

カテゴリ1 東京2020を『観せた』NTT R&Dの技術

バドミントン競技×超高臨場感通信技術 Kirari!

NTTは、東京2020組織委員会が主導する「TOKYO 2020 未来のスポーツ観戦プロジェクト」において、超高臨場感通信技術 Kirari!の技術協力を行いました。本稿では、武蔵野の森総合スポーツプラザで行われたバドミントン競技の様子を遠隔会場（日本科学未来館）へホログラフィックに映像伝送することで、あたかもその場にいるかのような臨場感と一体感を提供することをめざした次世代臨場感テクノロジー実証について紹介します。

むとう	まこと	はせがわ	けいすけ
武藤 誠	/	長谷川 馨亮	
なみかわ	だいち	こんや	せいいち
並河 大地	/	紺谷 精一	
ひらち	のぶひろ	なかわら	たいじ
平地 信博	/	中村 泰治	
すずき	けんや	きのした	しんご
鈴木 健也	/	木下 真吾	

NTT人間情報研究所

プロジェクトの概要

ライブの魅力は、実際の会場に行ったときに感じる臨場感や一体感を体験できることです。これは、スポーツや音楽ライブなど、アスリートやアーティストが平面の中の像としてではなく、目の前にいること・同じ空間を共有していること、この感覚を持てることと言い換えることができます。オンライン配信・ライブビューイングでは、画面のサイズや解像度などが飛躍的に向上してきましたが、依然として平面の映像しか送ることができません。そのため、アスリートのその一瞬にかけける緊張感、身体の力強さと美しさ、観客とアスリートの一体感などの臨場感を十分に伝えきれてはいません。本プロジェクトでは、バドミントン競技を対象に、超高臨場感通信技術 Kirari!を用いてホログラフィックな映像伝送を行うことで、競技会場に行けない人にも、あたかも会場にいるかのような体験を届けることを目的とした技術実証を行いました（図1）。

当初、一般参加者を招いてのライブビューイング実証を計画していましたが、新型コロナウイルス感染予防の観点から中止し、2021年7月30～31日、日本科学未来館において、メディア向けに「ス

ポーツ観戦の未来～次世代臨場感テクノロジー実証プログラム～」を公開しました。

システム構成

本技術実証における全体のシステム構成を図2に示します。東京2020オリンピックのバドミントン競技会場である武蔵野の森総合スポーツプラザに、8Kカメラを設置し（図3）、撮影したバドミントンの試合映像を、1 Gbit/sのネットワークで中継拠点に送ります。中継拠

点では、Kirari!の要素技術である「任意背景リアルタイム被写体抽出技術」を用いて、送られてきた8K映像から選手やシャトルの映像のみを抽出します（図4）。抽出映像と、OBS（Olympic Broadcasting Services）提供の映像や音声など複数の映像を、「超高臨場感メディア同期技術（Advanced MMT）」⁽¹⁾を用いて、遠隔会場である日本科学未来館へ同期伝送します。

遠隔会場では、ホログラフィック投影装置を備えた実物大のコートと約100席



図1 東京2020バドミントン競技のKirari!技術実証の様子

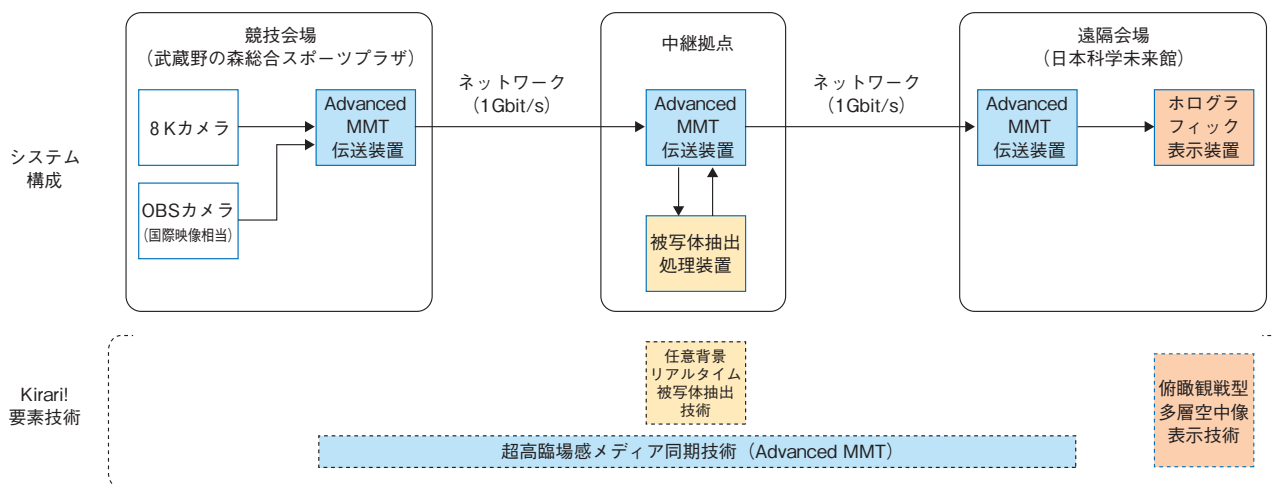


図2 全体のシステム構成



図3 競技会場に設置した8Kカメラ



図4 中継拠点の様子

の観客席（図5）を設置し、競技会場そのものを再現しました。伝送された選手とシャトルの抽出映像は、「俯瞰観戦型多層空中像表示技術」により、手前の選手はネットの手前に、奥の選手は奥に、ホログラフィックに表示されます。その結果、リアルなコートに2人の選手が降り立ったかのような空間を実現しました。

Kirari!の技術

■任意背景リアルタイム被写体抽出技術

「任意背景リアルタイム被写体抽出技術」は競技の撮影映像から、選手やシャトルなどの被写体映像のみを抽出する技術です⁽²⁾（図6(a)）。

通常、映像の中から被写体のみを抽出する場合、グリーンバックやブルーバックなどの背景を用意し、クロマキーによりその背景色を消す手法が取られます。本技術は、特別な背景環境を用意することなく、競技会場のそのままの映像からリアルタイムに被写体のみを抽出することが可能です。

今回、バドミントン競技への適応にあたり、5つの技術改善を行いました。

(1) コートの手前と奥の選手の個別抽出

従来技術では、手前と奥の選手を分けて抽出することができませんでした。個別抽出を行うために、バドミントンコートの空間を模擬した深度情報を入力し、奥と手前の選手について同時に学習・推論できる深層学習モデルを考案しました。これにより、バドミントンのように

コートの手前と奥で選手が分かれる競技においても、個別に選手を抽出することを可能とし、安定的かつ高精度な被写体抽出に成功しています(図7)。

(2) 選手映像の解像度・フレームレート向上

従来技術では、4K・30fpsまでにしか対応していませんでした。4Kカメラでの撮影では、選手像の解像度が最大でも約640ピクセルとなり、空中像として等身大に拡大投影した際に荒さが目立つ

てしまいます。また、30fpsでは、スマッシュ等の速い動きに追従できず途切れ途切れの映像となってしまいます。今回は、処理する映像フレームの多重化や計算リソースの平滑化などの工夫を行い、8K・60fpsカメラ対応を実現しました。その結果、選手の解像度を930ピクセル・毎秒60フレームのスムーズな高精細映像にすることができました。

(3) 微小かつ高速移動するシャトルの安定的な抽出

映像中に映るシャトルは、非常に小さくかつ高速に移動するため、これまでの画像認識手法では、ノイズが多くちらつきが目立ったり、シャトルの軌跡が途切れがちになったりと十分ではありませんでした。今回、シャトル専用の抽出方式を開発し、正確なシャトル位置の検出と、さまざまな形に変形するシャトルの高精度な抽出に成功しました。シャトル位置の検出には、映像に映り込むシャトルと同様の小さな物体(客席のガイド照明など)の影響を排除するため、連続したフレームを畳み込みニューラルネットワーク(CNN)に入力し、シャトルの位置情報および移動情報を同時に学習させる方式を考案しました。CNNから得られ

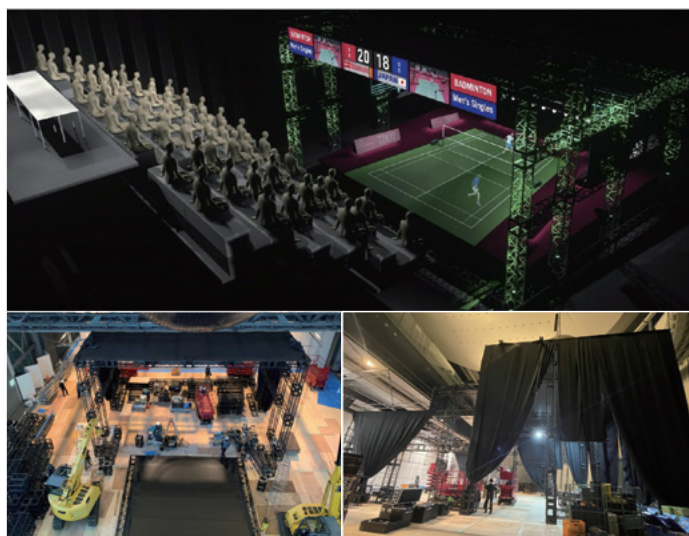


図5 遠隔会場における投影装置、構築の様子

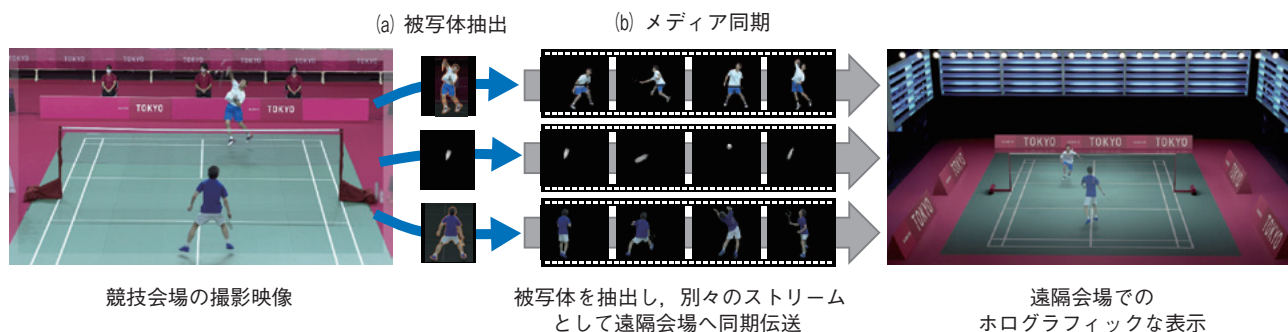


図6 Kirari!の被写体抽出、メディア同期の流れ

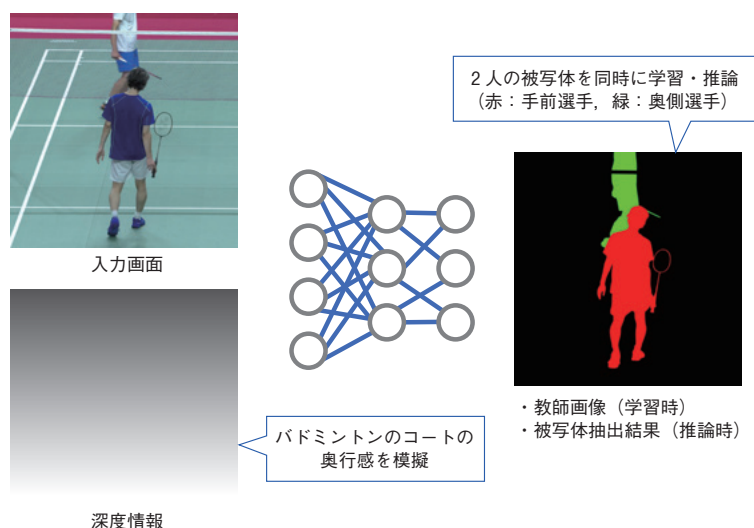


図7 コートの手前と奥の選手の個別抽出の仕組み

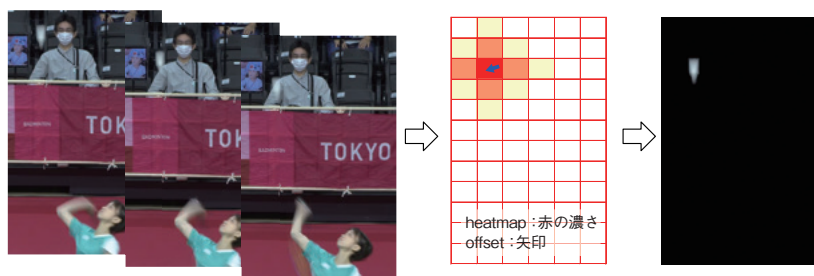


図8 シャトル抽出の仕組み

る大まかな存在位置 (heatmap) と補正値 (offset) を基にシャトル位置を求めます (図8)。シャトルの抽出には、検出されたシャトル位置・形状と、次フレームにおけるシャトル予測情報 (位置・形状・モーションブレイクの程度) でフィルタリングして再現率を高め、背景差分法により生成します。これらにより適合率90.7%、再現率90.3%を達成し、観戦に用いることができるレベルを達成しました。

(4) 選手映像の欠落部分の自動補完

図9のように、奥の選手がネットと重

なった場合、選手像に黒い帯状の欠落が発生します。欠落部分の上下の色情報から、欠落部分の色を推定・補完することにより、欠落のない映像をつくり出します。

(5) 選手の影の自動生成

実際の会場で現れる選手の影を、遠隔会場にも再現することで、より自然な選手像をつくり出すことができます。選手の抽出結果を基に、映像から影を検出しました。この際、ジャンプの高さや競技会場の照明条件から、選手の影が存在し得る範囲を定めることにより高速化を

図っています。

■超高臨場感メディア同期技術

「超高臨場感メディア同期技術 (Advanced MMT)」はメディア伝送規格MMT (MPEG Media Transport) をNTTが独自に拡張した技術で、映像、音声、照明情報といった複数の連続するデータ (ストリーム) を、時刻的な同期を保ったまま伝送できる技術です。今回の取り組みでは、競技会場で撮影した競技映像、音声、競技映像から選手だけを抽出した映像、シャトルだけを抽出した映像、OBSから提供された映像、など複数のストリームを本技術によって同期伝送し、遠隔会場に必要なデータを適切なタイミングで表示することで、臨場感の高い空間表現を実現しました (図6 (b))。

■俯瞰観戦型多層空中像表示技術

あたかも競技会場で観戦しているかのような感覚を体験いただくために、物理的に実会場に似せた空間を構築しました (図10)。具体的には、実物大のコートおよびネットを設置し、そこを見下ろす角度も同じとなるよう観客席の高さも揃えました。そうした空間の中にホログラフィックな選手を実会場と同じ場所に出現させるための技術が、「俯瞰観戦型多層空中像表示技術」です。

ペッパーズゴーストと呼ばれる従来のホログラフィック表示手法では、ディスプレイの映像を斜め45度に設置したハーフミラーに反射させることで空中像を表示させます。この手法では空中像は単一の層のみに表示が制限されます。しかし、単一層の表示では、バドミントンのように、ネットを挟んで、手前と奥の選手を異なる場所 (層) に表示することはできません。

本技術では、実際の競技会場に近い高

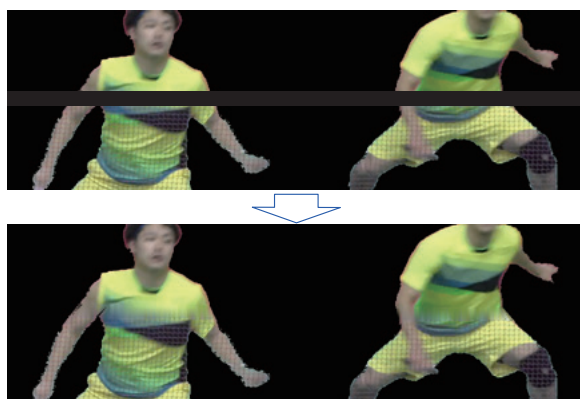


図9 ネットの黒帯状部分の推定

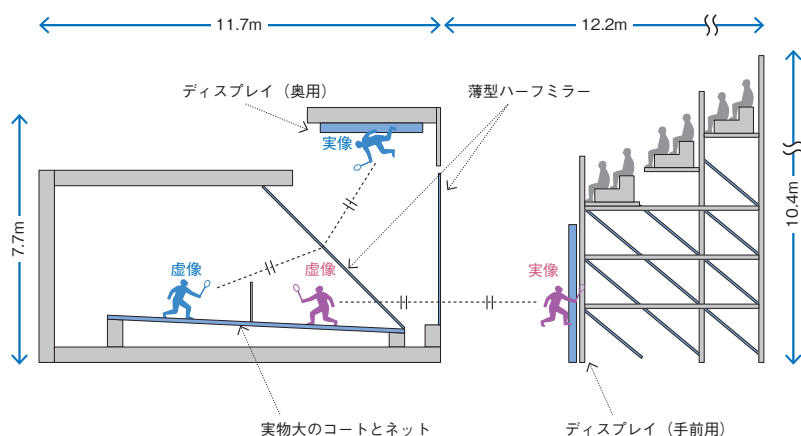


図10 装置のイメージと空中像表示の例

さの観客席を設置し、そこから見下ろす状況下において、手前と奥の選手が、ネットを挟んでリアルに見えるように、コート、ネット、2枚のハーフミラー、LEDディスプレイ・プロジェクタなどの配置を最適化しています。これにより手前と奥の選手は、それぞれ正しい位置でホログラフィックに表示できるようになります。

しかし、本技術では、2つのハーフミラーを用いているため、その間を行き来するシャトルが、不連続となってしまう

課題が発生します。この課題は、2つの層が物理的に分離しているという構造的な問題に起因するため、完全に解決することはできません。より自然に見せるために、シャトルの表示層を切り替えるタイミングの試行錯誤を重ねた結果、選手がシャトルを打った瞬間、別の選手側に切り替えるタイミングが適切であると判断しました。そのタイミング判定には、コート側面映像の解析によるシャトル位置の判定、競技音声からの打球音検出によるシャトル打球タイミングの判定、手

動判定を含む複数の判定方法を統合することで、ダイナミックなシャトル映像表示位置の切り替えを可能としました。

取り組みの結果

本技術実証は、バドミントンの決勝トーナメントが行われる2021年7月30日～8月2日の4日間行いました。予選などの録画映像だけでなく、男子ダブルス、女子シングルス、男子シングルのライブ中継を、報道関係者などに体験いただきました。体験者からは、高い臨場感を体験できたという感想を多く得ることができました（表1）。

実物のコートやネット、そこに映し出された等身大のホログラフィックな選手、こうしたリアリティの高い環境は、いつのまにか目の前に、実際の選手がいるという錯覚を感じさせる効用があるようです。その結果、競技中の躍動感や臨場感はもちろん、試合開始前の選手の緊張感、さらには勝敗が決まったときの喜怒哀楽までも、テレビとは異なるレベルで感情移入しやすくなっていた可能性があります。桃田賢斗選手が第2回戦で敗退したときの崩れ落ちた様子が、本当に桃田選手がそこにいるかのような感覚を受け印象的でした。

被写体抽出の性能は、8K・60fpsで撮影した映像をリアルタイム処理することに成功しました。

選手どうしが重なった場合や、ダブルスなど、一部抽出精度にまだ課題はありましたが、計6回の中継実験において、それぞれ異なるユニフォームを着用した選手に対しても、観戦に問題のある誤抽出や欠けはない処理に成功しました。ネットの黒帯状部分の補完や影の付与についても破綻のないことを確認していま

表 1 体験者からの感想

実物大の大きさで見ると、臨場感が全く違う
シャトルが高く上がったときに、少しシャトルの行方が見つらいときもあるが、まるで選手がその場にワープしてきたかのような、そんな感覚がある
まるで試合会場にいるようなリアル感を味わえる
実物大の選手は迫力があり、試合中、思わず会場にいるときと同じような歓声を上げてしまった

表 2 遅延時間の内訳

撮影系	17 ms以下 (1frame以下)
MMT伝送	1600 ms (4K 59.94p BT.709, エンコードレート40 Mbit/s, FEC5%)
被写体抽出処理	1000 ms (入出力バッファ込み)
表示系	~120 ms (30fps換算で4frames)

す。シャトル表示については、視聴場所によって、不連続さが目立つことも多く、大幅な改善が必要となります。

通信性能としては、競技会場から遠隔会場までのエンドツーエンドの遅延時間は、トータルで2800 ms以下でした。

表 2 にその内訳を示します。MMT 伝送遅延には、ネットワークの伝送遅延である、競技会場～中継拠点の約 1 ms、中継拠点～遠隔会場の約 0.1 ms も含まれます。被写体抽出処理の遅延には、処理時間のゆらぎ等を考慮したバッファも入っていますが、それを除いた抽出処理自体の遅延は 400 ms 以下です。

今回は、競技会場から遠隔会場への一方の映像伝送であったため、3 秒程度の遅延は大きな問題とはなりません。しかし、将来的に目標とする、両会場を双方向につなぎ、観客の声援を戻すような場合では、大きな遅延は問題となるため、さらなる処理時間の削減が課題となります。

これらの性能達成により、遠隔会場において、選手がそこにいるかのような体験を生み出すのに十分な性能を実現できました。

まとめ

今回、「TOKYO 2020 未来のスポーツ観戦プロジェクト」の中で、バドミントン競技へ Kirari! を適用し、あたかも競技会場にいるかのような臨場感を体験できることを実証しました。選手の抽出精度やシャトルの表示方法など課題もありますが、テレビを超える新しいスポーツ観戦の可能性をオリンピックという世界が注目するイベントにおいて、世界に発信できました。

今後は、抽出性能などの技術的改善だけでなく、他の競技や音楽コンサートなど他分野への展開、さらには、競技会場と遠隔会場を双方向につなぐための低遅延化などを検討するとともに、Kirari! をはじめとしたリモートワールドを具現化する研究開発を進め、未来提案へとつなげていきます。

謝 辞

本技術実証の実現に向け、共にさまざまな課題を克服していただいた東京 2020 組織委員会イノベーション推進室の皆様、主催者として先導いただいた日本

科学未来館の皆様、そして、技術的に支えていただいたパートナー企業の皆様に感謝します。

参考文献

- (1) 外村・今中・田中・森住・鈴木：“超高臨場感ライブ体験 (ILE) の標準化活動について,” ITU ジャーナル, Vol.47, No.5, pp.14-17, 2017.
- (2) 柿沼・長尾・宮下・外村・長田・日高：“機械学習を用いた任意背景リアルタイム被写体抽出技術,” NTT 技術 ジャーナル, Vol.30, No.10, pp.16-20, 2018.



(上段左から) 武藤 誠/ 長谷川 馨亮/
並河 大地/ 紺谷 精一
(下段左から) 平地 信博/ 中村 泰治/
鈴木 健也

◆問い合わせ先

NTT サービスイノベーション総合研究所
E-mail svkoho-ml@hco.ntt.co.jp