Vol.2015-MBL-77 No.7
Vol.2015-ITS-63 No.7
2015/12/2

CuraCopter: Automated Player Tracking and Video Curating System by using UAV for Sport Sessions

Elder Akpa[1,a) Kazuki Fujisawa[1,b)] Cedric Konan[1,c)] Marko Trono[1,d)] William Brou[1,e)]
Keiichi Yasumoto[1,f)]

Abstract: This paper presents CuraCopter, an automated athlete tracking system using drones, designed for sports training. While existing systems require navigation devices to control the drone movements, our proposed system has the particularity of using features-based drone control. To control drone movements, CuraCopter uses the color of the athlete's shirt as the feature to track. Furthermore, video scenes produced by CuraCopter are the result of video curating rules that automatically switch between video scenes, recorded by the drone-camera, and two grounds-cameras. The aim of CuraCopter is to help coaching staff and athletes by providing them a better video of their training session, in order to improve their analysis and results.

1. Introduction

Drones are becoming widely present in daily activities and environments. They are mostly used in activities such as entertainment, delivery, search-and-rescue, agriculture, and military missions. Moreover, in the ubiquitous computing environment, some studies have showed the use of drone as a jogging partner [1] and the possibility for humans to interact with drones using gestures [2].

More recently, drones have been introduced in the sport domain as video capture tool. Some studies have shown that the use of video enables coaches to focus on the coaching tasks with video providing additional information for assessment of technical skills, feedback in training and reports [3]. For training, the main requirement is to get video of high quality from multiple angles, for real-time movements analysis and immediate behavior changes tracking. With this in mind, coaches and athletes are using drones to record video of their training sessions for performance analysis and skills improvement. However current drone-assisted systems are either controlled using PCs, tablets, smartphones [4] [5] [6]; or require a wireless wearable device to move the drone [7]. These requirements can hinder users in some activities, where athletes are not allowed to carry objects or need to have their hands free. Furthermore, it is difficult to setup many fixed cameras and edit video later.

In this paper, we propose a self-controlled drone system called CuraCopter that addresses the needs of athletes and coaches. CuraCopter can record sport videos during training sessions at the desired angle of view, and the obtained videos scenes are composed by automatically switching between ground-cameras and the drone-camera. We developed CuraCopter to support soccer, basketball and other sports especially individual sports such as snowboarding, canoeing, and tennis.

2. CuraCopter

CuraCopter is a combination of an automated drone and a video curating system that automatically follows an athlete and record his or her training session. The obtained video from Curacopter can be used for performance analysis and skills improvement. Our proposed system uses three kinds of video cameras: the drone camera video and two ground-cameras, which are a Soloshot2 [8] camera video and a fixed camera video. For the safety of the athlete and the materials, all cameras are positioned out of the game field, at a predefined distance for each camera.

Fig. 1: Example of an athlete using CuraCopter in close range.

1 NAra Institute of Science and Technology (NAIST)
a) akpa.elder.zx6@is.naist.jp
b) fujisawa.kazuki.ey4@is.naist.jp
c) konan.cedric.js5@is.naist.jp
d) marko.trono.mg8@is.naist.jp
e) brou.william$_{f}abrice.br3@is.naist.jp$
f) yasumoto@is.naist.jp

(a) Original image (b) HSV image (c) Thresholded image (d) Contour detection

Fig. 2: Image processing

The contributions of the CuraCopter are as follows.

- Automatically control drone with high accuracy by binary color extraction.
- Capture multiple video streams without user burden.
- Create video that user wants automatically based on defined rules

CuraCopter, with its tracking module, allows the drone to move automatically, while its curating module provides the rules for switching between cameras. Figure 1 shows CuraCopter being used in close range, during a training session of soccer.

We developed a prototype of CuraCopter using AR.Drone from Parrot Inc.

2.1 Drone tracking module

The tracking module controls and manages the drone flying motion (drone navigation) using a set of vision processing functions on some predefined features. In our current implementation of CuraCopter, we used the athlete shirt color as the feature to be tracked. We could use the athlete face or uniform number as the feature to be tracked. However, Curacopter is the system for one player and the tracking by face and uniform number is not stable. So we use only the athlete shirt color. Color has been widely used in real-time tracking systems [9], [10]. It offers several significant advantages over geometric cues such as computational simplicity, robustness under partial occlusion, rotation, scale and resolution changes. Color has the ability to discriminate various objects in the same environment. Our tracking module relies on this discriminative ability to track the athlete shirt. Our current implementation uses the orange color as color feature because there are few objects of similar color in ordinary training areas. To realize the tracking module, we have to answer the following questions: (1) how can the drone recognize a specified color, and (2) how is the drone instructed to move based on the real-time extracted color features? For the color extraction, the module converts the drone images from the RGB color space to the HSV color space, and thresholds the images, through a range check of the discriminative color, to display only the targeted color. Then, we compute the dimensions and extract the coordinates of the targeted within the actual Field of view (FoV) of the drone. Figure 2 describes the different steps of the video frames processing realized by the tracking module, using the open source computer vision library OpenCV.

For the drone navigation instructions, because we are implementing CuraCopter with the AR.Drone 2, we need to provide the parameters that will make the drone spin, gain or reduce altitude automatically, based on our targeted color motions. From the previous step of color extraction , we obtain the target dimensions and coordinates within the drone image, and then move the drone to put the target in the center of the FoV. As mentioned before, to ensure the safety of the athlete and the materials, the drone should track the athlete from outside the game field. Therefore, in the tracking module, we do not consider the pitch and the roll movements of the drone to prevent the drone to go front (within the game field) or back (far from the game field). We use only the yaw and the throttle flying movements of the drone to track the athlete. Section 3 gives more details about the drone self-navigation based on the target motions.

2.2 Video curating module

The video curating module automatically controls the switching between the drone video camera and the ground video cameras based on the training video preferences of the athlete. There are two kinds of automatic curation: rule-based curation and machine learning-based curation. In rule-based curation, video is selected by the predefined flowchart rule. Fujisawa et al. [11] proposed a system for automatic curation from multiple baseball video streams using machine learning. In their work, they used a live TV broadcast as ground truth data and constructed a switching timing estimation model. Machine learning-based curation is suitable if the number of video streams and metadata is very large, but it has the disadvantage of not being able to obtain sufficient accuracy if the number of training data is small. In this paper, we choose the rule-based curation because the number of video streams and metadata is not large.

CuraCopter uses three video cameras:

- Drone camera
- Soloshot2 camera
- Fixed camera

Figure 3 describes the two main components of Soloshot2. Soloshot2 is a stand-alone robot that tracks (i.e. tilt and pan) a person wearing a certain TAG. Soloshot2 uses GPS for tracking, if the distance between the BASE and TAG is within a 10 m range, or it is put in-door environment or around tall building and trees, the accuracy is very low. But, it can automatically track a person within a 600 m range and suitable for focusing on a certain moving target. The speed of tilt is 30° / sec (max tilt is 60° / sec) and the speed of pan is 80° / sec (max pan is 360° / sec), so it keeps matches and subject in frame accurately even in a situation such as surfing and bike where the subject is moving in a wide range and high speed.

(a) Soloshot2 BASE (b) Soloshot2 TAG

Fig. 3: Soloshot2 components

Figure 4 gives an example of curation of multiple video streams.

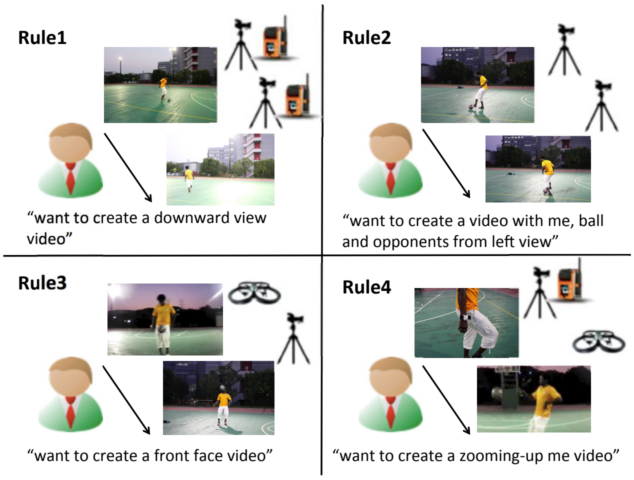

Fig. 4: Example of curation of multiple video streams

The curating module switches automatically among our three video streams based on their metadata in real time. The rules of content curation can be various. The player or the staff can set the rules by specifying the different view modes and scenarios such as "want to create a downward view video," "want to create a video with player, ball and opponents from left/right view," and so on. Based on these rule settings, the module realizes automatic content curation.

3. System configuration

We developed a prototype of CuraCopter by using AR.Drone 2 from Parrot Inc, and Soloshot 2 with an ordinary camera (Sony HDR-CX390) as ground cameras. Figure 5 describes the overall system flowchart.

3.1 Drone tracking module configuration

The tracking module controls and manages the drone navigation using a set of vision processing functions (Fig.6).

The system is configured to set an initial flight altitude H to the drone when taking off. Also for the security and safety of the

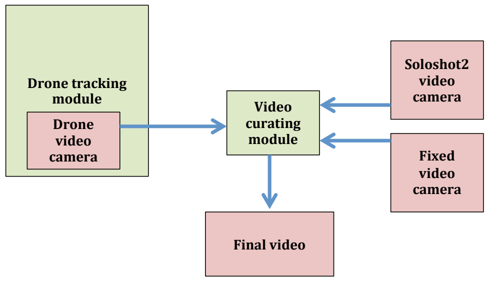

Fig. 5: CuraCopter's flowchart

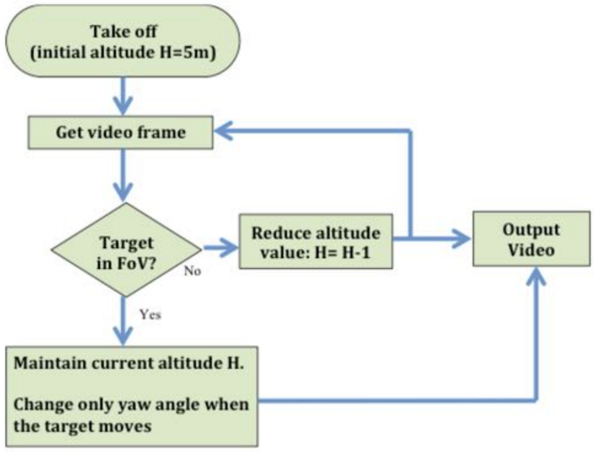

Fig. 6: Drone tracking module

athlete, we put the drone 3m to the field boundary (d=3 m) and positioned the target within the drone FoV (Fig. 7).

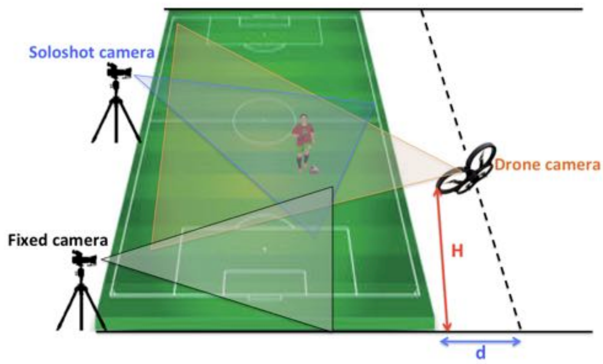

Fig. 7: Field view

CuraCopter uses the 720p (1280 x 720 pixels) image resolutions of the AR.Drone 2 for the tracking processes at 15 fps. Once the drone takes off and hovers, from the initial altitude, the system checks whether or not the target is in the FoV of the drone. If the target is not found, the system reduces the drone altitude (also known as throttle value), sends the video frames, and checks again. When the target is found, the system maintains the drone altitude, sends the video frames, and determines its yaw angle by computing the spinning amount, only when the target moves. Whenever the target moves, Curacopter calculates the drone-spinning amount by these formulas:

$$SpinningamountX = \frac{(-(ctX - imgX))}{imgX} \quad (1)$$

$$SpinningamountY = \frac{(-(ctY - imgY))}{imgY} \quad (2)$$

$$\sum yaw < 180^\circ \quad (3)$$

Where ctX and ctY are the x and y coordinates of the center of the detected target obtained from the vision processings, $imgX$ and $imgY$ are the size (width and height) of the drone image. These formulas move the drone in a way that the target is kept in the center of the FoV of the drone. Limiting the spinning angle to 180° keeps the drone recording within the game field boundaries. Since the drone spin and throttle speeds are limited to a particular range, there may be situations when the drone movements are too slow to follow a fast moving athlete. To solve this problem, we implemented a lost and found strategy as follows: when a previously found target is lost, the drone spins approximately 20° for 5 seconds, in the same angle orientation (clockwise or counter-clockwise) and with the same speed it was moving before losing the target. If the lost-search yields no results the drone checks in the opposite orientation from where the target was lost within the game field boundaries.

3.2 Video curation module configuration

The video curation module consists of 4 steps.
Step 1: Making curation rules
Step 2: Collection and processing of multiple video streams
Step 3: Assignment of metadata to each of video segments
Step 4: Automatic switching among cameras

These steps are described in the following subsections. In this session, f1 represents the video streams from the drone camera, f2 are the streams of the Soloshot2 and f3 are from the fixed camera.

3.2.1 Making curating rules

Step 1 consists of making the curating rules. For example in the Rule4 of Figure 4, the module selects f1 as a default. But, if f1's metadata is labeled as "player=1, overlook degree=low" but f2's metadata is labeled "player=1, overlook degree=high", the module selects f2 as next interval video. In the same way, we make a set of rules in order for the module to determine next interval video to be selected.

3.2.2 Collection and processing of multiple video streams

In Step 2, the module collects three video streams and the collected streams are divided into segments with a fixed interval (e.g., t = 3 seconds).

3.2.3 Assignment of the metadata to the video segments

In Step 3, metadata is assigned to each segment. In this paper, we assume that the metadata is manually assigned. However, it can be assigned automatically by using a special camera [12], a special wearable device [13], or using existing annotation techniques [14][15]. The proposed system will support the above automatic metadata assignment in the future. The candidate labels of metadata are the following 2 labels: *player, overlook degree*. If the player appears in a fixed interval video, we give a value of "1" for the label. Otherwise, give a value of "0" for the label. For the overlook degrees: "low" means that the face or the upper body of a player is shown, "middle" means that the whole body of a player is shown, "high" means that multiple players are scattered on the screen. One of the two label (Player or overlook degree) is labeled manually.

3.2.4 Automatic switching among multiple cameras

In Step 4, the module selects the next interval video by inputting the metadata (Step 3) assigned to the current segments captured by all cameras by following the set of rules (constructed in Step 1). By selecting the next interval video every point of time, it becomes possible to switch to the camera at each point of time and thus conduct real-time automatic video curation. Figure 8 describes the example of automatic switching among three cameras (Rule4 in Figure 4).

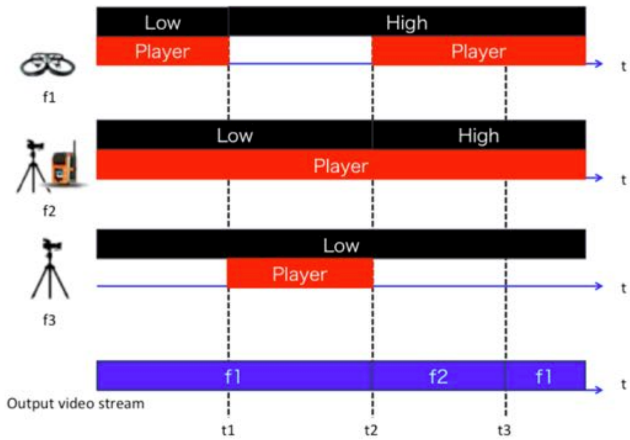

Fig. 8: Example of automatic switching among three cameras (Rule4)

4. Conclusion

In this paper, we proposed Curacopter, a drone-based, self-automated athlete tracking and video curating system. We developed a prototype system that uses vision processing on pre-defined features to control the drone, and a video-curating module to manage the changes between cameras. For future work, we will improve our current implementation and conduct a field evaluation in real-life training environments.

Acknowledgments This work was partly supported by CICP (Creative and International Competitiveness Project) in NAIST (Nara Institute of Science and Technology). We are deeply grateful.

References

[1] Mueller, F.F. and Muirhead, M. 2015. "Jogging with a Quadcopter," In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15), 2023-2032.
[2] Jessica R.Cauchard, Jane L.E, Kevin Y.Zhai and James A.Landay.: "Drone and Me: An Exploration Into Natural Human-Drone Interaction," 17th ACM International Conference on Ubiquitous Computing (UBICOMP2015).
[3] Wilson, B. D. (2008), "Development in video technology for coaching," Sports Technol., 1: 3440. doi: 10.1002/jst.9
[4] Keita Higuchi, Yoshio Ishiguro and Jun Rekimoto.: "Flying Sports Assistant: External Visual Imagery Representation for Sports Training," Proceedings of the 2nd Augmented Human International Conference, AH 2011.
[5] Keita Higuchi, Yoshio Ishiguro and Jun Rekimoto.: "Flying Eyes:

Free-Space Content Creation Using Autonomous Aerial Vehicles," ACM CHI 2011.

[6] Kazuya Yonezawa and Takefumi Ogawa.: "Flying Robot Manipulation System Using a Virtual Plane," Proc. of IEEE Virtual Reality Conference (VR2015), 2015.

[7] Lily, "https://www.lily.camera/," Oct.2015

[8] Soloshot2, "http://shop.soloshot.com/," Oct.2015

[9] Tang Sze Ling, Liang Kim Meng, Lim Mei Kuan, Zulaikha Kadim and Ahmed A. Baha'a Al-Deen, "Colour-based Object Tracking in Surveillance Application" in Proceedings of the International Multi-Conference of Engineers and Computer Scientists 2009 Vol I IMECS 2009, March 18-20, 2009, Hong Kong.

[10] T. Fitzgibbons and E. Nebot "Bearing only SLAM using colour-based feature tracking", 2002 Aust. Conf. Robot. Autom., 2002

[11] Kazuki Fujisawa, Yuko Hirabe, Hirohiko Suwa, Yutaka Arakawa, Keiichi Yasumoto.: "Automatic Content Curation System for Multiple Live Sport Video Streams," International Workshop on MIPR Colocated with Multimedia (ISM), 2015 IEEE International Symposium on.

[12] Ryohei Suzuki, Daisuke Sakamoto, and Takeo Igarashi.: "Annotone: Record-time audio watermarking for contextaware video editing," In CHI ' 15 Extended Abstracts on Human Factors in Computing Systems, pp. 57-66. ACM, 2015.

[13] Thomas Plötz, Chen Chen, Nils Y Hammerla, and Gregory D Abowd.: "Automatic Synchronization of Wearable Sensors and Video-Cameras for Ground Truth Annotation–A Practical Approach," Wearable Computers (ISWC), 2012 16th International Symposium on , pp.100-103, 2012.

[14] Takuro Yonezawa, Masaki Ogawa, Yutaro Kyono, Hiroki Nozaki, Jin Nakazawa, Osamu Nakamura, and Hideyuki Tokuda.: "Sensetream: enhancing online live experience with sensor-federated video stream using animated two-dimensional code," In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp.301-305. ACM, 2014.

[15] Risto Sarvas, Erick Herrarte, Anita Wilhelm, and Marc Davis.: "Metadata creation system for mobile images," In Proceedings of the 2nd international conference on Mobile systems, applications, and services, pp.36-48. ACM, 2004.