Inexpensive Colour Tracking to Overcome Performer ID Loss

Bob Pritchard
UBC School of Music
6361 Memorial Road
Vancouver, Canada
bob@mail.ubc.ca

Ian Lavery
Algorhythmic Software Design
308 - 988 16th Ave W.
Vancouver, Canada
ian_lavery@hotmail.com

ABSTRACT

The NuiTrack IDE supports writing code for an active infrared camera to track up to six bodies, with up to 25 target points on each person. The system automatically assigns IDs to performers/users as they enter the tracking area, but when occlusion of a performer occurs, or when a user exits and then re-enters the tracking area, upon rediscovery of the user the system generates a new tracking ID. Because of this any assigned and registered target tracking points for specific users are lost, as are the linked abilities of that performer to control media based on their movements. We describe a single camera system for overcoming this problem by assigning IDs based on the colours worn by the performers, and then using the colour tracking for updating and confirming identification when the performer reappears after occlusion or upon re-entry. A video link is supplied showing the system used for an interactive dance work with four dancers controlling individual audio tracks.

Author Keywords

Colour tracking, Kinect, Orbbec, NuiTrack

CCS Concepts

•Applied computing~Arts and humanities~Sound and music computing•Applied computing~Arts and humanities~Performing arts•Applied computing~Arts and humanities~Media arts

1. INTRODUCTION

Video tracking of moving athletes within a constrained playing area is common, with various commercial and research systems available, such as [1][3][5][6][7][8]. The identification of the individuals tracked can be important for analysis of performance effort and abilities, for pedagogical support in improving performance decisions, or for making strategic and tactical decisions. Many of these systems use on-body tracking units [10][17] or identify athletes optically through their jersey numbers [14][19].

In interactive performance systems and consumer level gaming, optical tracking is often used [16][9]. The ability to maintain identities when there are multiple users is important, and while it is possible to place markers or transmitters on individuals to assist in such tracking, this may not be desirable in performance if it interferes with costuming or with contact between performers. Some systems capable of tracking moving subjects across a wide area use multiple cameras, enabling constant registration of subjects from one or more angles [15].

However, for many dance groups such systems are prohibitively expensive, and may be beyond the technical abilities of support staff. Single camera active infrared systems usually operate by auto-assigning a unique tracking ID to each subject as they enter the tracking area and then continuously generate coordinate data for that particular ID. A problem arises as the number of performers being tracked increases, along with the amount of interaction in the tracking area. As a result, the ability of the system to maintain accurate IDs degrades.

2. THE SYSTEM

2.1 KiCASS 2016+

The Kinect-Controlled Artistic Sensing System (KiCASS) system was first developed as an undergraduate Engineering project in 2016 and then upgraded in following years. Using a Kinect-for-Windows© camera, a Surface Pro II©, a router, and custom written software, it is capable of tracking up to six performers and 25 target points on each but has generally been used to track one or two dancers or musicians. In this configuration it has been part of dozens of studio and public performances and gallery installations, such as [4][16].

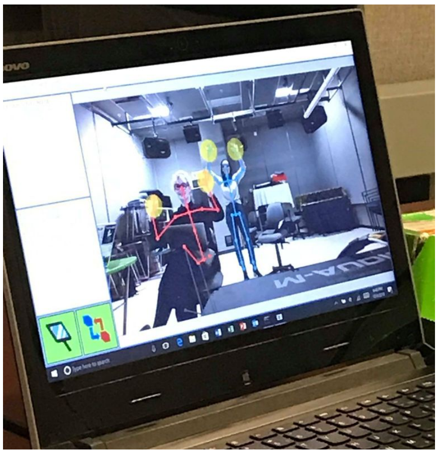

Figure 1. Two subjects tracked with KiCASS

While the system is quite successful, when a performer is briefly occluded by another performer or object, or when they exit and re-enter the tracking area, (in other words, when the system loses and rediscovers them) it assigns the subject a new tracking ID. As a result, the pre-registered target points on the subject are lost, and the subject loses the ability to generate data from the movement of the target points and thereby control media using the resulting data.

The desire to confidently track more than two dancers in performance and to maintain tracking IDs through the exiting and re-entering of performers in the tracking area led us to investigate colour tracking as a possible means to achieve that goal.

NIME'20, July 21-25, 2020, Royal Birmingham Conservatoire, Birmingham City University, Birmingham, United Kingdom.

2.2 OrCASS

At the time that this investigation commenced, Microsoft had not announced the release of the Verge© camera, so in an effort to move forward we switched from Kinect© to Orbbec Pro© cameras, applying the same name logic to the system to arrive at OrCASS. The new software was written using the NuiTrack SDK, which supports the development of tracking software for Kinect©, Orbbec©, LIPS©, Intel RealSense©, and Asus© units. (At some point our software names of KiCASS and OrCASS may be neutralized to show that the software can handle different optical tracking hardware.) We continued to make use of a single camera in an effort to keep costs down, and to develop a simple system that is useable by non-technical users.

To overcome KiCASS's loss of tracking IDs in performance, we rewrote the software to use the incoming data from the Orbbec© to generate tracking IDs based on colour, and to override the autogenerated tracking IDs. As a result, with the new software each dancer is required to wear a significantly different colour of costume, with the system's colour detection focussing on the parallelogram formed by the shoulders and hips of each performer. Due to this focus, after some testing with coloured tee-shirts it was found advantageous to have dancers wear leotards with long sleeves that were the same colour as the torso. Thus, when a dancer turned sideways to the camera, a coloured parallelogram was still available to be tracked. It was found when using tee-shirts that the skin colour of the dancer's arms interrupted the colour of the parallelogram, increasing the tracking difficulty for the system.

3. REGISTRATION OF COLOURS

3.1 Pre-performance

The pre-performance preparation begins with the registration of the colour of each dancer's leotard, with the dancer using a posture reminiscent of the original Kinect ©start pose. ID numbers are assigned to each dancer as they register, beginning with ID 1 and ascending from there. For the colour registration we take a 6-frame analysis of each dancer's dominant hue, using as the area to be measured the quadrilateral created by the joint-tracked shoulders and hip and NuiTrack's user mask. We convert the image from RGB to HSV (Hue Saturation Value), using the Open C library Emgu CV, and also use the library to create a histogram of all hue values in the torso region. The colour value that occurred most frequently (taken as the bin with the highest number of pixel occurrences) is then selected to be dominant hue of the torso. This colour is used as the model for the dancer and is used as the model -- the colour the system expects to see when the performer has been lost and then reacquired during tracking). During the colour registration – using the same Max/MSP© GUI – we select the target points to be tracked on each dancer. Each client is able to connect to the server and sends a configuration map of the target points to be tracked. The server then sends only the requested data from those points to the client, reducing the communication bandwith with all clients. If necessary each client can change or update their configuration map in real time, allowing for different target points to be tracked at different times in a performance, supporting different gestures and movement in the choreography.

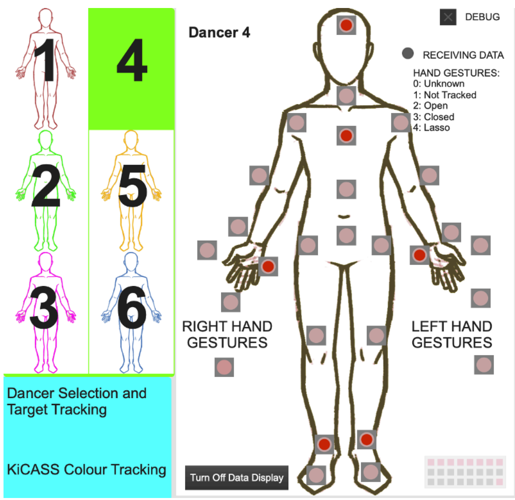

Figure 2. GUI for colour registration and target selection

3.2 Performance

All of the tracking logic for the OrCASS system is handled by the C# wrapper for the NuiTrack SDK. In performance the system checks incoming HSV values from each performer every 500 msec, and measures how far the results are from the registered colours.

For instance, suppose Performer 1 is registered as being pink (a value of 233). In performance Performer 1 is measured as having a value of 224, a difference of 9 from the originally registered colour. From this a Confidence Value is generated, where the Confidence Value = 100 – distance-from-match. A Confidence Value of 0 indicates that the performer has been lost, occluded, or has exited the tracking area. For this example the Confidence Value would be 100 – 9 = 91.

When a missing dancer is rediscovered in the tracking area, the software determines the dominant colour using the same 6-frame analysis procedure used to create HSV data. That data is then compared to the registered values, and the colour with the shortest distance to a registered value is chosen as the matching colour.

3.3 Optimization

A constant optimization is run on the matches once they are made. Every 500 msec each performer is reanalyzed for colour. If a performer receives a match with a registered colour that has a shorter distance than their current match, (in other words, a better colour match than the one currently assigned) the colour matches are swapped between the two users identified with the two colours. If the performer gets the same match, the match distance is updated to be the average of the previous distance and the new distance. If the performer gets a new match, but it isn't a closer match than the current assigned one, the system adds 1 to the distance of the match: this is to prevent ruts in the processes, where the system would go into oscillation checking the two performers.

With all performers recognized in the tracking area, once a performer has received three confirmed matches with a distance of 0, the system locks their match. This means that the system will stop trying to confirm that match and no other user can take their match. However, once a change in the Confidence Value for a performer falls below a threshold set in the software (i.e., the performer is "lost"), the system will automatically run the optimization routines.

4. PERFORMANCE

In performance, the tracking data and Confidence Values are sent out over Wi-Fi in OSC format. The dancer ID and the X Y Z coordinate values of each tracked target point are sent as one message, and a separate message sends the dancer ID and a Confidence Value. Max/MSP© (and Pd) GUIs receive the data and assign it to control different performance parameters, but since the data is OSC formatted many other means could be used to accept and use the data.

At the users end the Confidence Value can be tracked to set thresholds for accepting or rejecting incoming OSC data. If the system is hunting for a particular colour or is oscillating between two dancers for assigning a colour, the Confidence Value will be low, and the user may not want to accept data that will jump around during this process. In such cases the user can elect to ignore the low confidence data, treating the dancer as currently being "lost" or offscreen.

The depth of the tracking area with OrCASS has been extended to 8 metres (Z) with a width at that distance of about 6 metres (X). To date height (Y) has not been a concern since no dancers have been elevated when at the far limits of the X and Z tracking. Current work and studio testing seem to indicate that colour tracking remains more accurate than joint/skeletal tracking at the outer limits of the tracking space, but confirmation of this requires more work and testing. This may have future implications for choreography, in which target tracking may not be as sensitive or as useful as colour tracking in more distant parts of the performance area.

5. OUTCOMES

The OrCASS system proved to be mostly successful in its ability to consistently track four dancers moving in complex patterns on a professional music stage with standard concert (non-theatrical) lighting.

In the choreographed work *Cheap Loops* [11], a high tempo rhythmic bed track ran constantly. Each dancer's Z position on stage controlled the volume of her own rhythm tracks, with all tracks synchronized in tempo. The closer a dancer was to the Orbbec, the louder her track would be. Occlusion or exiting of the tracking area would cause her track to quickly fade out, while re-discovery or entry into the tracking area caused the track to quickly fade in. Each dancer could control four tracks of increasing complexing, which were gradually introduced during the performance, or during solo or duet sections.

In tests and in the performance work, choreography was created to specifically test the system's ability to deal with constant occlusion of different dancers, to deal with reacquisition of dancers, and to handle the tracking of different dancers at various depths and widths. In the choreography no tracking of target points on limbs was carried out: colour tracking was the only concern.

Figure 2. Performance of *Cheap Loops*

6. DIFFICULTIES

One difficulty experienced in the initial testing involved the choice of material for the dancers' costumes. Originally it had been decided to dress the dancers in very bright fluorescent-coloured leotards. Under the lighting of the testing space (a sound studio) and the performance space (a music concert stage) a glare or washout of the colours was discovered, due to the sheen on the synthetic fabrics. This resulted in low Confidence Values from the software, even when no occlusion or loss of subject had occurred. To correct this, leotards of more contrasting colours were substituted for a few of the dancers.

Another difficulty was the quality of lighting in the studio and on stage. The rehearsal studio had both fluorescent work lights and dimmable incandescent pot lights. As such it was possible to vary the lighting balance somewhat. In the studio testing colour tracking errors hovered around 10% but were quickly corrected within a second or two. Things were slightly different when we moved to the concert stage for the dress rehearsal.

Since the stage was designed with music performance in mind, it was set up for a general wash, with a few side spots. As such, it was difficult to achieve consistent light levels across the entire stage, even within the restricted tracking area. Additionally, when dancers were close together, shadows could be cast by a dancer onto another's leotard, changing the perceived colour values. These differences in lighting caused difficulties for the tracking software, which necessitated reducing the sensitivity of the colour registration to allow for a greater variance in the HSV data. The problems with lighting on the leotards still meant that depending on the lighting and a dancers' posture/presentation, the system might require two or three seconds to properly register a specific dancer during performance. At times in our testing during the dress rehearsal we estimated that under certain lighting and shadow conditions the error rate could rise as high as 25%. This resulted in a slower response to triggering and amplification/ attenuation of a dancer's rhythm track. However, even with this difficulty and the lowered sensitivity, with some lighting adjustments the software was able to provide corrections and re-acquire and make correct matches.

Difficulty was also found in tracking the dancers who were to the side of the tracking area. This was exacerbated as the dancers moved forward, as the tracking area narrows significantly, causing the loss of the dancer if she remains to the side while moving forward, or else significantly lowering the Confidence Value as the quadrilateral parallelogram for colour tracking is gradually lost.

7. CONCLUSIONS

The OrCASS system proved to be capable of tracking multiple dancers using colour information, even in lighting environments that were not of the highest quality nor consistent in brightness. Its overall low error rate means that it is capable of being used in a variety of performance environments. With hardware consisting of a single Orbbec Pro, a Windows computer, and a router, it is an inexpensive option for performance groups that wish to include interactive/ responsive systems in their performance repertoire.

8. FUTURE WORK

Since the initial use of OrCASS was only for colour tracking of four performers we are now focussing on fine tuning the interaction of the physical lighting and the colour tracking algorithms, and extending the sensitivity of the tracking at the current distance limits. Once that has been resolved, the next stages involve tracking six dancers with various types of choreography and costumes in order to test the system even more, while maintaining artistic interest. After that we will combine the existing tracking of limb target points with the colour tracking to see how well the system works dealing with the greater demands. We will also combine the OrCASS system with other tracking research of the P.I. [2][13] increasing both the complexity of the system but also the performance possibilities.

9. ACKNOWLEDGMENTS

We acknowledge that University of British Columbia is located on the traditional, ancestral, and unceded territory of the hən̓q̓əmin̓əm̓ speaking xʷməθkʷəy̓əm (Musqueam) people.

10. ETHICAL STANDARDS

This research is supported by the Social Sciences and Humanities Research Council of Canada (SSHRC), and by the University of British Columbia.

11. REFERENCES

[1] H. Ben Shitrit, J. Berclaz, F. Fleuret, and P. Fua, "Tracking Multiple People under Global Appearance Constraints," in ICCV, 2011.

[2] K. Bhumber, R. Pritchard, and K. Rodé. "A Responsive User Body Suit (RUBS)". Proceedings of the International Conference

on New Interfaces for Musical Expression 2017, May 15 - 18, Copenhagen, Denmark

[3] "ChyronHego announces TRACAB Gen5", https://chyronhego.com/chyronhego-introduces-tracab-gen5/ retrieved Jan. 29, 2020

[4] M. Ducharme, E. Fredriksson, D. Lee, C. Guardiola, and R. Pritchard, *The Spaces In Between*, for dancer, bass clarinet, DMX lighting, and TaSTE system. Premiere Dec. 12, 2018, WinterBang! Festival, Vancouver, BC.

[5] M.-C. Hu, M.-H. Chang, J.-L. Wu, and L. Chi, "Robust Camera Calibration and Player Tracking in Broadcast Basketball Video," IEEE Transactions on Multimedia, vol. 13, no. 2, pp. 266–279, 2011.

[6] IVCNZ '14: Proceedings of the 29th International Conference on Image and Vision Computing New Zealand, November 2014 Pages 13–18 https://doi.org/10.1145/2683405.2683422 Retrieved Jan. 13, 2020

[7] J. Liu, X. Tong, W. Li, T. Wang, Y. Zhang, and H. Wang, "Automatic player detection, labeling and tracking in broadcast soccer video," Pattern Recognition Letters, vol. 30, pp. 103–113, 2009.

[8] W.-L. Lu, J.-A. Ting, K. P. Murphy, and J. J. Little, "Identifying Players in Broadcast Sports Videos using Conditional Random Fields," in CVPR, 2011.

[9] Microsoft: Kinect for Windows SDK 2.0: https://developer.microsoft.com/en-us/windows/kinect. Retrieved Jan. 29, 2020

[10] S. Milici, A. Esposito, and E. Staderini. "Evaluating Athletic Performances with a Real Time Location and Tracking System." Materials Science Forum. 879. 2352-2354, 2016 10.4028/www.scientific.net/MSF.879.2352.

[11] R. Pritchard, E. Fredriksson, I. Lavery, *Cheap Loops*, for four dancers and OrCASS tracking system. Premiere Dec. 11, 2019, WinterBang! Festival, Vancouver, BC. https://youtu.be/EpLqE7oQSpA

[12] R. Pritchard."T(w)o Nearly Touch: You" for two dancers, RUBS interfaces, and sound clips. June 3, 2018, NIME 2018, (June 3 – 6), Blacksburg, Virginia.

[13] R. Pritchard."Tracking And Smart Textiles Environment (TASTE)". International Computer Music Conference 2019, June 16 – 23, New York, New York, New York University.

[14] M. Saric, H. Dujmic, V. Papic, and N. Rozic, "Player Number Localization and Recognition in Soccer Video using HSV Color Space and Internal Contours," in ICSIP, 2008.

[15] G. Thomas, R. Gade, T.B. Moeslund, P. Carr, and A. Hilton. "Computer vision for sports: Current applications and research topics." In *Computer Vision and Image Understanding*, Vol. 159, June 2017, pp. 3-18.

[16] B. Topp, E. Fredriksson, S. Wasik, E. Richardson, and R. Pritchard, *...to travel the distance of a changing feeling...*, for dancer, flute, video, and TaSTE system. Premiere Dec. 12, 2018, WinterBang! Festival, Vancouver, BC

[17] Troika Ranch. https://troikaranch.org. Retrieved Jan 29, 2020.

[18] Trace tracking: https://www.traceup.com/soccer/how-it-works. Retrieved Jan. 13, 2020.

[19] Q. Ye, Q. Huang, S. Jiang, Y. Liu, and W. Gao, "Jersey number detection in sports video for athlete identification," in SPIE, 2005.