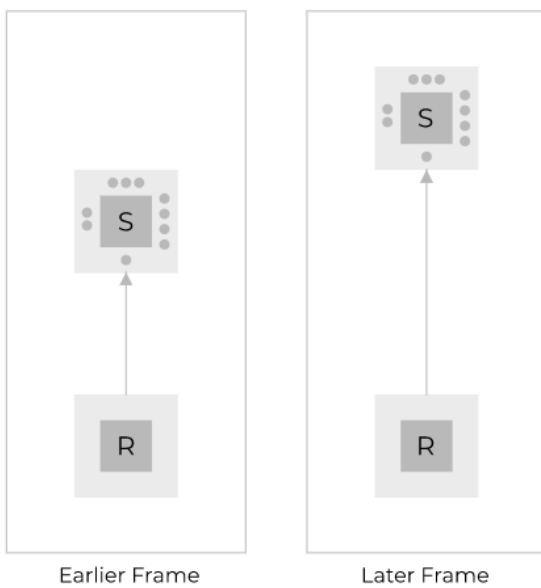

Figure 3. Away (AW) movement

5.2.3 Across (AC)

In this motion, the subject moves in the space between two reference points. Figure 4 shows a visual representation of the "Across" relative movement.
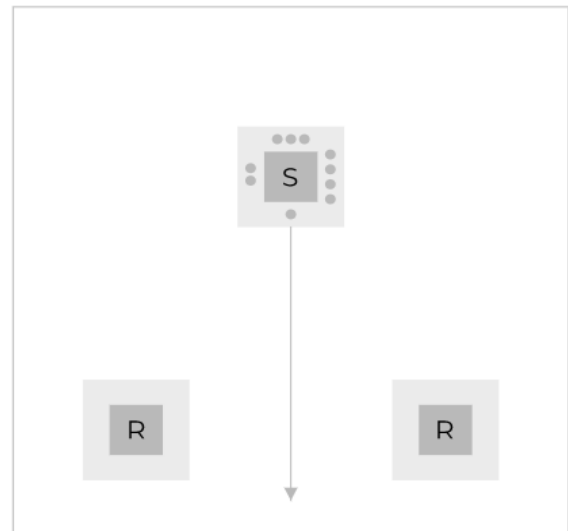

Fig 4: Across (AC) movement

5.3 Static Positioning

It is also very important to take note of the position of the subject in relation to one or more reference points when they are not in motion. Staying still is also a part of coordinated movement. A number of positions are presented here.

5.3.1 Beside (BS)

This refers to when the subject is positioned to the side of the reference point. Figure 5 visualizes the "Beside" positioning.
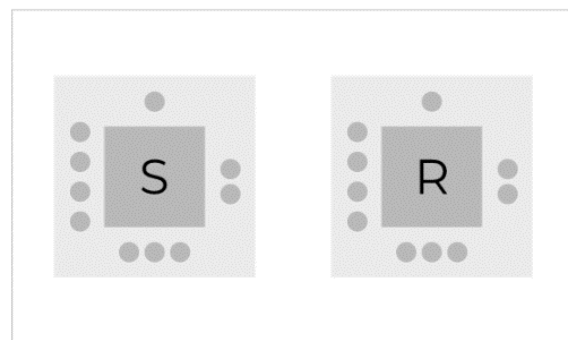

Fig 5: Beside (BS) positioning

5.3.2 In Front (IF)

This refers to when the subject is positioned at the front of the reference point. Figure 6 visualizes the "In Front" positioning.
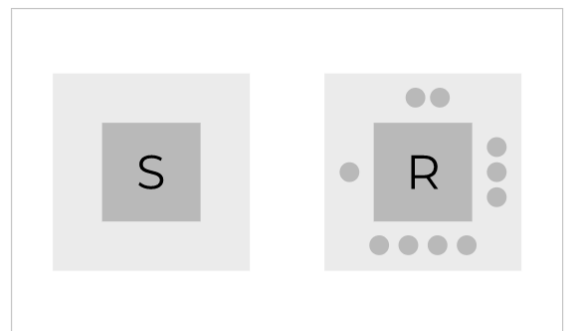

Fig 6: In Front (IF) positioning

5.3.3 *Between (BW)*

This refers to when the subject is positioned somewhere between two reference points. Figure 8 visualizes the "Between" positioning.

Fig 7: Between (BW) positioning

5.3.4 *Behind (BH)*

This refers to when the subject is positioned at the back of the reference point. Figure 7 visualizes the "Behind" positioning.
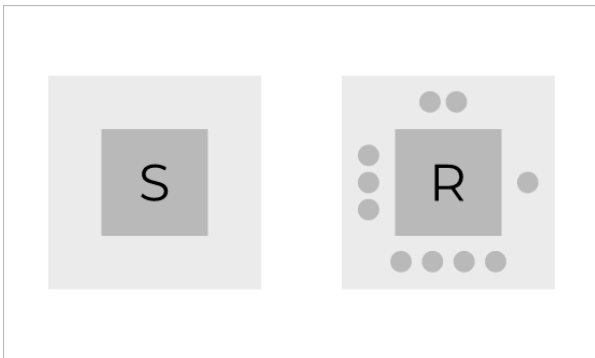

Fig 8: Behind (BH) positioning

5.4 Motion Sequences

The motions of the subjects in a coordinated movement are timed to occur either in sequence or concurrently. Sometimes there is a delay in action time after one motion is completed. The types of timing to take into consideration are presented here.

- Before (BF): As the name implies, this refers to timing a movement before another one begins. However, the "before" timing is harder to spot in a progressive analysis. Instead it is more easily spotted in retrospect.

- After (AF): This refers to timing a movement to begin when another one has ended. In a sequence of motions, it is always easier to spot "after" than "before" timings.

- Concurrent (CC): This refers to timing two movements to begin at the same time.

6. DATA REPRESENTATION

This section proposes structures for the component data units for the proposed analysis framework.

6.1 Motion Sequences

The motions of the subjects in a coordinated movement are timed to occur either in sequence or concurrently. Sometimes there is a delay in action time after one motion is completed.

The types of timing to take into consideration are presented here.

Units of data in the tactical movement represent either the motion or positioning of a subject or reference point in the progress of the tactical movement. Each motion data unit would indicate the subject being tracked, the motion undertaken, the points of reference and, if necessary, the timing of the motion.

Since there may be more than one reference point against which a motion is measured, the field for the reference points needs to be a data structure that supports multiple values, such as an array or list.

Listing 2 shows the class structure of the tactic data unit.

```
class Tactic {

    enum Motion {

        TOWARDS,

        AWAY,

        ACROSS,

        BESIDE,

        IN_FRONT,

        BEHIND,

        BETWEEN

    }


    Subject subjects;

    List<Reference> reference_points;

    Motion movement;

    Timing timing;

}
```

6.2 TIMING DATA UNIT

By default, it can be assumed that the next data unit in a list of data units contains a motion or positioning that is timed after the previous data unit. That makes the timing field of the primary data unit optional.

However, in the event of a concurrent timing, it becomes necessary to indicate that the action occurs while another subject is in motion. Storing the timing data can be achieved by specifying the type of timing (before, after, concurrent) and a reference to the tactics data unit which is the reference for the timing. Listing 3 shows the class structure for the timing data unit.

```
class Timing {

    enum Type {BEFORE, AFTER, CONCURRENT};

    Type timing_type;

    Tactic reference;

    Integer delay;

}
```

6.3 Full Data Structure

he full data for the coordinated movement should comprise of the list of reference points, the list of subjects and the list of tactics. Listing 4 shows the class structure for the full data.

```
class CoordinatedMovement {

    List<Subject> subjects;

    List<Reference> reference_points;

    List<Tactic> tactics;

}
```

7. Real-time Analysis of Coordinated Movement

This section discusses the means for implementing the real-time automated analysis algorithm.

7.1 Motion Library/Database

For real-time analysis of coordinated movements, a comparison of the profile of an ongoing movement can be matched against a library or database of coordinated movements. To achieve more efficient matching, the library can be read into memory in a tree data structure with each path from root to leaf representing a single profile. Also, each node in the tree denotes a single movement or positioning.

7.2 Matching

Matching a coordinated movement to a predefined profile involves locating the path in the motion library tree that most closely matches the profile of the currently playing coordinated movement. A trace of the predefined profile can be conducted by traversing the tree from its root, following the movements that are currently playing out until a leaf is reached or the path being traversed doesn't have any child nodes that match the coordinated movement that just played out.

7.3 Potential Applications

The automated analysis and profiling of coordinated movements algorithm can be applied in a predefined space with multiple overhead cameras and easily identifiable points of reference. It can be applied on sports fields for group sports such as basketball, soccer, volleyball and handball. During practice, the automated analysis can be used as a tool for refining the team's tactics to near perfection.

In addition, the automated analysis and profiling algorithms can be applied to assisted surveillance systems in places such as train stations, banking halls and airport terminals, to identify suspicious coordinated behavior that may turn out to be security threats.

7.4 Feasibility

As mentioned earlier, among the main components for running the algorithm are multiple overhead cameras and easily identifiable reference points. In addition, a plan of the grounds is also useful for rendering an aerial view for human observation.

The main challenge for running this algorithm has to do with the variable number of subjects being tracked. For instance, in the case of a sports play with a definite number of players, a definite number of threads will be dedicated to tracking the movements of the subjects and only one matching operation is required to identify the coordinated movement that is being assessed at the time.

However, in the case of monitoring a space such as a train station where the associations of the subjects are not known beforehand, multiple combinations of the subjects' movements need to be ran concurrently, requiring much processing power. As such, a very powerful system of computers needs to be dedicated to running the assisted surveillance system.

8. REFERENCES

[1] Garganta, J. (2009). Trends of tactical performance analysis in team sports: bridging the gap between research, training and competition. Revista Portuguesa de Ciências do Desporto, 9(1), 81-89.

[2] Helsen, W. F., Starkes, J. L., & Hodges, N. J. (1998). Team sports and the theory of deliberate practice. Journal of Sport and Exercise psychology, 20(1), 12-34.

[3] Sohn, V. Y., Miller, J. P., Koeller, C. A., Gibson, S. O., Azarow, K. S., Myers, J. B., ... & Rush Jr, R. M. (2007). From the combat medic to the forward surgical team: the Madigan model for improving trauma readiness of brigade combat teams fighting the Global War on Terror. Journal of Surgical Research, 138(1), 25-31.

[4] Smith-Jentsch, K. A., Zeisig, R. L., Acton, B., & McPherson, J. A. (1998). Team dimensional training: a strategy for guided team self-correction.

[5] History of Aerial Photography Professional Aerial Photographers Association (retrieved 5 October 2016)

[6] Cheung, G. K., Kanade, T., Bouguet, J. Y., & Holler, M. (2000). A real time system for robust 3D voxel reconstruction of human motions. In Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662) (Vol. 2, pp. 714-720). IEEE.

[7] Ali, A. (2015). Visual object tracking - classical and contemporary approaches. Frontiers of Computer Science.

[8] Hu, W., Tan, T., Wang, L., & Maybank, S. (2004). A survey on visual surveillance of object motion and behaviors. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 34(3), 334-352.

[9] Yin, F., Makris, D., & Velastin, S. A. (2007, October). Performance evaluation of object tracking algorithms. In IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, Rio De Janeiro, Brazil (p. 25).

[10] Danelljan, M., Shahbaz Khan, F., Felsberg, M., & Van de Weijer, J. (2014). Adaptive color attributes for real-time visual tracking. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1090-1097).