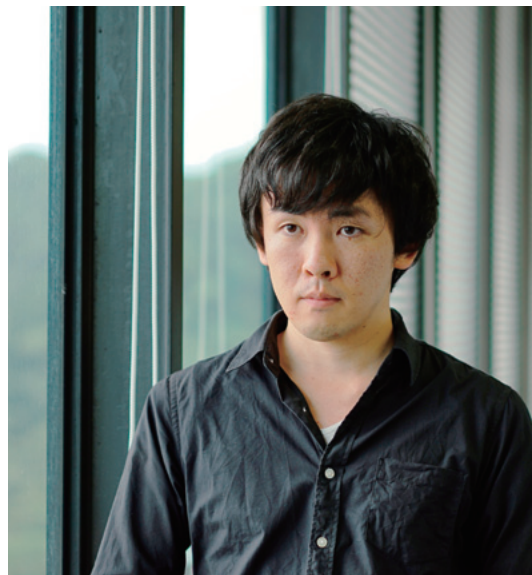


主役登場

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る別の空間に伝送・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の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、研究所の中だけで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。