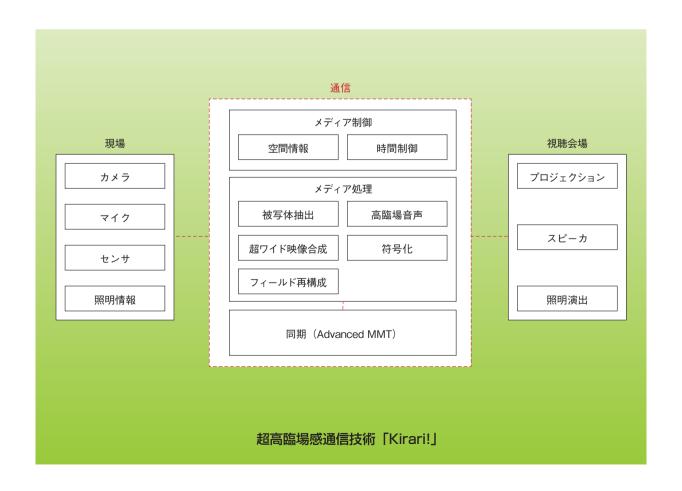
## 特集

# 超高臨場感通信技術

高臨場 機械学習 空中像 だりニアブレンディング VR/AR

本特集ではNTT R&Dが掲げる「新たな感動を生む高臨場UXサービスの創出」のめざす方向性とサービスイメージ、そのサービスを実現するために、NTT研究所で取り組んでいる超高臨場感通信技術Kirarilについて紹介する.



# Kirari!

超高臨場感通信技術Kirari! Be	Vond 2020
. 中国标场级用号权MIN II di I! DE	vuilu Euzu

12

Kirari!の技術概要およびライブビューイングや今までにないような驚きの演出による公演等に加え、Beyond 2020の方向性について紹介する.

#### ■ 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

16

擬似3D映像などを用いて選手の存在感までも遠隔地に提示する際に必須となる被写体抽出技術において、任意背景からリアルタイムに被写体を抽出するシステムを紹介する.

#### ■ 競技空間を取り囲んで観戦する "Kirari! for Arena"

21

簡易な構成で実現可能な、視覚的効果を活用した2D空中像による映像提示、および 全周囲から取り囲んで観察できる表示システムの概要を紹介する.

#### ■ 360度テーブルトップ型裸眼3D映像表示技術

25

3D映像をテーブル上に浮かび上がらせる360度テーブルトップ型裸眼3D映像表示技術の基本となる視覚的な知覚メカニズムとプロジェクタ配置の光学構成を中心に紹介する.

# ■ 「VR&ARを活用したスタジアム観戦」における競技フィールド再構成のための映像処理・提示技術

30

VR(Virtual Reality)とAR(Augmented Reality)で共通的に利用できる映像コンテンツを作成する映像処理技術や,目の前に実物感のある映像を表示するための映像表示技術を紹介する.

#### 主役登場

柿沼 弘員 (NTTサービスエボリューション研究所)

Kirari!における被写体抽出技術の今とこれから

Q.

## 超高臨場感通信技術Kirari! Beyond 2020

NTTグループでは、超高臨場感通信技術Kirari!の研究開発を通じて、「空間の壁を超える」価値創造に貢献していきます。高音質、高画質という進化の枠を超え、あたかも目の前に人がいるかのような超高臨場感をめざします。本稿では、Kirari!の技術概要およびライブビューイングや今までにないような驚きの演出による公演等に加え、Beyond 2020の方向性について紹介します。

### 阿久津 明人 /南 憲一

ひだか こうた†2

#### 日高 浩太

NTTサービスエボリューション研究所 所長<sup>11</sup> NTTサービスエボリューション研究所<sup>12</sup>

#### Kirari!

平成28年1月22日に閣議決定された第5期科学技術基本計画<sup>(1)</sup>では、新たな概念「Society 5.0」が掲げられています.狩猟社会 (1.0)、農耕社会 (2.0)、工業社会 (3.0)、情報社会 (4.0)に続く、人類史上5番目の社会Society 5.0<sup>(2)</sup> (図1)は、科学技術イノベーション総合戦略(平成28年5月24日閣議決定)<sup>(3)</sup>において、次のように定義されています.

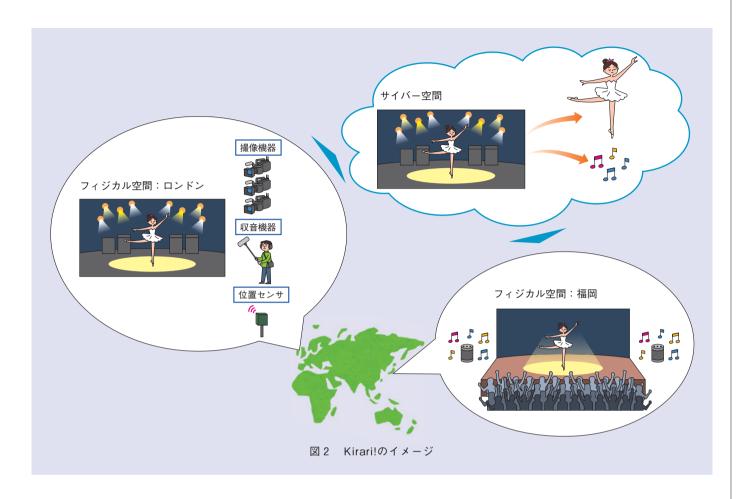
- ・サイバー空間とフィジカル空間を 高度に融合させることにより、
- ・地域、年齢、性別、言語等による 格差なく、多様なニーズ、潜在的 なニーズにきめ細かに対応したモ ノやサービスを提供することで経 済的発展と社会的課題の解決を両 立し、
- ・人々が快適で活力に満ちた質の高 い生活を送ることのできる,人間 中心の社会

これまでも情報通信技術 (ICT) の 進展に伴い,離れた場所の人との音 声・映像コミュニケーションが実現さ れ,より高音質,高画質の通信へと遷 移してきました。また,各種センサの 技術進歩により,人がいるフィジカル 空間そのものをセンシングできるようになってきました。センシング情報をサイバー空間でメディア処理して伝送し、"あたかも目の前に人がいるかのように"別のフィジカル空間に再現させることが、Society 5.0では一般的になるかもしれません(図2)。この

再現を高臨場感と呼ぶこととすれば、 NTTグループが進めるKirari!とは、 通信によるメディア伝送の先をいくテ クノロジで、高音質、高画質という進 化の枠を超えた超高臨場感を実現する ものです。

Kirari!によって、離れた場所の人





や空間を、距離を超えてリアルタイムに伝送することができます。これにより、例えば、スポーツ観戦のために現地に足を運ぶことなく競技を体験することや、公演会場に遠隔地から演者が参画するような演出も実現可能になります。来るSociety 5.0では、「空間の壁を超える」観点で、Kirari!が貢献できると考えています。

Kirari!の利活用例を図3に示します. 観戦が簡単ではない海外でのスポーツ大会を地元でライブビューイングする例では、錯視効果を用いて立体と見まがうような視聴体験や、従来の画角をはるかに凌ぐ視野角いっぱいの視聴体験を提供します。また、公演等へ、今までにないような驚きの演出を提供することも可能です。例えば、実

際の演者をセンシングおよびメディア 処理し、錯視効果による演者を創出し て重畳します. 重畳する時刻を制御す ることで、本来であればあり得ない、 現在の演者と少し過去の演者との共演 が実現します. このような取り組みを 通じて、さまざまなコンテンツを、空 間の壁を超えて超臨場感で体験いただ くことをめざしています.

#### Kirari!の技術概要

Kirari!の技術概要を図4に示します。スポーツや公演をしている現場,通信,視聴会場の3つの観点で考えると,通信で示される領域が超高臨場感通信技術Kirari!になります。カメラ,マイク,センサによって抽出された情報を,ネットワークを介してメディア

制御,メディア処理,同期して視聴会場に伝送します.

メディア制御は、カメラで撮像された人物とセンサによって得られた位置情報、および照明情報を関連付ける空間情報と、人物の配信時刻を絶対時刻で制御するための時間制御から構成されます.

メディア処理は、撮像された画像情報から人物の領域を被写体抽出し、マイクの音響情報から波面合成音響技術(4)等により高臨場音声とします.これらは、図3のスポーツ中継や公演において、錯視効果により人物を立体的に表示させるとともに、音を客席まで飛び出させるような音響再現に必要な処理です。一方、超ワイド合成では、図3の視野角いっぱいの画角での視聴体験

に向け、複数のカメラ画像を合成します<sup>(5)</sup>. また、フィールド再構成において、被写体の構造や状態を解析して3次元的に再構成することを可能にします。さらに、効率的な伝送のために符号化処理を行います。

メディア伝送規格MMT (MPEG Media Transport) に独自の拡張を行ったAdvanced MMTにて同期伝送します<sup>(6)</sup>. 絶対時間で同期できることから,世界中の任意の地点に同時刻に伝送可能となります. Advanced MMT

は、画像、音響、照明情報等を基に、 超高臨場感を実現するための設計図の 役割を担います.

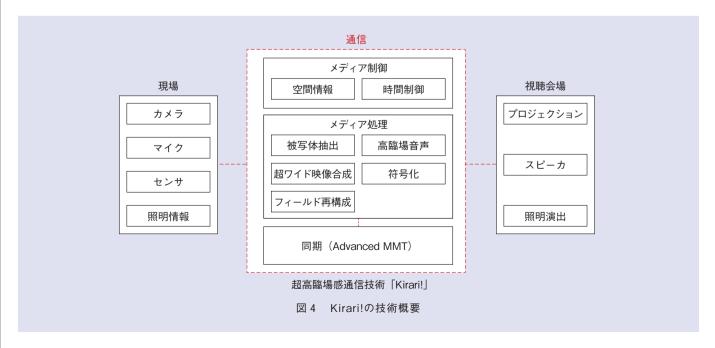
視聴会場では、Advanced MMTの 設計図に従い、プロジェクション、ス ピーカ、照明演出等を施します。

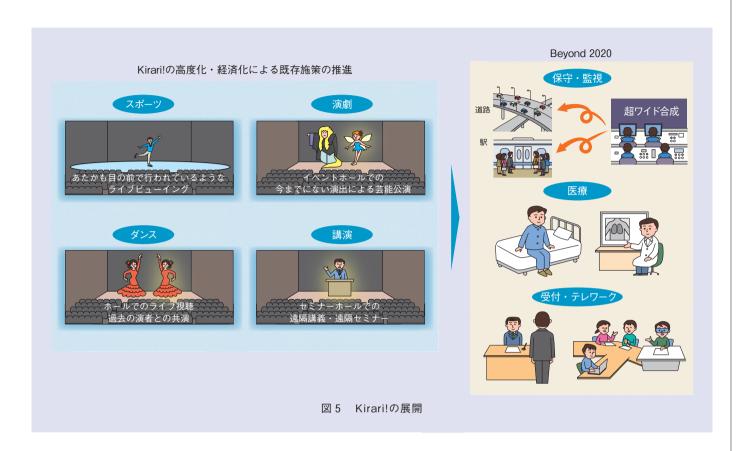
#### 今後の展開

超高臨場感を実現するKirari!を多くのユーザに体験いただくため、あたかも目の前で行われているかのようなスポーツのライブビューイング、今までにないICT演出による芸能公演、過去の演者との共演に焦点をあてたダンス、遠隔地の講演中継等のPoC(Proof of Concept)、トライアル等を実施してきました(図5). これらの施策を推進するためには、Kirari!の高度化として、人が行き交うような環境においても、注目する被写体を高精度に抽出する必要があります。また、より高解像度の映像を超ワイド合成することも課題となります.

一方で、社会実装に向けては、コス







ト削減施策も重要課題です。例えば、 高解像度映像をより圧縮することや、 メディア処理のサーバ数削減について も、検討を進めています。さらに、メ ディア制御が高度化することにより、 過去の演者との共演等において、実物 の演者の隣に、正確に過去の演者を自 動重畳可能となりますので、演出のた めの人的コスト削減にも貢献できます。

Kirari!のスポーツ、公演等への社会実装を促進すると同時に、Beyond 2020を見据える時期にきています。例えば、保守・監視分野に適用すれば、道路や駅等を超ワイド合成映像で監視できます。あたかも目の前で監視対象が存在しているかのような体験につなげることができます。医療分野における医師の遠隔診断、エンタープライズ分野での受付の無人化、テレワークの促進にも検討を広げていきます。今後

もこれらの取り組みを通じて、Society 5.0における「空間の壁を超える」価値創出に貢献していきます.

#### ■参考文献

- http://www8.cao.go.jp/cstp/kihonkeikaku/ index5.html
- ② 日高・長谷川・布施田: "新たな経済社会としてのSociety 5.0を実現するプラットフォーム," オペレーションズ・リサーチ, Vol.61, No.9, pp.551-555, 2016.
- (3) http://www8.cao.go.jp/cstp/sogosenryaku/2016.html
- (4) 堤・高田: "客席まで飛び出す音響を実現する波面合成音響技術," NTT技術ジャーナル, Vol.29, No.10, pp.24-28, 2017.
- (5) 佐藤 難波・小野・菊池・山口・小野: "競技空間全体の高臨場ライブ中継に向けたサラウンド映像合成・同期伝送技術," NTT技術ジャーナル, Vol.29, No.10, pp.19-23, 2017.
- (6) 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.



(左から) 日高 浩太/ 阿久津 明人/ 南 憲一

NTTでは、Society 5.0における「空間の壁を超える」価値創出に向けて、あたかも目の前に人がいるかのような超高臨場感通信技術を、スポーツ、エンタテインメント分野で先行的に社会実装し、これをレバレッジにBeyond 2020を見据えた他分野適用も推進していきます。

#### ◆問い合わせ先

NTTサービスエボリューション研究所 ナチュラルコミュニケーションプロジェクト TEL 046-859-2509 E-mail kota.hidaka.yk@hco.ntt.co.jp 超高臨場感通信技術Kirari! 「ペメージセグメンテーション 機械学習 高臨場感

## 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サービスエボリューション研究所では、遠隔のスポーツ選手があたかも目の前にいるかのように感じさせる超高臨場感通信技術Kirari!の研究開発を進めています。本稿では擬似3D映像などを用いて選手の存在感までも遠隔地に提示する際に必須となる被写体抽出技術において、任意背景からリアルタイムに被写体を抽出するシステムを紹介します。

ながた ひでのぶ ひだか こうた 長田 秀信 /日高 浩太

NTTサービスエボリューション研究所

#### はじめに

映像中から人や物体の正確な領域を 把握することは、高品質な画像編集や 合成を実施するうえで必須の技術のた め. コンピュータビジョンにおける主 要な研究テーマの1つです。 超高臨場 感通信技術Kirari!においても、正確 な被写体領域を抽出することは擬似 3D映像提示を行い、高い臨場感を得 るために必須となります. NTTサー ビスエボリューション研究所では、任 意背景リアルタイム被写体抽出技術<sup>(1)</sup> において、グリーンバックなどのスタ ジオ設備を用いずに試合会場や演技し ている舞台映像からリアルタイムに被 写体の領域のみを抽出するシステムを 提案しています、本稿では、そのシス テムを①従来では判別できなかったよ りわずかな特徴量の差までも判別する 機械学習を導入すること、②赤外線光 などを用い抽出したいオブジェクトの 特徴量を生成すること、でより正確に 被写体領域を抽出できるようにシステ ム開発しましたので紹介します.

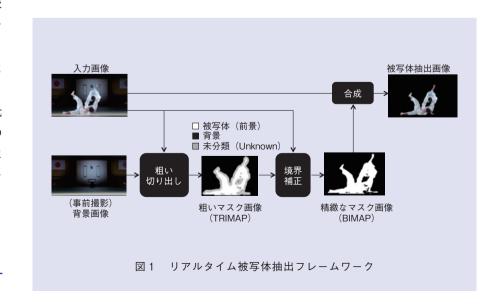
#### 機械学習を用いたリアルタイム 被写体抽出フレームワーク

リアルタイムに対象領域を抽出する

処理には背景差分法が知られています。背景差分法は、入力画像と背景画像との差分をとり、しきい値処理することで変化のあった被写体の領域として抽出するため、高速な処理ができ、事前の準備が少ないことから広く用いられてきました。しかし、適切なしきい値決定は難しく、また、背景の変化などに対応できないなどの問題がありました。そこでNN(Neural Network:ニューラルネットワーク)を用い、入力される特徴ベクトルを別の特徴空間に変換し、識別を行う被写体抽出方法を開発しました。NNを用いることで、識別のための特徴空間はあらかじめ与

えられる教師データによりNN内で決定され、より適切な特徴空間に自動で変換されることが期待されます。また、識別に用いる入力も、対象の画像だけでなく、リファレンスとなる被写体の特徴を有した画像や、時間の異なる画像や、領域情報、赤外線映像などを入れても、NN内で適切に特徴空間に変換され、背景や被写体の高次な特徴量を用いた被写体抽出が実行できるため、背景の変化などにロバストになることが期待されます。

開発システムのワークフローを**図** 1に示します.開発システムは2段 階の被写体抽出を行います.1段目で



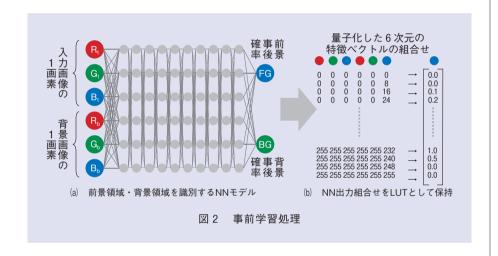
は、被写体領域を粗いマスク画像 (TRIMAP\*1)として抽出し、2段目 ではTRIMAPを頼りにしたマッティン グ処理\*2にて正確な被写体領域を抽出 します、機械学習は1段目のTRIMAP を導出するのに導入しました。

機械学習処理は事前学習処理と適用 処理に大別できます、事前学習処理は、 教師データからNNモデル中のパラ メータを学習させることを目的として 実行されます. 事前学習処理フローを 図2に示します. 事前学習処理では, まず、教師データを用意します.被写 体の含まれない背景画像と被写体の含 まれるサンプル画像をあらかじめ取得 しておき、正解となるマスク画像を作 成します. 次に、作成したマスクの前 景領域に対応するサンプル画像中の注 目画素と、背景画像中の対応画素を組 み合わせたものを入力特徴ベクトルと し、その組合せが前景領域であるとし て学習させます. また同様に、マスク の背景領域に対応するサンプル画像中 の注目画素と、背景画像中の対応画素 の組合せについても、入力特徴ベクト ルが背景領域であるとして学習させま す. このようにして. ある入力注目画 素に対応する背景画素の組合せに対し て、それが前景領域であるか背景領域 であるかを識別するNNモデルを得ま

\*1 TRIMAP: 既知の領域と未知の領域を示した 領域マップ. 既知の前景領域を白, 既知の 背景領域を黒, 未知領域を灰色に設定し ます. す(図 2 (a)). しかし、NN処理は一般に演算量が多いことから、NN演算処理をLUT(Look Up Table: ルックアップテーブル)実装することで高速に処理します. そこで、入力特徴ベクトルを量子化処理によって少ない階調数に削減し、量子化された入力特徴ベクトルのすべてのNN出力組合せを

LUTとして保持します(図 2 (b)). なお、図 2 では分かりやすいようにRGBの各画素について行う処理を記載しましたが、画像位置情報を入れるなど入力ベクトルは色情報に限定されません。

LUTを適用し、TRIMAPを生成する適用処理を図3に示します。適用は





<sup>\*2</sup> マッティング処理:被写体を抽出するアルファマスクを導出する処理.アルファマスクは0から1までの値を持ち、入力画像に掛け合わせて抽出映像が得られます.

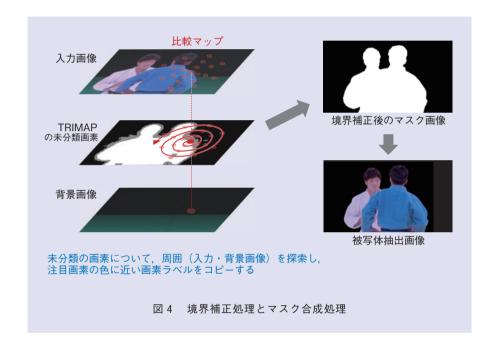
機械学習処理と同様の処理で量子化し た入力特徴ベクトルを導出し、導出し た特徴ベクトルに応じてLUT参照す ることで、注目画素の前景画素である 事後確率を高速に導出します(図3 (b)). 得られた前景事後確率画像から、 前景か背景かがあいまいである領域に 対しては未分類領域と設定することで TRIMAPを生成します (図3 (c)). TRIMAPの未分類領域に対しては. 注目画素と特徴ベクトルの近い周辺の 画素が、前景と背景どちらの領域と識 別されているかの情報を用いて、未分 類領域を識別する境界補正処理を実施 します。境界補正処理の詳細について 図4に示します。未分類領域の注目画 素について距離が近い領域をらせん状 に探索を行い. 注目画素が前景と背景 にどちらが似ているかでアルファ値を 決定しています。導出されたアルファ 値を使い、被写体は抽出されます. こ の境界値補正処理を導入することによ り, 画素情報だけでなく, 周りの画素 の情報を考慮した被写体抽出が実行で きるほか、粗い切り出しは低解像度画 像に対して処理を行い、 境界補正のみ すべての画素に対して実行するなど. 処理の効率化も図れるフレームワーク となっています.

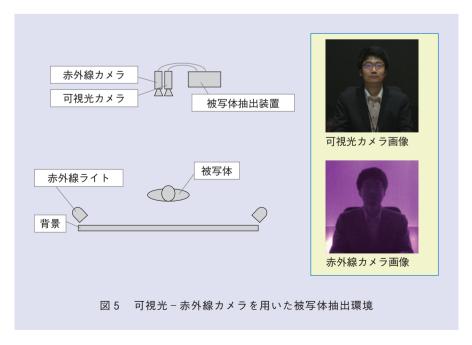
#### 赤外線光を用いた同一色背景から のリアルタイム被写体抽出

機械学習を用いることで、NN内で 自動的に高次の特徴空間に変換してく れるものの、色情報や形状情報を頼り に抽出するため、入力特徴ベクトルが 同じ場合は理論的に分離不可能となります。そこで、新たな特徴を増やす試みとして肉眼で見えない赤外線を利用した可視光 - 赤外線カメラによる被写

体抽出システムを開発しました.

可視光 - 赤外線カメラを用いた被写体抽出撮影環境を図5に示します. 平行に設置された可視光カメラと赤外線





カメラはそれぞれ、可視光のみおよび 赤外線のみの画像を撮影します.赤外 線ライトは背景に赤外線が照射され被 写体には極力照射されないように設置 します.これにより赤外線カメラ画像 において背景は比較的明るく、被写体 は比較的暗く写すことができます.図 5の可視光カメラ画像では、人物と背 景が同色のために色情報を用いた分離 は難しいところ、赤外線カメラ画像で は、人物のシルエットが抽出できていることが確認できます。次に、赤外線カメラ画像を精度良く被写体抽出するため、可視光カメラと赤外線カメラの視差補正を行います。補正は、あらかじめキャリブレーションボードを撮影しておき、赤外線カメラ画像と可視光カメラ画像の同一特徴点が重ね合わせられるように射影変換行列を導出・適用することで補正を行います。こうし

て得られた赤外線カメラ画像を機械学習の入力特徴ベクトルに入れることで,色情報や形状情報ではない新たな特徴を取り込んだ被写体抽出ができます.

#### 評価実験

機械学習を用いたリアルタイム抽出 システムを開発し、今年4月に幕張 メッセにて行われたニコニコ超会議 2018の超歌舞伎「積思花顔競(つも るおもいはなのかおみせ)」の中で使 用しました. 開発システムは3840× 2160画素, フレームレート60 fpsの映 像に適用することができますが、他の システムとの連携のために、本トライ アルの中では1920×1080画素、フレー ムレートは59.94 fpsの映像に適用し ました. 超歌舞伎本編のクライマック スのシーンにて、中村獅童さん演じる 惟喬親王 (これたかしんのう) を背景 が変化する中、舞台映像からリアルタ イムに抽出し、初音ミク演じる小野初 音姫(おののはつねひめ)の対決を盛 り上げることができました. そのとき の映像サンプルを図6に示します. 惟 喬親王の背景にある戸板は人手で持っ ているため揺れ、背景差分法ではうま く抽出できていませんが (図6(b)), 機械学習を用いることにより正確に抽 出できている様子が確認できます(図 6 (c)).

次に、赤外線カメラを用いる効果を確認しました(**図7**). 赤外線カメラの利用により、可視光カメラだけでは抽出が難しい場合も被写体を正確に抽



(a) 入力画像



(b) 背景差分法による被写体抽出画像



(c) 機械学習による被写体抽出画像

図6 機械学習を用いたリアルタイム被写体抽出結果



(a) 可視光カメラ画像

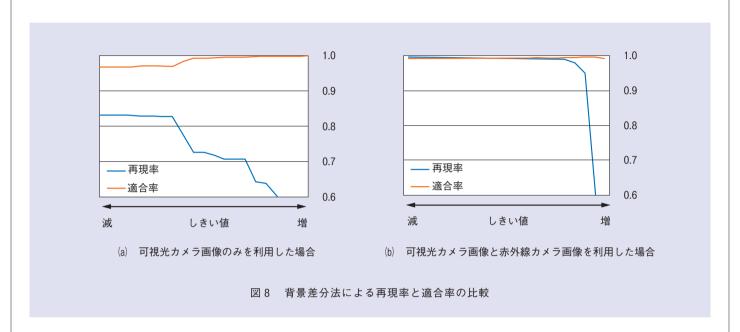


(b) 赤外線カメラ画像



(c) 被写体抽出結果

図7 可視光 - 赤外線カメラを用いた機械学習による被写体抽出結果



出できていることが確認できます.機械学習では、NNの学習の際に自動的に高次特徴空間への変換が行われ、しきい値処理などは発生しませんが、今回、赤外線カメラを用いることでどの程度頑健になるのか、背景差分法のしきい値を変化させることによって評価しました(図8).赤外線カメラを用いることで再現率・適合率がともに高く、また、背景差分法のしきい値を大きく変動させた場合にも安定して動作することが確認できます.

今回は、赤外線カメラを用いた場合を紹介しましたが、私たちのシステムは、ステレオカメラやLiDARから生成される深度マップなど、条件に応じて被写体をうまく抽出できる特徴を入力することもできます.

#### 今後の展開

本稿では、機械学習を用い、入力特

徴ベクトルを新たな特徴空間にNN内で変換・識別することで、被写体の高次の特徴量を基に高精度に抽出する方法、および可視光カメラだけではどうしても抽出することが難しいユースケースに対応するため、肉眼で見えない赤外線カメラを利用した手法について紹介しました。

今後は、セマンティクスを考慮した 被写体抽出を行うため、深層学習を利 用した被写体抽出のリアルタイム化を 進めるとともに、被写体が重なるなど のオクルージョン発生時も精度良く被 写体を抽出する方式の検討を進める予 定です。

#### ■参考文献

長田・宮下・柿沼・山口: "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.29, No.10, pp.33-37, 2017.



(左から) 長田 秀信/ 外村 喜秀/ 宮下 広夢/ 柿沼 弘員/ 日高 浩太/ 長尾 慈郎

見る人・使う人に驚きと感動をもたらすようなサービスやシステムの実現のため、今後も技術をさらに進化させていきます、被写体抽出技術を使うことで、今までになかったような観戦・観劇のスタイルや、SF映画で見たようなコミュニケーションスタイルが実現できると信じています。

#### ◆問い合わせ先

NTTサービスエボリューション研究所 ナチュラルコミュニケーションプロジェクト TEL 046-859-3780 E-mail nagata.hidenobu@lab.ntt.co.jp 超高臨場感通信技術Kirari! 高臨場 奥行き知覚 空中像

## 競技空間を取り囲んで観戦する "Kirari! for Arena"

NTTサービスエボリューション研究所では、競技空間などを取り囲んで観察する高臨場なライブビューイングを実現する表示システムを開発しました。被写体があたかも目の前にいるかのような映像提示を実現するため、人の視覚的な効果を活用することで空間中に被写体を知覚させ、奥行き方向への移動も表現しています。本稿では、簡易な構成で実現可能な視覚的効果を活用した2D空中像による映像提示、および全周囲から取り囲んで観察できる表示システムの概要を紹介します。

い さ か たける まきぐち もとひろ

井阪 建 /巻口 誉宗

た か だ ひであき

高田 英明

NTTサービスエボリューション研究所

#### 高臨場ライブビューイング

街頭TVからはじまりライブビューイングに至るまで、一堂に会した観客がスポーツ中継を楽しむメディアは観客どうしのつながりを重視することで感動を共有してきました。また、2020年には、4K・8K放送の普及により、世界中の多くのスポーツがTV中継されるだけでなく、インターネットによるライブ配信や街中での手軽なライブビューイングといった多様な視聴スタイルによって、一層多くの感動が場所や時間を問わず広く共有されると予想されます。

こうした世の中の視聴スタイルの多様化を踏まえ、NTTサービスエボリューション研究所では、あたかもその場にいるかのような高臨場なライブビューイングをあらゆる場所でリアルタイムに実現するイマーシブテレプレゼンス技術 "Kirari!" の研究開発を推進してきました。

"Kirari!" における映像提示手法については、いくつか研究が進められています。その中でも虚像を活用する2D空中像表示手法は、汎用の2Dディスプレイと入射光の一部を透過・反射するハーフミラーなどの光学素子を組み合

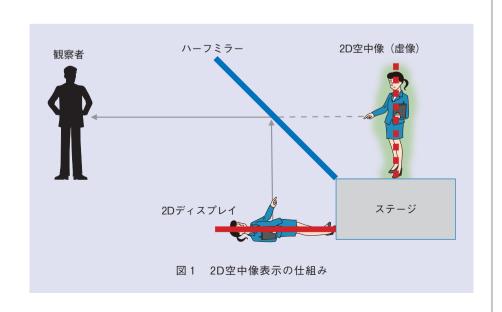
わせることで、簡易な構成にもかかわらず等身大の人をあたかもその場のステージ上に存在するかのように提示できます(図1). この手法では映像が提示される位置が表示装置から離れることから、観察者は自分と同じ空間に映像が浮かび上がったかのように見え、映像中の被写体を実物として知覚します<sup>(1)</sup>. 本手法をスポーツのライブビューイングに応用すれば、競技を実際に目の前で観戦しているかのような高い臨場感を与えることができます.

#### 空中像による臨場感向上

ライブビューイングでは、複数の観

察者が同じ場に集まりコンテンツを同時に視聴することが多いため、広い視点範囲に対して被写体の空間位置を正しく提示する必要があります.しかし、これまでのTVやスクリーンでは全観察者に同じ視点の映像を提示するため、被写体の正しい位置や距離感まで再現することができませんでした.これは、2D空中像を用いた場合でも同様の問題が生じます.

これまでに、広い視点範囲に被写体の空間位置を提示可能なディスプレイ技術が多く研究されています<sup>(2),(3)</sup>. しかしこれらの研究では、多くのプロジェクタや特殊な光学系を必要として



おり、装置構成が複雑で大規模になり、 表示面の大型化も困難なため、目的と するライブビューイングへの適用は困 難です.

一方、私たちが注目している2D空 中像表示手法では、装置構成がシンプ ルなため、大型化はもちろん、スマー トフォンやタブレットと組み合わせた 手軽な空中像表示も可能です(4).

2D空中像表示では、虚像の表示面 の位置が固定されるため、奥行き方向 の移動表現はできません. これまでに、 2D空中像表示に対して奥行きを付与 する研究も行われてきましたが(5)、観 察者は1名に限られ、特殊なメガネも 装着する必要があります. これまで. 広い範囲にいる複数の観察者に対し て、被写体の奥行き移動を表現すると いう試みはされていません.

本稿では、シンプルな構成で複数の 観察者に対して被写体の奥行き移動を 表現するため, 心理的な奥行き知覚の 効果を積極的に用いる手法を提案し、 全周囲から取り囲んで観察できる表示 システムへの適用について紹介し ます.

#### 遠近法と奥行き知覚

心理的要因による奥行き表現手法と して、遠近法が古くから広く知られて います. 遠近法は、被写体の大きさや 位置を眼に映るものと同様に描くこと で、絵画等の固定された2D平面への 描画であっても観察者に大幅な距離感 を認識させることが可能です. この遠 近法を2D空中像表示に適用すること

で、奥行き方向への移動を提示するこ とができます.

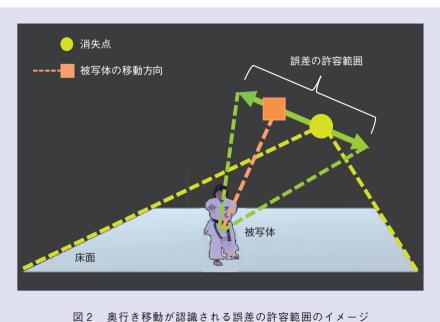
遠近法では、被写体までの奥行き距 離が遠くなるほど描画される位置や大 きさはある1点に収束します. この点 は消失点と呼ばれ、例えば奥行き方向 に無限に伸びる床面が持つ左右の稜線 も消失点で1点に交わります。被写体 がまっすぐに奥行き方向へ移動する場 合も、被写体の移動方向はこの消失点 に向かいます. 複数の視点がある場合、 視点ごとに消失点が異なるため、被写 体の移動方向はそれぞれの視点ごとに 設定しなければなりません. 2D空中 像表示ではステージ等の実物体による 消失点は視点ごとに正しいものが与え られますが、空中像として表示する被 写体には1つの視点の消失点しか設定 できず、異なる視点位置から見た際に

本来の向きへの奥行き移動が表現でき ないという問題がありました.

#### 奥行き付与手法

図2のように、破線と四角形で示す 被写体の移動方向が、丸印で示す床面 が持つ消失点に向かう方向と一致せず 誤差を有する場合でも、経験的に人は 奥行き方向への移動を認識します. こ の奥行き移動が認識される誤差の許容 範囲を、被験者実験により検証し定義 しました<sup>(6)</sup>. そして、視点ごとに許容 範囲を設定し、それらの共通範囲に収 まるように被写体の移動方向を決定す ることで、複数の視点に対して奥行き 移動を表現する手法を提案しました.

奥行き移動が認識される誤差の許容 範囲は、被写体と床面との位置関係に 大きく依存します。被写体が被写体領



域の足元,床面奥側の左右 2 端点を結ぶ三角形の領域内に向けて移動している場合は、実際の奥行き方向への動きに違和感が小さくなります.この知見に基づき、2D空中像の奥行き移動を表現するための手法を、被写体の位置、観察範囲の情報から、フレームごとの被写体位置を決定するアルゴリズムとして定めました。また、ステージ等の実空間の床面に虚像による被写体が接地しているかのように表現するための調整も行います.

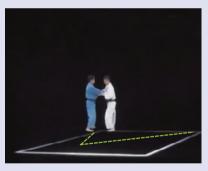
本手法によって、被写体を違和感なく床面に接地させつつ、奥行き移動を表現できるようになりました。奥行き付与手法の適用前後における被写体の移動軌跡を図3に示します。実線が被写体の移動軌跡を、破線と床面で囲まれる領域が許容範囲を示しています。この図から、提案手法の適用前では移動軌跡が許容範囲から逸脱しますが、適用後は許容範囲内に収まることが分かります。これは、等身大の被写体が10 m×10 mの床面内を移動範囲とす

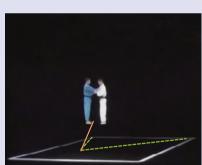
るとき, 約7.5 m~9.0 m相当の奥行き 移動を表現できる可能性を示してい ます.

#### 競技空間を取り囲んで観戦する "Kirari! for Arena"

2Dディスプレイとハーフミラーから構成される光学構成を1つの床面を共有するように4つ組み合わせ、それぞれの虚像面に4方向から撮影された被写体を表示することで、全周囲から観察可能な"Kirari! for Arena"のプ

--- 許容範囲の境界





被写体の移動軌跡

提案手法適用後

提案手法適用前







移動開始時

図3 提案手法の適用前後における被写体の移動軌跡

移動終了時

時間



ロトタイプシステムを構築しました (図4). このシステムでは、被写体の みが抽出された映像と、LiDARによ り計測した被写体の距離情報を入力と し、本手法による画像処理をリアルタ イムに行って表示します. 多数のプロ ジェクタや特殊な光学素子が必要な従 来の表示装置に対して、汎用のディス プレイとハーフミラーだけの簡易な構 成で全周囲への表示を実現でき、高い 実用性を備えていることから、複数人 で競技空間を取り囲んで観戦するス ポーツのライブビューイングなどへの 応用が期待されます.

#### 今後の展開

今回、ディスプレイとハーフミラー を組み合わせた簡易な構成で実現でき る2D空中像表示に着目し、心理的奥 行き効果を積極的に活用して被写体の 奥行き移動を複数の観察者に対して提 示する奥行き表現手法を提案しまし た. また、それを応用し、全周囲から 空間位置も含めた被写体を観察可能な "Kirari! for Arena" のプロトタイプ を構築しました.

今後、被写体位置のより正しい表現 や表現エリアの拡大について、陰影表 現など多くの映像表現手法の適用に よって実現していく予定です.

#### ■参考文献

- (1) 宮崎: "三次元像の空中表示技術の現状と展 開,"光学, Vol.40, No.12, pp.608-615, 2011.
- (2) S. Yoshida: "fVisiOn: 360-degree viewable glasses-free tabletop 3D display composed of conical screen and modular projector arrays," Optics Express, Vol. 24, No. 12, pp. 13194-13203, 2016,
- (3) B. Alex, H. Otmar, I. Shahram, H. Steve, M. David, K. David, and K. Danny: "Vermeer: Direct Interaction with a  $360^{\circ}$  Viewable 3D Display," Proc. of UIST 2011, pp.569-576, Santa Barbara, U.S.A., Oct. 2011.
- (4) 巻口・高田・新島・松井・横山: "モバイル 端末で利用可能な小型多層空中像投影装置の 提案と実装,"情処学論, Vol.58, No.11, pp. 1787-1798, 2017.
- (5) O. Bimber, B. Fröhlich, D. Schmalstieg, and L.M. Encarnação: "The Virtual Showcase,"

ACM SIGGRAPH 2005 Courses, No.3, Los Angels, U.S.A., July-August 2005.

床面

(6) 井阪・巻口・高田・小野:"2D空中像の奥行 移動知覚に関する一考察," 2017年映像情報 メディア学会年次大会予稿集, 34D-5, 2017.



(左から)巻口 誉宗/ 井阪 高田 英明

人々の日常を変革する視聴体験を早期に 提供できるよう、研究開発を日々推進して いきます.

#### ◆問い合わせ先

NTTサービスエボリューション研究所 ナチュラルコミュニケーションプロジェクト TEL 046-859-3675 E-mail 3dimg-ml@hco.ntt.co.jp

超高臨場感通信技術Kirari! 裸眼3D ポヴリニアブレンディング 運動視差

## 360度テーブルトップ型裸眼3D映像表示技術

NTTサービスエボリューション研究所では、遠隔地の競技空間などを取り囲んで観察する視聴体験の実現に向け、3D映像をテーブル上に浮かび上がらせる360度テーブルトップ型裸眼3D映像表示技術を開発しました。本稿では本技術の基本となる視覚的な知覚メカニズムとプロジェクタ配置の光学構成を中心に紹介します。

まきぐち もとひろ たかだ ひであき

#### 巻口 誉宗 /高田 英明

NTTサービスエボリューション研究所

#### 裸眼3Dによる超高臨場感な 映像提示

将来の究極の映像提示手段の1つと して、3Dメガネ等を必要とせずに運 動視差も含めた自然な立体視が可能な 裸眼3Dディスプレイが期待されてい ます. 特に. 表示対象があたかもテー ブル上に実在するかのように提示でき るテーブルトップ型の3D映像表示技 術は、スポーツ競技のフィールド全体 を俯瞰して観戦するライブビューイン グや工業製品のモデリング等、幅広い 応用が期待されます. 複数人でテーブ ルを囲んでスポーツを観戦できれば, 各自が自分の応援するチームのプレー を望む方向から自由に覗き込むことが 可能となり、観察者どうしのコミュニ ケーションも飛躍的に高まります. こ うした体験を実現するには、観察者の 表情やアイコンタクトを妨げる特殊な 3Dメガネなどを用いずに3D映像を観 察でき、さらに360度の視点移動にも 対応した裸眼3D映像表示技術が必要 となります.

360度の視点移動が可能な裸眼3D映像表示技術としては、円錐形の特殊スクリーンに対し、微小な間隔で設置した数百台規模のプロジェクタで3D映

像を提示するシステムなどが提案されています<sup>(1)</sup>. こうした,複数のプロジェクタで各視点位置に応じた視点映像を投影する手法では, 3Dコンテンツを複数人が同時に裸眼で視聴できるという利点がある一方, 視点移動に伴う映像の切替えをなめらかにするためにプロジェクタを密な間隔で大量に配置する必要があり,機材コストの増加や装置が複雑かつ大規模になるという問題がありました.

NTTサービスエボリューション研 究所ではこれまでに、隣り合う視点映 像の輝度を視点位置に応じた輝度比率 で合成し提示することで、中間視点の 映像を視覚的に補間して知覚させる 「リニアブレンディング」という視覚 系による知覚メカニズムと、そのリニ アブレンディングを光学的に実現する 「空間結像アイリス面型光学スクリー ン」を用いることで、従来よりも4分 の1~10分の1の少ないプロジェク タ数でなめらかな視点移動を実現する 裸眼3D映像表示の基本技術を確立し、 対角50インチスクリーンと13台のプ ロジェクタを用いた垂直型の視点移動 対応裸眼3D映像スクリーンシステム を開発してきました(2). また、そのス クリーンを水平方向に応用したテーブ ルトップ型の実現に向けた検討を進め、プロトタイプの実装によってテーブルトップ型に対してもリニアブレンディングのプロジェクタの削減効果があることを示してきました<sup>(3)</sup>.しかし、これまでの検討では垂直型に向けて開発したスクリーンと投影システムをテーブルトップ型に展開したため、光学設計の制約から視域が1方向の左右約65度に限られ、360度の全周囲化には至っていませんでした.

本稿では、リニアブレンディングを 用いたテーブルトップ型裸眼3Dスク リーンシステムの視域を360度に拡張 するための新たな光学構成と、60台の プロジェクタを用いたテーブルトップ 型の新たなプロトタイプについて紹介 します.

#### 視覚系による知覚メカニズムの 活用

左右方向に微小にずれて重なり合った同一の2つの像において、それぞれの像の輝度の比率を変化させると、人の網膜上には図1(a)のような2重エッジが左右についた像として映ります。しかし、人の視覚系では、像の2重エッジの幅がある程度小さい場合には、2重エッジではなく1つのエッジ

として知覚されることから、図1(b) に示すように輝度の比率に応じてエッ

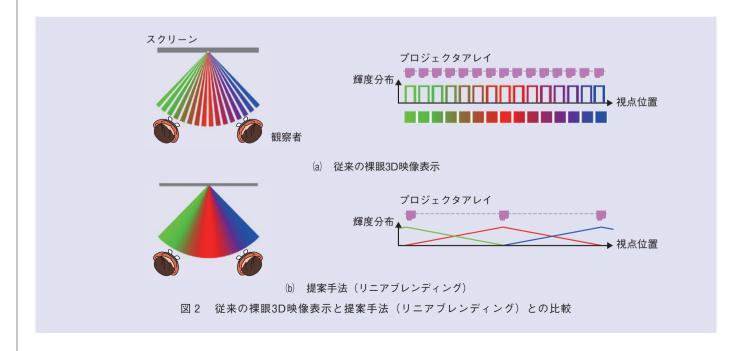
ジの位置がなめらかに遷移することが 知られています<sup>(3)</sup>. この2つの像の輝

度の比率分配に応じた像のエッジ知覚 のメカニズムを隣接するプロジェクタ の映像の重ね合わせに適用した表示原 理をここではリニアブレンディングと 呼んでいます。

これまで提案されている複数のプロ ジェクタを用いた裸眼3D映像表示で は、 両眼視差となめらかな運動視差を 提示するために**図2**(a)のようにプロ ジェクタを両眼間隔よりも狭い間隔で 設置するため、広視域化するほど大量 のプロジェクタが必要になります. 一 方、リニアブレンディングでは、視差 が融合限界内に収まる2つの視点映 像を視点位置に応じた輝度比率で合成 して提示することで、中間視点が知覚 的に補間されるため中間視点分のプロ ジェクタが不要となります (図 2 (b)).



図1 リニアブレンディングにおけるエッジ知覚メカニズム



これにより、両眼の間隔より広いプロジェクタ間隔でも両眼視差となめらかな運動視差を持つ3D映像を表示することができるようになりました.

#### テーブルトップ型に向けた 光学構成最適化

リニアブレンディングを実現する空間結像アイリス面型スクリーン<sup>②</sup>は反射層,フレネルレンズ層,拡散層から構成され,プロジェクタから投影された光はスクリーン上で反射し,テーブルトップ型の水平配置ではスクリーンを挟んで対面に位置する空間上に集光します。これによりプロジェクタのアイリス面(投影光学系の絞り相当の部分)が空中に結像され、観察者はこの範囲内でのみ投影映像を視認できます。これまでに検討を進めてきたテー

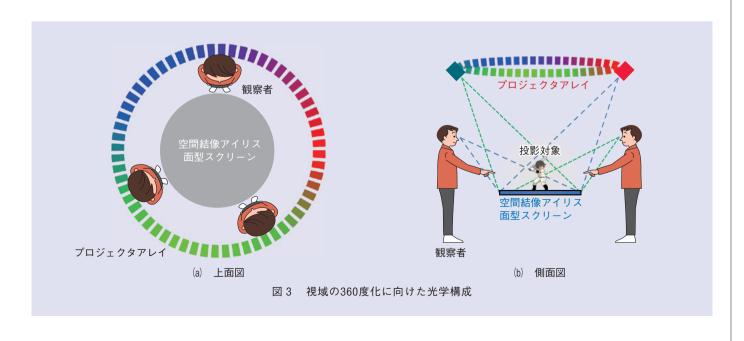
ブルトップ型の構成<sup>(4)</sup>では、プロジェクタを空間結像アイリス面型スクリーンに対し水平方向に直線に並べた直線状のプロジェクタアレイを用いていました。しかし、この直線状プロジェクタアレイでは視点がプロジェクタアレイの中心から水平方向に離れるほど投影する視点映像の歪みが大きくなり、全周囲化に向け直線状プロジェクタアレイを4方向に組み合わせて拡張するだけでは全視点に対して対象な3D映像を提示することができません。

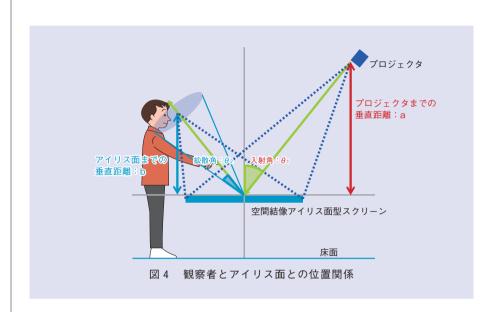
そこで今回、視域の全周囲化をめざし、図3のようにプロジェクタを円形に配置する円形プロジェクタアレイの光学系を設計しました。観察者とアイリス面との位置関係の側面図を図4に示します。アイリス面の輝度分布は拡散層の特性によって、中心を最大とし

中心からの距離が離れるにつれて減衰するように制御できます。隣り合う視点映像が視点位置に応じた輝度比率で合成されるように拡散層の拡散角 $\theta_2$ とプロジェクタの水平方向の配置間隔を最適設計することで、リニアブレンディングの効果によって中間視点映像を補間できます。

# 360度テーブルトップ型裸眼3D映像表示システムのプロトタイプ実装

前述した円形プロジェクタアレイの 光学系を用いた全周囲360度裸眼3D映 像表示システムのプロトタイプを構築 しました(図5).空間結像アイリス 面型スクリーンの直径は110 cm, ス クリーンの床面からの高さは70 cmと しました.プロジェクタは解像度1280 ×800 px (輝度600 lm)を60台, 6 度





間隔で設置しています。各プロジェクタから投影する視点映像は、それぞれのプロジェクタに対応する60台のクライアントPC上でリアルタイムに時刻同期してレンダリングしました。各クライアントPCは共通のCG空間を有しており、対応するプロジェクタ位置に設置した仮想カメラの映像をレンダリングします。視点映像の時刻同期は3Dモデルのアニメーションのスタート情報や位置情報をサーバPCからUDP(User Datagram Protocol)パケットで送信することで行いました。





図 6 周囲 5 方向から撮影した3D映像

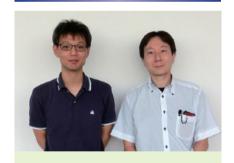
構築したプロトタイプの周囲 5 視点から撮影した写真を図6に示します. この写真から視点位置の変化による運動視差が再現されていることが分かります.

#### 今後の展開

今回、全周囲へなめらかな視点移動 を伴う自然な裸眼3D映像を提示する 技術の実現性をプロトタイプにより実 証しました. しかし今回のプロトタイ プでは、リニアブレンディングの原理 上生じる隣り合うプロジェクタの映像 が重なる微小な2重像に加え、隣接し たプロジェクタ以外の映像が視認され ることによる多重像が画質の劣化の原 因となっています。原理上生じる2重 像を視覚的な効果に影響のない範囲で 打ち消す取り組み(5)も開始しており、 今後、スクリーンの光学特性の見直し や最適化によるさらなる画質改善を進 め、将来的にはスポーツバーのような 場所でのテーブルを囲んだ超高臨場感 スポーツ観戦などの実現につなげてい く予定です.

#### ■参考文献

- S. Yoshida: "fVisiOn: 360-degree viewable glasses-free tabletop 3D display composed of conical screen and modular projector arrays," Optics Express, Vol.24, No.12, pp.13194-13203, 2016.
- (2) M. Makiguchi, T. Kawakami, M. Sasai, and H. Takada: "Smooth Motion Parallax Glassless 3D Screen System Using Linear Blending of Viewing Zones and Spatially Imaged Iris Plane," SID 2017 Digest of Technical Papers, Vol.48, No.1, pp. 903-906, 2017.
- (3) S. Suyama, S. Ohtsuka, H. Takada, K. Uehira, and S. Sakai: "Apparent 3-D image perceived from luminance-modulated two 2-D images displayed at different depths," Vision Research, Vol.44, No.8, pp.785-793, 2004.
- (4) 巻口・高田: "光学リニアブレンディングを 用いたテーブル型裸眼3Dスクリーンシステム の提案と基礎検討," マルチメディア, 分散, 協調とモバイル (DICOMO2018) シンポジウム予稿集, pp. 509-513, 2018.
- (5) M. Makiguchi, H. Takada, T. Fukiage, and S. Nishida: "Reducing Image Quality Variation with Motion Parallax for Glassless 3D Screens using Linear Blending Technology," SID 2018 Digest of Technical Papers, Vol.49, No.1, pp.251-254, 2018.



(左から)巻口 誉宗/ 高田 英明

超高臨場感を実現する自然な裸眼3D映像は人々の生活を変革するインタフェースの1つとなり得ると考えています。あたかもそこに実在するかのような体験をどこでも手軽に実現できるように研究開発を推進していきます。

#### ◆問い合わせ先

NTTサービスエボリューション研究所 ナチュラルコミュニケーションプロジェクト TEL 046-859-2539 E-mail 3dsc-mem-ml@hco.ntt.co.jp 超高臨場感通信技術Kirari! VR/AR スタジアム観戦 映像処理・提示技術

## 「VR&ARを活用したスタジアム観戦」における 競技フィールド再構成のための映像処理・提示技術

近年、VR(Virtual Reality)/AR(Augmented Reality)を活用したスポーツ観戦への期待が高まっています。NTTメディアインテリジェンス研究所では、VR/ARのそれぞれの利点を活かし、試合展開に応じて再構成された競技フィールドの映像を、観戦者が操作することなくVR/AR表示することで、より試合を応援したくなる「VR&ARを活用したスタジアム観戦」の実現をめざしています。本稿では、VRとARで共通的に利用できる映像コンテンツを作成する映像処理技術や、目の前に実在感のある映像を表示するための映像表示技術を紹介します。

NTTメディアインテリジェンス研究所

#### VR&ARがもたらす新たなスタジ アム観戦の世界

近年、VR(Virtual Reality)/AR(Augmented Reality)を支えるセンサデバイス、コンテンツ制作および映像配信技術の発達に伴い、VR/ARを活用したスポーツ観戦のトライアルやサービス導入が開始されています。NTTドコモは、観戦者がARデバイスとして装着したスマートグラス上の試合を阻害しない位置に、選手情報や別視点のカメラ映像を重畳表示しながら、ラグビーを観戦するトライアルを実施しています<sup>(1)</sup>. また、KDDIは、野球場に設置された5台のカメラを切り替えながらVR視聴するVRスポーツ観戦サービスを開始しています<sup>(2)</sup>.

しかし、スタジアム観戦では、目の前で繰り広げられる試合展開に応じて VR/ARに適したシーンが切り替わる ため、VR/ARのいずれかだけではすべてのシーンを満たすことができません. 試合観戦そのものを阻害しないよう、観戦者が意識して操作することなく、VR/ARが切り替えられる仕組みが必要になります. 私たちは、VR/ARのそれぞれの利点を活かし、試合

展開に応じて再構成された競技フィールドの映像を、観戦者が操作することなくVR/AR表示することで、観戦者がより試合を応援したくなる「VR&ARを活用したスタジアム観戦」の実現をめざしています。このために、VRとARで共通的に利用できる映像コンテンツを作成する映像処理技術や、目の前に実在感のある映像を表示するための映像表示技術の研究開発を進めています。

本稿では、VR&ARを活用したスタ ジアム観戦を実現するためのシステム 概要を説明した後、VR/AR表示用に 競技フィールドを再構成する映像処理 技術として、カメラを設置することが できない競技フィールド内の映像を生 成するための「3次元再構成技術」, 注目選手にフォーカスさせるため注目 選手以外を目立たなくした映像を生成 する「隠消技術」について紹介します. また、観戦席以外の場所でもスポーツ 観戦を盛り上げることを目的とした将 来のARデバイスとして、テーブル上 に競技フィールドを立体的に表示する ための「視覚的に等価なライトフィー ルド映像生成技術」について紹介し

#### VR&ARを活用したスタジアム観 戦システムの概要

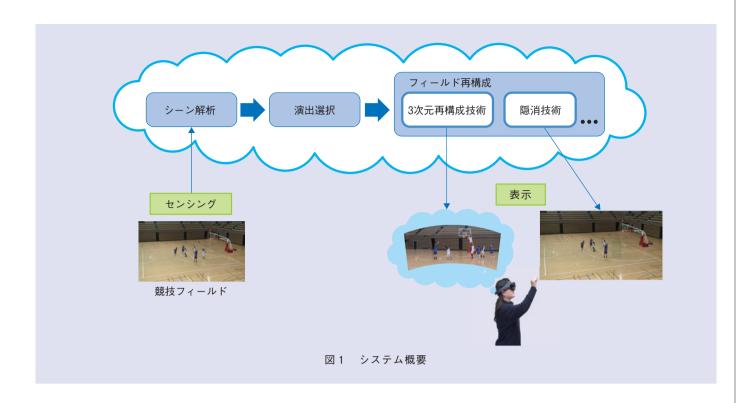
VR&ARを活用したスタジアム観戦を実現するためのシステム概要を図 1に示します.

スタジアムを取り囲むように設置された複数カメラは、競技フィールド全体をセンシングし、カメラ映像をクラウドにアップロードします。クラウド上では、カメラ映像から競技フィールドで起こっている事象を解析し、その情報を基に観戦者にどのような映像を提示するか演出が選択されます。そして、演出に応じてカメラ映像から競技フィールドを再構成し、観戦者の視聴端末に映像が配信されます。

#### 競技フィールド再構成のための 3次元再構成技術

はじめに、競技フィールドを再構成する映像処理技術の1つである、カメラを設置することができない競技フィールド内の映像を生成する「3次元再構成技術」について紹介します.

スタジアムやアリーナなどでは観戦 席が決まっているため、限定された方 向からしか観戦できなかったり、競技



フィールドまで遠く選手が見えなかっ たり、観戦者が試合展開に応じて期待 した視点から観戦することができませ ん. 自由な視点をつくる取り組みとし て、これまで競技フィールドを取り囲 むように配置したカメラから、カメラ を配置できない競技フィールド内の視 点の映像を生成する自由視点映像合成 技術が提案されてきました. しかし, 広域な競技フィールド全体をセンシン グするためには、 多数のカメラを会場 に配置しなければならず、既設施設へ の設営負荷が高いという問題がありま す. また、カメラ位置より前方の競技 フィールド内の視点映像は、カメラ映 像の空間分解能の限界や、隠蔽領域の 不良設定問題のため、スポーツ観戦に 耐え得る品質の映像を生成することが できません.

この問題を解決するため私たちは、 撮影した映像から選手の動きを推定 し、その推定結果を事前に作成してお いた当該選手の高品質なCGモデルに 割り当てることで、人物の擬似的な再 構成を行う手法を提案しています.近 年の映画やゲームコンテンツなどで見 られるように、CGの品質は非常に高 い水準に達しているため、本手法でも 現実と見紛うような品質の自由視点映 像を生成することが可能となります.

本手法の実現のためには、映像から 選手の動きを推定する技術が必要とな ります. 近年、DeepLearningを用い てカラー画像から人物の動作を高精度 に推定する技術<sup>(3)</sup>が提案されています が、これらの技術は遮蔽物により被写 体である人物が部分的に隠れてしまう 場合に動作の推定精度が大きく低下す るという問題があります。 スポーツ シーンでは、1つのシーンに多数の選 手が入り乱れ、互いに遮蔽し合う状態 は容易に発生するため、遮蔽物に対す る頑健性を高めることは重要な課題で す、そこで私たちは、カラー画像から 直接被写体の動作を推定するのではな く、まず、遮蔽物に隠された被写体の シルエットをカラー画像から推定し, 推定したシルエット画像とカラー画像

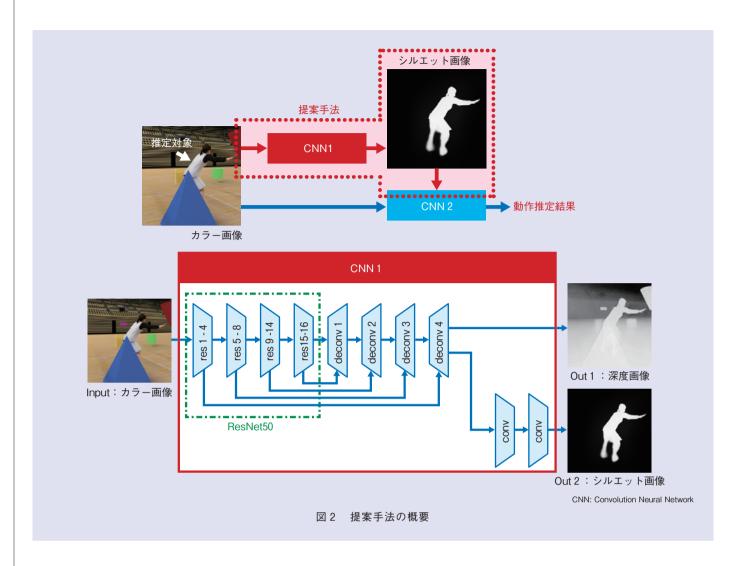
を利用して人物の動作を推定する手法 を提案しました(図2).シルエット 画像を用いて人物の動作の候補範囲を 限定することで、従来技術と比較して より高精度な動作推定が実現できます.

#### 競技フィールド再構成のための 隠消技術

次に,注目選手にフォーカスさせる ため注目選手以外を目立たなくした映 像を生成する「隠消技術」について紹 介します.

映画や漫画では、注目させたい人物 のみを残し、さらにスポットライトな どの効果を付与することで、そのシー ンで特に注目させたい人物にフォーカ スを当てる演出が行われています.私 たちは、このような演出をスタジアム やアリーナなどで開催される試合にお いて適用することで、勝敗を左右する 決定的なシーンで得点にかかわる選手 への注目が高まり、試合観戦をより盛 り上げることができると考えています.

実空間に存在する物体を目立たなく



する技術として、対象領域の背景映像を対象領域に重畳することで、対象領域を取り除いた映像を生成するDR (Diminished Reality) 技術が提案されています。スポーツ競技に適用するためには、重なり合うようにプレイしている複数の選手の前後関係(3次元情報)を考慮しながら、対象領域のみを目立たなくするDR技術が必要となります。また、観戦者はさまざまな位置姿勢でスポーツ競技を観戦しているため、観戦者の視点ごとに対象領域を目立たなくする処理を施した映像を提示しなければなりません。

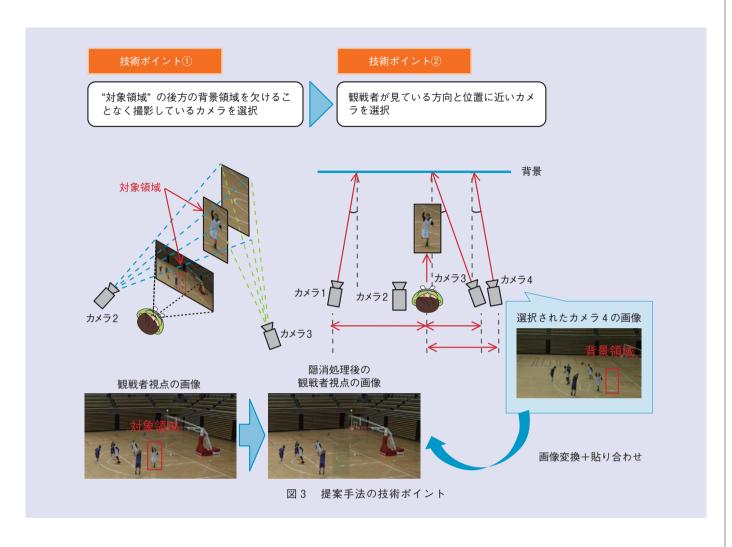
このような課題に対して、私たちは、 スタジアムやアリーナなどの競技会場 を取り囲むように設置した複数カメラ から、対象領域の背景を撮影している カメラを選択し、観戦者から見て競技 フィールドに一致するように画像変換 した画像を貼り合わせることで、対象 領域を目立たなくする手法を提案して います<sup>(4)</sup>.

提案手法の技術ポイントを図3に示します。第一の技術のポイントとしては、観戦者から競技フィールドまでの距離が遠く、競技フィールドを再構成するのに必要な奥行きが狭いことに着目し、競技フィールドを離散的な多層平面と仮定して3次元情報を推定しています。これにより、競技フィールド全体の3次元情報を推定する手法と比較して少ない計算量で、対象領域の前後関係(3次元情報)を考慮し

ながら、対象領域の背景を撮影しているカメラ群を抽出することができます。第二の技術ポイントは、対象領域の背景を撮影しているカメラ群から、観戦者が見ている方向と位置に近いカメラを選択していることです。これにより、背景画像に施す画像変換量が小さくなり、幾何歪みを抑えながら対象領域を目立たなくすることができます。

#### 視覚的に等価なライトフィールド 映像生成技術

最後にARデバイスとして、テーブル上に競技フィールドを立体的に表示するための「視覚的に等価なライトフィールド映像生成技術」について紹介します。



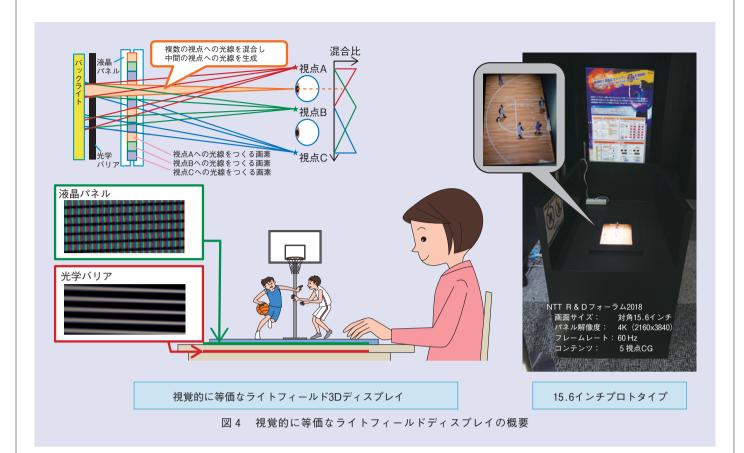
私たちはチケット売り場やロビーなど、スタジアムやアリーナの観戦席以外のスペースで試合のハイライトシーンを提示することで、選手やチームに興味を持ってもらい、リピーターを増やすことができると考えています。観戦体験をより魅力的に提供するには、その場に選手がいるかのような実在感を提示することが重要になります。

選手, コート, ボールなどの物体は 周囲の照明などの光を反射したり, 自 身が発光したりすることで, 方向によ り色や明るさが異なる光線を発してい ます. この多数の光線がつくる場はラ イトフィールドと呼ばれ, 正確に提示 することができれば, 高い実在感で シーンを提示できます. しかし, 方向 によって異なる光線を表示するには. 方向の数だけ画素を割り当てることになります。例えば、100方向の光線を生成するには、同じ解像度の2次元ディスプレイの100倍の画素が必要になってしまいます。

人間の視覚は眼の瞳に入る光線から物体を知覚します.これは、カメラで写真を撮るような単純な画像の取得だけではなく、左右の眼という異なる視点で得られる光線情報の違いや、身体の揺れや眼球運動による繊細な視点位置の変化に対応した微小な光線の変化を利用しています.この知覚は一見すると高度ですが、人間は必要のない要素には鈍感です.実物からの光線とは異なっていても、人間の知覚にとって等価と感じるものを提示することができれば、人間は実物と同じように知覚

するというのが視覚的に等価なライトフィールドディスプレイの考え方です。本技術では、離散的な視点への光線の情報を基に、ディスプレイ内部で複数の光線を光学的に加重平均することで、中間の視点への光線を等価的に補間生成しています<sup>(5)</sup>.

装置の概要を図4に示します. 左上に示したように, バックライト光源の前面にストライプ状の光学バリアを設置し, その前面に液晶パネルを設置します. 例えば, バリアの隙間から視点 Aに至る光線は液晶パネルの視点Aへの光線をつくる画素だけを照明するので, 視点Aからはその画素群で構成された画像が見えます. 視点Bでも同様です. 視点AとBの中間から見ると, 視点からの距離に応じて加重平均され



た光線が見えることになり、中間の視点と等価に知覚されます。今回はスポーツということで、左下の図のように平面のコートに対応して画面を水平に置き、コートの隅々まで高画質で表示可能としました。

ライトフィールドを再現しているため、大人と子どものように左右の眼の間隔が異なっても、正しい奥行きで知覚できます。また、補間の効果は画素数の節約だけではなく、そのアナログな滑らかさが実在感の向上に貢献していると考えています。

#### 今後の展開

VR/AR関連技術の向上に加えて、2020年に向けてスポーツが注目されていることから、今後も、VR&AR技術を活用したスタジアム観戦への期待は高まっていくと予想されます。NTTメディアインテリジェンス研究

所では、試合展開に応じて演出および VR/AR提示を選択するためのシーン 解析技術、競技フィールドを再構成す るための映像処理技術と視覚的に等価 なライトフィールド映像生成技術の研 究開発を推進し、新たなスポーツ観戦 体験の創出に貢献していきます。

#### ■参考文献

- (1) https://www.nttdocomo.co.jp/binary/pdf/ info/news\_release/topics/topics\_171221\_01.
- (2) http://news.kddi.com/kddi/corporate/ newsrelease/2018/07/25/3281.html
- (3) Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh: "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," CVPR 2017, pp.7291-7299, 2017.
- (4) M.Isogai, D.Ochi, and H.Kimata: "Diminished Reality Using Plane-based Reconstruction Method," IMID 2017, F38-4, 3DSA, Busan, Korea, Aug. 2017.
- (5) 伊達・越智・木全: "視覚的に等価なライトフィールドフラットパネル3Dディスプレイ," 第22回日本バーチャルリアリティ学会大会講演予稿集, 1B4-04, 2017.



(後列左から) 能登 肇/伊達 宗和/ 木全 英明/亀田 明男 (前列左から) 岡見 和樹/松村 誠明/ 磯貝 愛

観戦者がより試合を応援したくなる「VR&ARを活用したスタジアム観戦」の実現をめざし、競技フィールドを再構成するための映像処理技術と視覚的に等価なライトフィールド映像生成技術の研究開発に取り組んでいきます。

#### ◆問い合わせ先

NTTメディアインテリジェンス研究所 画像メディアプロジェクト TEL 046-859-2938 FAX 046-859-2829

E-mail megumi.isogai.ks@hco.ntt.co.jp

### 主役登場

## Kirari!における被写体抽出技術の 今とこれから

## 柿沼 弘員

NTTサービスエボリューション研究所 研究員



近年、AR(Augmented Reality: 拡張現実)やMR(Mixed Reality: 複合現実)と呼ばれる、映像、音声などのデジタル情報を目の前の現実空間と融合させて提示するコンテンツの観せ方が広まりつつあります。

私たちの取り組んでいる超高臨場感通信技術Kirari!は、スポーツの試合、音楽や演劇のステージ、著名人の講演など、大勢で観戦・観賞・聴講するような空間一体型のコンテンツを、通信と映像、音声の技術を組み合わせて別の空間に伝送・再構築する技術であり、これまで現場でしか味わえなかった臨場感を、遠く離れた会場でも体感できるような世界を実現しようとしています。

私はKirari!の要素技術群の中でも、映像から被写体領域を抽出する研究に取り組んでいます。これは映像中の対象人物を背景から切り離し、あたかも別の空間に存在しているかのように表示するための技術です。具体的には、ニュース番組の天気予報や、ハリウッド映画のCG合成などで用いられるクロマキーと同じ効果をねらいながらも、スタジオやグリーンバックなどの環境を必要とせず、かつリアルタイムで被写体のみを抽出することを実現しようとしています。

私たちはこの技術を用いて、直近2年間で10以上の異なる会場で、一般のお客さま向けの実証実験やプロモーションイベントを一部事業会社の方々と協力しながら行ってきました。中でも多方面から好評いただいているニコニコ超会議の超歌舞伎では、毎年新しい技術を取り入れてお

り、3年目の今年は機械学習(ニューラルネットワーク)を利用した被写体抽出技術を一部のシーンに組み入れて、 演出を盛り上げることができました。しかしその裏側では、 本番と同じ条件での撮影が直前までできない舞台環境とい うこともあり、最終リハーサル時点でも精度の良い抽出結 果が得られていませんでした。そのため、公演期間中も毎 回本番映像を追加して学習を繰り返し、後半になってよう やく最高精度に達するという綱渡り的な運用で成立させて いたのが実際のところで、不確定要素の多い現場での利用 にはまだまだ多くの課題が残っています。

また、臨機応変な調整や判断、検証とプロトタイピングのスピード感など、現場で求められるものは決して簡単なものではありませんが、本気で良いものを創り上げようという熱意のあるNTTグループ外の方々と協力して仕事ができたり、アウトプットに対してエンドユーザから生のフィードバックが得られたりと、研究所の中にいるだけでは得られない知見やモチベーションが得られることが多く、とてもやりがいのある環境だと感じています。

今後は、撮影や学習における制約条件を緩和し、被写体抽出の適用シーンを増やしていくことを研究の軸としながらも、そもそも臨場感とは何なのか、どうしたら人間は高い臨場感を感じるのか、その本質に迫れるよう研究開発を進め、今までにない臨場感のかたちを提供していきたいと考えています。