



超高臨場ライブ体験を実現する通信技術の国際標準制定

ながお じろう

長尾 慈郎

NTTサービスエボリューション研究所

2019年11月、ITU-T (International Telecommunication Union - Telecommunication Standardization Sector) にて、離れた場所のイベントをリアルタイムに伝送し、非常に高い臨場感で再現する「超高臨場ライブ体験 (ILE)」に関する標準が制定されました。ILEに関する一連の標準によって、超高臨場なライブ中継を世界中どこにいても楽しめるようになることが期待されます。ここでは、ILEに関する標準について紹介します。

超高臨場ライブ体験

近年、注目度の高い音楽コンサートやスポーツイベントをリアルタイムに会場外の映画館や大型スクリーンなどに中継し、本会場と同時に遠隔の視聴会場においても多数の観客がイベントを観覧する、いわゆる「パブリックビューイング」がさかんに行われるようになってきました。パブリックビューイングの特徴は、会場が遠いなどの理由で本会場に行けない人々が、より便利な場所（視聴会場）に集まって、中継映像を見ながらお互いに感動を共有できることです。

しかし、パブリックビューイングによる視聴会場での観覧は、本会場に比べると臨場感が低くなりがちであるという課題があり、より高い臨場感で観覧を楽しめる技術が求められています。例えば、本会場でパフォーマンスしている人物があたかも視聴会場にいるかのように表示されたり、映像中の

物体の位置から音が聞こえてきたり、さらには振動や熱気といった雰囲気再現されたりすれば、パブリックビューイングの臨場感を高められると考えられます。このように映像や音声などの状況が、あたかもその場にいるかのように感じられることを「超高臨場感」と呼び、これを実現するための技術の研究開発が進められています。特にリアルタイムに超高臨場感を体験することを「超高臨場ライブ体験 (ILE: Immersive Live Experience)」と呼びます。ILEによって、世界中のどこにいても、世界のどこのイベントもその場にいるかのように体験することが可能になります。結果を知ってからでは観覧の楽しさが損なわれてしまうスポーツ競技のような、リアルなイベントとの同時性が感動の重要なファクターである場合に、ILEは特に効果的です。

ILEの実現のためには、映像や音声だけではなく、本会場のさまざまな状況（環境情報）を詳細にセンシングし、リアルタイムで遠隔地の視聴会場に伝送して再現する技術が必要です。センシング対象の例としては、例えば人物の3次元位置や、音の発生位置、照明制御情報などがあります。また、リアルタイム性を確保するためには、異なるシステム間でのフォーマット変換の負荷を軽減するとともに、遠隔地までのコンテンツ伝送の相互接続性を確保する必要があります。世界のどこへでもリアルタイムで伝送し再現するために、国境を越えた相互接続性を確保す

る国際標準が求められていました。しかし、映像や音声とともに環境情報をリアルタイムに伝送し、視聴会場で超高臨場に再現することに特化した国際標準はこれまでありませんでした。

ILEの国際標準

本会場側システムと視聴会場側システムを国際的に接続し、世界のどこにいてもリアルタイムで多くの観客に同時に超高臨場感を提供することを目的とし、2016年にITU-T (International Telecommunication Union - Telecommunication Standardization Sector) SG (Study Group) 16^{*1}にILEを研究課題とするQ (Question) 8が設置され、検討が進められてきました。その結果制定されたのがITU-T H.430シリーズです^{(1)~(4)}。これまでに制定されたILE関連のITU-T標準を表1に示します。

■H.430.1

H.430.1ではILEの定義と要求条件が規定されています。ILEの要求条件を表2に示します。要求条件には、再現された情景が目の前で起きているかのように感じるための映像の実物大表示や音像定位^{*2}、人物等を視聴会場で再現するための物体像のリアルタイム抽出、映像音声その他の情報を時間

*1 ITU-T SG16: ITUは国連の専門機関であり、ITU-Tは電気通信を世界規模で標準化する部門です。SG16はマルチメディアに関する標準化を行っています。

*2 音像定位: 複数のスピーカーなどを使って、音の方向を再現すること。



表 1 ILE関連ITU-T標準

番号	タイトル	制定年月
H.430.1	Requirements for immersive live experience (ILE) services (ILEのための要求条件)	2018/ 8
H.430.2	Architectural framework for immersive live experience (ILE) services (ILEのためのアーキテクチャフレームワーク)	2018/ 8
H.430.3	Service scenario of immersive live experience (ILE) (ILEのサービスシナリオ)	2018/ 8
H.430.4	Service configuration, media transport protocols, signalling information of MPEG media transport for immersive live experience (ILE) systems (ILEのためのサービス構成, メディア伝送プロトコル, MMTシグナル情報)	2019/11

表 2 ILEの要求条件

要求条件	内容	必須／推奨／オプション
1	実物大の表示	推奨
2	音声方向の再現	必須
3	照明等の効果による雰囲気再現	推奨
4	視聴会場の端末の性能・3次元位置に合わせた、空間環境の再現	必須
5	複数資産（映像、ビデオ、位置情報、構成情報等）の同期表現	必須
6	拡張情報の表示機能	オプション
7	物体のリアルタイム抽出	必須
8	空間情報の計測	推奨
9	映像、音声、照明、空間、構成情報の同期伝送	必須
10	映像、音声、照明、空間、構成情報の同期データの蓄積	オプション
11	再構築のための複数メディア（映像、音声、照明、空間情報等）の処理	必須
12	リアルな音像定位	推奨
13	ワイド映像生成	オプション

ずれなくリアルタイムに会場間で伝送するための同期伝送などが盛り込まれています。

■H.430.2

H.430.2は、ILEの基本アーキテク

チャと一般的な役割モデルを規定するとともに、ILEの機能を実現するための候補技術を紹介しています。ILEの基本アーキテクチャを図1に示します。アーキテクチャは以下の要素か

ら構成されています。

(1) 環境情報取得

本会場の映像、音声、人物の位置、照明状態等をカメラ、マイク、センサ等で取得します。

(2) 同期メディア伝送

同期メディア伝送機能は、複数のメディアを同期して配信する機能です。

(3) 伝送レイヤ

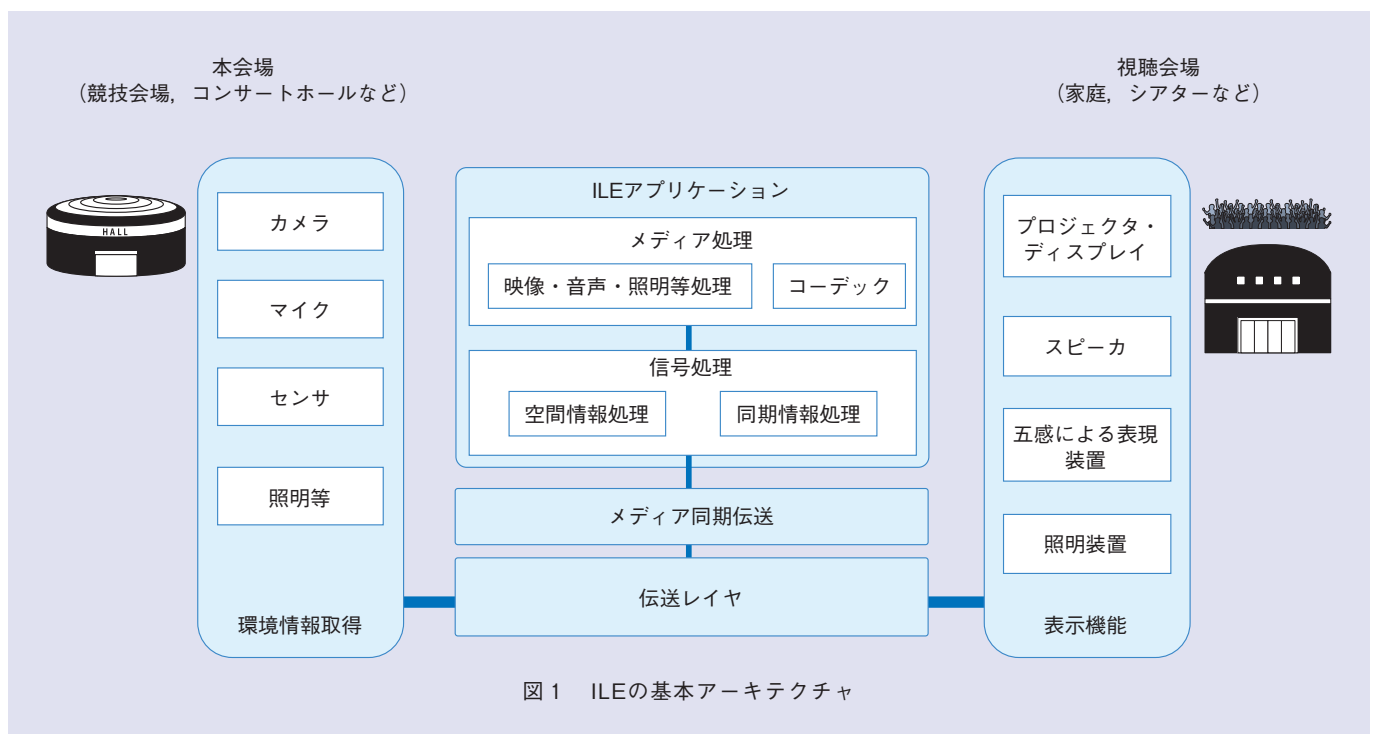
コンテンツ保護やネットワーク遅延を考慮しつつ、本会場から視聴会場にメディアを伝送します。

(4) ILEアプリケーション

映像処理では例えば、映像から人物像を抽出したり、複数の映像を統合してより臨場感の高い映像を生成したりします。コーデックでは、コンテンツのエンコード・デコード処理を行います。空間情報処理では、例えばカメラから得られた人物の位置情報と、センサで計測した人物の3次元位置情報を統合したりします。同期情報処理では、表示機能での再現に向けて、ビデオ、音声、照明、センサ情報などのコンテンツを同期させる処理を行います。

(5) 表示機能

視聴会場において、映像の表示、音声方向の再現、照明の再現などを行います。これらを視聴会場同期して実施することで、視聴者は超高臨場感を得られます。



■H.430.3

H.430.3ではサービスシナリオが類型化され、それぞれにユースケースが例示されており、ILEサービスによって可能となる新しい体験を具体的にイメージしやすくなっています。例えば以下のようなサービスシナリオが示されています。

(1) 選手目線映像

レーシングカーのドライバー目線の映像をリアルタイム伝送し、視聴会場に再現します。さらにレーシングカーのスピードや加速度、順位や残り走行

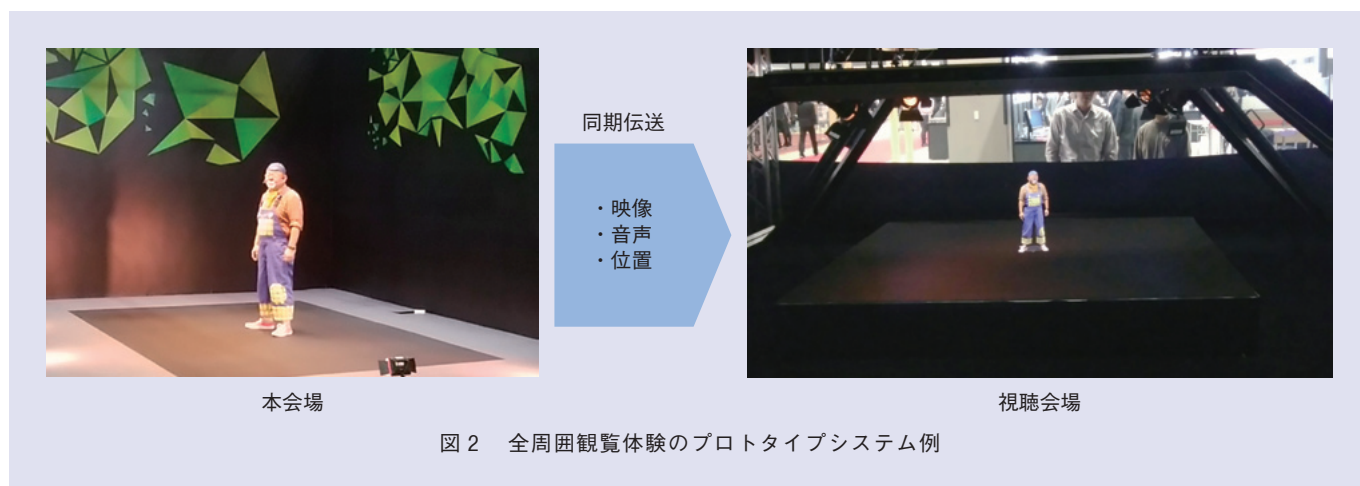
距離などの追加情報を表示することで、レースを盛り上げます。

(2) 音楽コンサートの立体表示

音楽コンサートを複数の高精細カメラ、高音質マイク等によって収録します。収録したデータを（超高臨場感を与えるように）処理し、視聴会場に同期伝送します。視聴会場では、伝送されたデータを基に3次元映像や擬似3次元映像を表示し、高音質な音声や特殊効果を用いて本会場にそっくりのイベントを再現することで超高臨場感を提供します。

(3) 全周囲観覧体験

スポーツ選手やコンサートのアーティスト等の高精細な映像を複数の角度から撮影し、高音質な音声、人物の3次元位置情報等とともに遠隔地の視聴会場に伝送します。視聴会場には人物像などを全周囲から観覧できる特殊な表示装置が設置されており、伝送されてきた3次元位置情報を用いて映像を表示し、どの方向からでもイベントを観覧することができます。複数のスピーカーを用いて、音声の方向も忠実に再現されます。このシナリオに



従って構築されたプロトタイプシステムの例を図2に示します。視聴会場の表示装置は、正面から観覧するだけでなく、横から見た場合は横からの映像を観覧することができ、あたかも本会場を歩き回りながら観覧しているかのような感覚を得られます。

■H.430.4

H.430.4では具体的な通信技術として、映像と音声に加え臨場感を高めるために必要な、物体の位置等の空間情報も同期伝送するために、MMT (MPEG Media Transport)^{*3}を拡張したメディア伝送方式が規定されています。

伝送される空間情報には、以下の項

目が含まれています。

- ・本会場のサイズ（幅，奥行き，高さ）
- ・本会場の機材の情報（機材の種類，3次元位置，向き，サイズ）
- ・ステージ上の人物等の3次元位置，画像上での位置，外接矩形情報

視聴会場においてこれらの空間情報を利用することで、本会場と視聴会場のサイズや機材の配置の違いを吸収しながら、人物等の位置を再現することが可能になります。



今後もITU-Tでは、視聴会場での表示機能の構築に関する参照モデルおよびガイドラインの整備や、VR (Virtual Reality) 等の関連技術の取り込みな

ど、他の標準化機関とも連携してITU-T勧告のさらなる拡充を検討していく予定です。

■参考文献

- (1) <https://www.itu.int/rec/T-REC-H.430.1/en>
- (2) <https://www.itu.int/rec/T-REC-H.430.2/en>
- (3) <https://www.itu.int/rec/T-REC-H.430.3/en>
- (4) <https://www.itu.int/rec/T-REC-H.430.4/en>

*3 MMT：国際標準化団体である、ISO/IEC JTC 1/SC 29/WG 11 MPEGで制定するメディアトランスポート規格です。