Software Defined Media WG
視聴空間サービスのソフトウェア制御

WIDEプロジェクト
塚田 学（東京大学）

IIJセミナー@東京

Software Defined Media WGとは

- 目的
  - インターネット環境を前提として、オブジェクト志向のディジタルメディアを用いたサービスとしての映像音響空間を提供するためのアプローチを検討する
- 開始
  - 2014年初頭 (1/22, 第一回ミーティング)
- コアメンバー
  - 東京大学
  - 慶應義塾大学
  - ヤマハ
  - KDDI
  - NTT研究所
  - Dolby
  - Panasonic
  - 他、37名

月1ミーティング (今日)

アウトライン

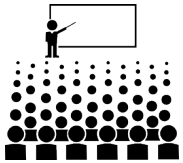
- 視聴空間サービスとは
- SDM WGの目指すもの
- 2014年のWG活動
- 今後の予定
- まとめ

視聴空間の変遷

コンテンツ　視聴空間

演劇・ダンス

スポーツ

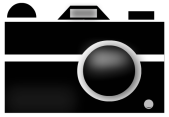

講義

- 感動や劇的な感情を共有
- 知識の伝承

コンテンツ　メディア　視聴空間

スピーカ

照明


カメラ

蓄音機

CD, DVD

- 効率化：スピーカ・照明
- 時間的に延長：カメラ・蓄音機
- 空間的に延長：CD、DVD、映画、テレビ
- 大量複製・大量消費：デジタル化

コンテンツ　インターネット　視聴空間

- インターネットが前提：収録環境も視聴空間もインターネットが浸透
- インターネット前提の視聴空間として設計されるべき

視聴空間の潮流

VOCALOID® 「初音ミク」

- ディジタルネイティヴの伝達・共有・加工
- 「見たままのビットマップ情報」から「オブジェクトごとのデータ」へ

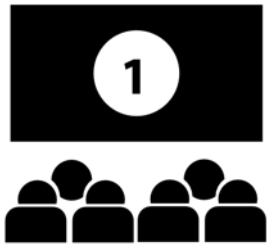
- 映像と音声において急速に進むオブジェクト化によるメタデータ活用
  - Dolby Atmos
  - ゲーム
  - 複数ビデオカメラからの立体オブジェクト抽出
  - 自由視点映像（KDDI）
- 業務用音響機器の仮想化志向
- 家庭内ネットワークを活用したハイレゾオーディオ再生

オブジェクト化された映像音声のデータをソフトウェアで制御することで、新しい表現方法とビジネスへの期待

視聴空間の現状

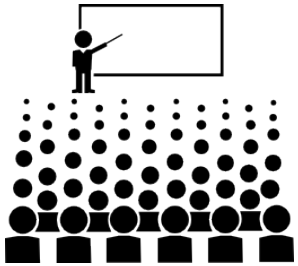
- 映像と音声のシステム
  - コンサートホールや劇場における拡声システム
  - 放送・映画・音楽制作等で使われる制作システム
  - BGMや館内放送のための音響設備など
- IPを利用した1Gbps環境の一般化により、インターネット前提のシステムが可能
- オブジェクト志向の３Dの映像と音声が登場
- 専用の設備とシステム
  - 会議室・講義室・劇場・ホームシアターなど用途別に特化した映像音声の設備
  - 計算処理ごとに専用のハードウェア

ホームシアター



劇場

テレビ会議



講演会・講義

Software Defined Media

- デジタル3次元の映像と音声が登場
- 会議室・講義室・劇場・ホームシアターなど用途別に特化した映像音声の設備
- 機能毎にハードウェアベースのシステム

1. 3次元の映像・音声の演出をソフトウェアで制御
- 汎用施設の多目的利用を想定
- 入出力機能の仮想化による装置と演出（アプリ）の分離

TV放送

- インターネットと放送が混在
- リアルタイムとオンラインのコンテンツが混在
- オブジェクトベースの映像と音声データ
- 視聴者は放送側が作成したコンテンツを単一のソースで受信

2. 映像／音声を複数ソースのミキシング可能
- リアルタイムとオンラインのコンテンツを管理・選択
- クラウド側での共同ミキシング
- 放送側でなく受信側でのミキシング
- ソフトウェアレンダリングによる拡張演出

- 特定の方向の音、特定のアングルの映像など、視聴者の特定の関心事に集中できない
- 視聴者の相互作用は時間と空間の制約を受ける

3. 視聴者の関心事への集中可能性
- フィードバックにより動的な演出制御・ソース選択
- 時間と空間の制約を超えた視聴者の相互作用

ユースケース

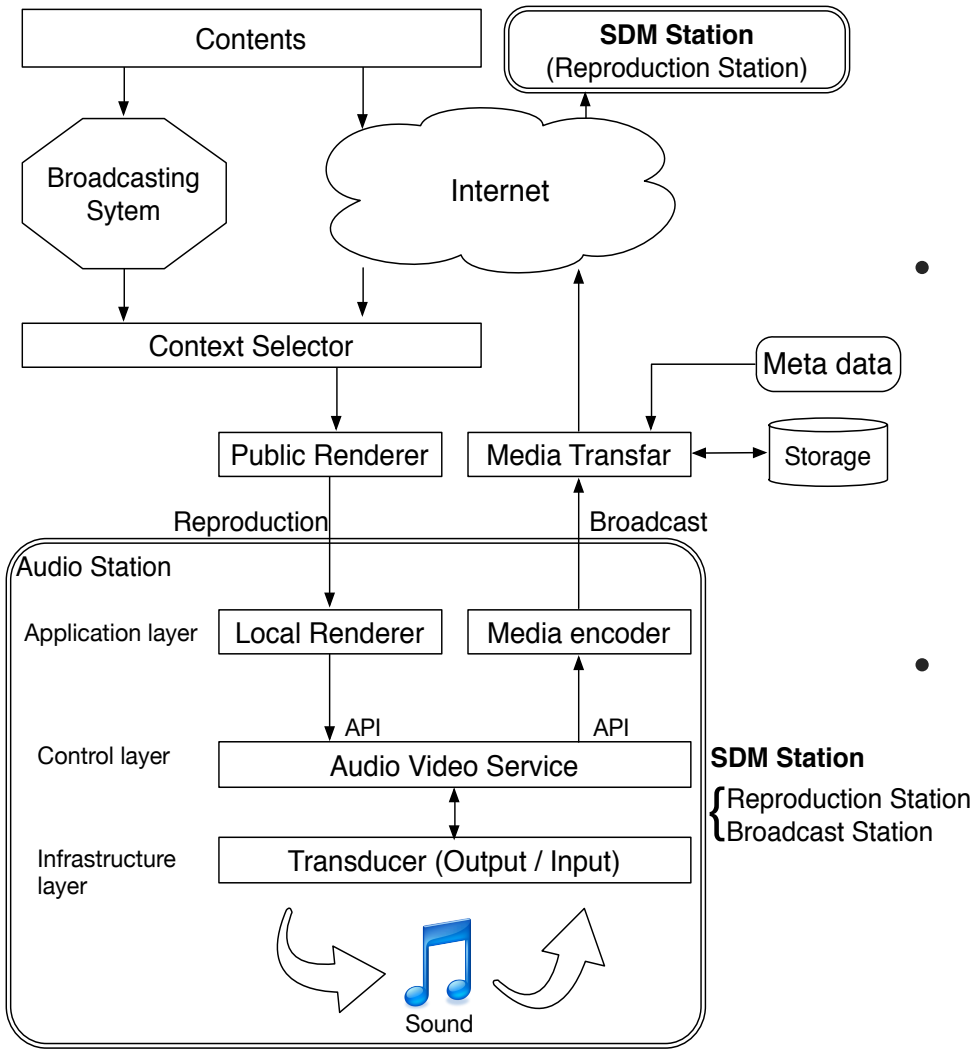
1. 3次元の映像・音声の演出をソフトウェアで制御

2. 映像／音声を複数ソースの選択ミキシング可能

3. 視聴者の関心事への集中可能性

- 汎用設備の多目的利用
  - 演劇、軽音楽、講演といった様々な演目への対応をソフトウェア制御
  - 講演会、遠隔会議、小規模音楽イベント様々なアプリケーション向けに最適化
  - 映画館を遠隔音楽教室として利用
- 柔軟なミキシング
  - リアルタイムとアーカイブを含む多チャンネルの映像と音のソースを受信側でミキシング
  - 多チャンネルの映像と音のソースを第三者クリエーターがクラウド側でミキシング
  - コンピュータグラフィクスとのミキシング（*i.e.* 拡張現実）
- 視聴者の関心事のフィードバック
  - 映像と音声の3D視聴点の指定
    - 例）スポーツ選手の一人称視聴点
    - 例）応援側スタンド方面からの視聴
  - 言語／ナレーター／サブタイトルの指定
  - 同一のイベントに関心を持つ視聴者の相互作用を促進

システム



- コンテンツ
  - リアルタイムとアーカイブ
  - 映像と音声のデジタル3次元コンテンツ
  - 柔軟な管理システム
- 編集と放送
  - インターネットと放送システム
  - コンテキストによるコンテンツ選択
  - 双方向も想定
  - クラウド側での共同編集
  - 関心事の近い視聴者の相互作用を支援
- 再生環境
  - ローカルでの処理はアプリケーションとして扱う
  - 再生設備は、機能毎に映像と音のサービスとして抽象化され、演出と入出力を分離
  - 視聴者フィードバックをローカル再生環境または放送元へ送信

メンバー活動紹介

- インテリジェントマイク技術（NTT）
- 自由視点映像の生成技術 (KDDI)

NTT

Innovative R&D by NTT

DANTEを用いたマイクアレイ系（プロトタイプ）の構築

NTTメディアインテリジェンス研究所　音声言語メディアプロジェクト
丹羽 健太，齊藤 翔一郎

Copyright©2014 NTT corp. All Rights Reserved.

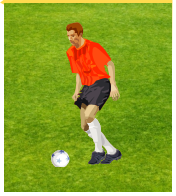
背景

NTTでは，欲しい音を欲しいユーザに届けるために，音声強調系の研究開発を進めている。

遠隔会議

特定の話者の
発言を拾いたい。

放送



キック音等を
拾いたい。

車

走行音がある中でも
音声コマンド入力したい。

工場

高騒音下でも
音声通話したい。

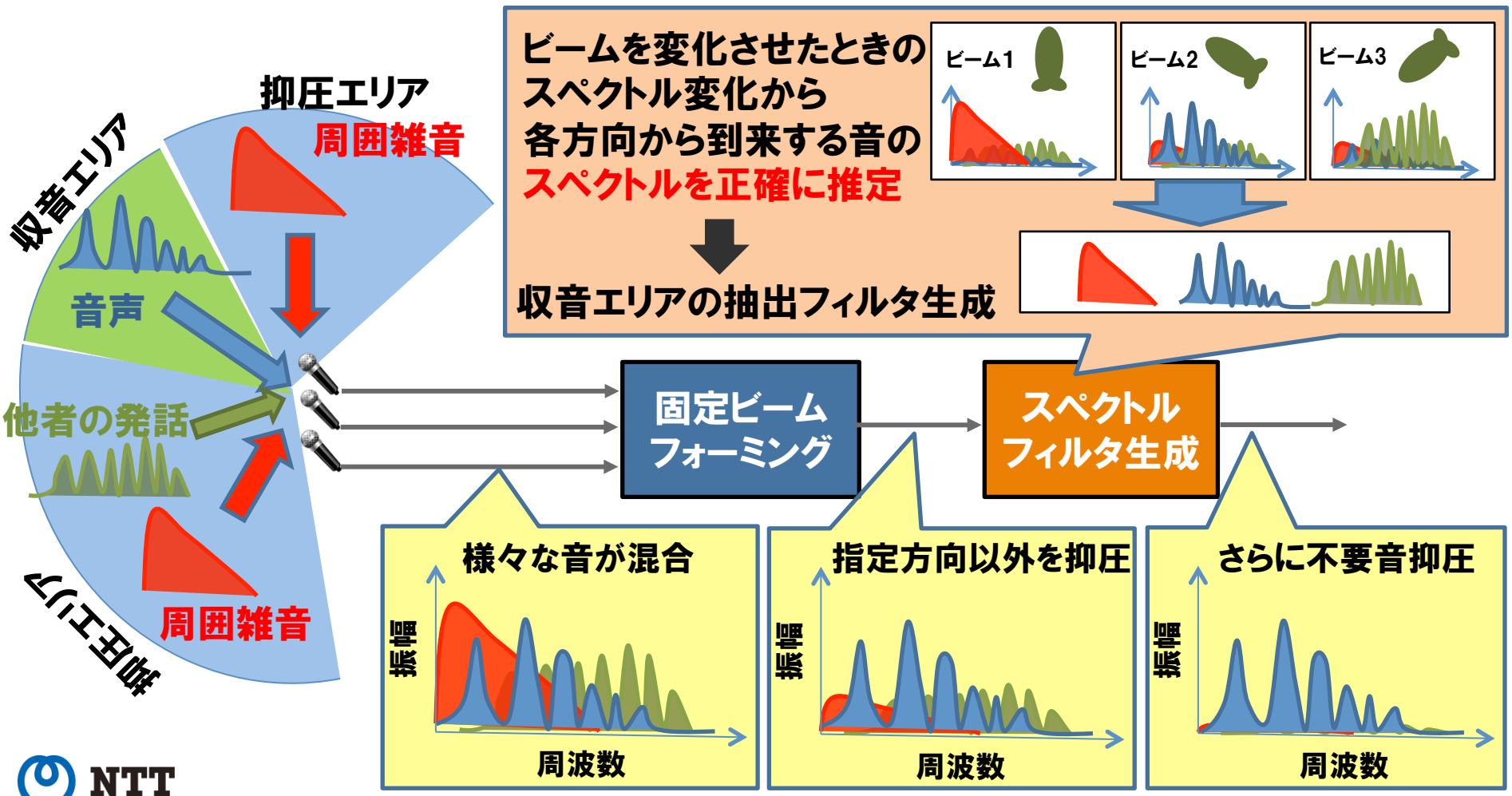
ロボット

話者を区別して
音声認識したい。

Copyright©2014 NTT corp. All Rights Reserved.

NTT技術の特徴

マイクアレイを利用し，指定した方向の音声のみをクリアに収音。
高い周囲音抑圧性能と，正確なスペクトルの復元を両立。



Copyright©2014 NTT corp. All Rights Reserved.

プロトタイプ系の構築

Innovative R&D by NTT

単一指向性マイク3本で収音した信号を音声強調処理して出力する系を構築（DANTEベースで動作）

NTT

Copyright©2014 NTT corp. All Rights Reserved.

（参考）ユーザインタフェースに関するKDDI研究所の取組事例

大空間3D映像の リアルタイム映像合成技術

スタジアム規模の大空間3D映像を自由な視点から視聴可能

【動作原理】

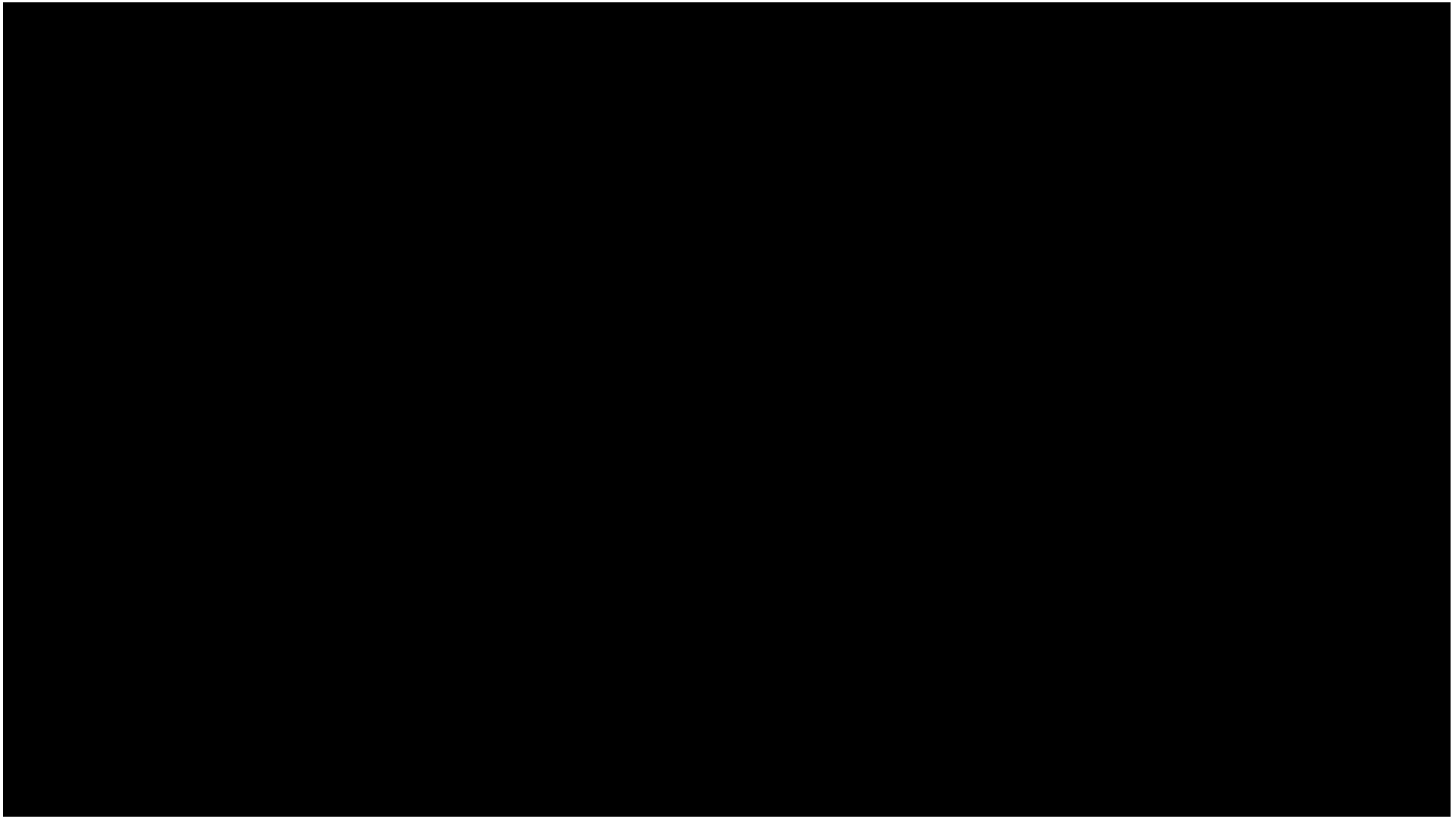
- 複数台カメラで撮影した映像を配信サーバに蓄積
- ユーザの選択したアングルに応じて、合成に必要な映像のみ抽出
- 配信受信側で、3D自由視点映像を合成して視聴

複数台のカメラで撮影

様々な視点から視聴可能

出典：情報通信審議会 情報通信政策部会 研究開発戦略委員会 「ICT分野における今後の研究開発戦略」

スポーツ自由視点映像の生成



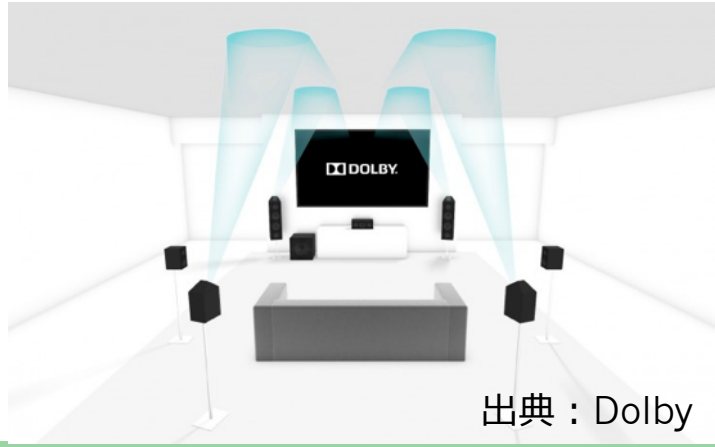
出典：Youtube サッカー中継を自由なアングルで見れる スタジアム3D自由視点技術：DigInfo

WG活動紹介

- 立体音響
  - Dolby Atmos
  - Virtual Sound Scape System
- 協力関係の構築
- 技術標準化動向

Dolby Atmos

- 2012年4月, 劇場版の登場
- 2014年6月, ホームシアター版の登場
- 技術的特徴
  - 環境音は従来のチャネルベース、独立した音響要素をオブジェクト管理
  - 音響設備に応じてリアルタイムレンダリング
  - 最大128の音響要素を口スレスで記録、最大64の独立スピーカーに出力可能
  - ホームシアター版は天井にスピーカーを設置しないオプションも
- 普及動向（2015 Jan）
  - 200本以上の映画が対応
  - 全世界900以上のスクリーン
  - 日本でも8劇場＋2予定 (2014.9)



出典：Dolby

Virtual Sound Scape System (VSSS)

- バンダイナムコスタジオとの協力
  - 8月21日 SDMミーティングに招待
  - 10月17日 開発環境兼デモブースを見学
- VSSS
  - 最先端の音響演出を再現・体感できるイベントおよび施設向けの音響制御システム
  - バンダイナムコのゲーム製品で実際に使用されている音響制御プログラムを応用
  - プレイヤーの操作に応じて即時に反応するためのインタラクティブな技術
  - Windows上で動作するアプリケーション

©BNGI

図の出典：(株)バンダイナムコスタジオ
第1回ライブ&イベント産業展出展資料

AEONとの協力

- 4月9日 打ち合わせ
- 5月9日イオン幕張の劇場を訪問
  - Dolby Atmosの体感ビデオ
- SDMとの協力関係
  - AEONは劇場の資産価値向上のため、映画とは別にどんなことができるかを考えている
    - ライブイベント
    - パブリックビューイング
    - 音楽教室など
  - AEONは80館で上映できる出口を持ち、興業のコミットメントができる

パイオニア、Intelとの協力

- パイオニアとの協力関係
  - 由雄 淳一様（IEC TC100 TA11 Quality for audio, video and multimedia systems議長）
  - 7月3日 東京大学にて打ち合わせ
  - 9月11日 WIDE合宿にて招待講演
- Intelとの協力関係
  - 8月29日　中川様、Roberts様と打ち合わせ
  - CamCom：イメージセンサを用いた高精度室内位置特定
  - 2015年1月からIEEE標準化開始(802.15)

研究協力関係

- インドIIT-Hにおいてスマートキャンパス・プロジェクト向け提案として、新キャンパスの建物への導入を計画中（東京大学）

- ロンドン大学クイーン・メアリーのデジタル音楽センターと共同研究でSDMと連携（YAMAHA）

- EUパートナーとSDMに関連する共同提案を模索中（東京大学）

SDMの構成と各組織の強み

SDM WGコアメンバ

東京大学 THE UNIVERSITY OF TOKYO — IPv6とネットワーク仮想化

Keio University 1858 CALAMVS GLADIO FORTIOR — メディアデザイン

YAMAHA — オーディオ製品と音響の研究

DOLBY — 立体音響のオブジェクト管理

NTT — インテリジェントマイク技術

KDDI R&D LABS — 自由視点映像

Panasonic ideas for life — オーディオ製品とオリンピック2020公式スポンサー

展開パートナー

AEON — 80以上の劇場

BANDAI NAMCO — ゲーム開発環境とイベント向け音響制御システム

intel

Pioneer

研究パートナー

Queen Mary University of London centre for digital music

MANDAT INTERNATIONAL

UPMC PARIS UNIVERSITAS

भारतीय प्रौद्योगिकी संस्थान हैदराबाद Indian Institute of Technology Hyderabad

uni.lu UNIVERSITÉ DU LUXEMBOURG

オーディオ・ビジュアル標準化動向

MPEG-2 (H.222/H.262, ISO/IEC 13818)
ISO/IEC 13818-1
ISO/IEC 13818-2
ISO/IEC 13818-3
ISO/IEC 13818-4
ISO/IEC 13818-5
ISO/IEC 13818-6
ISO/IEC 13818-7
ISO/IEC 13818-8
ISO/IEC 13818-9
ISO/IEC 13818-10
ISO/IEC 13818-11
ISO/IEC MPEG (ISO/IEC JTC1/SC29/WG11)
IEC TC 100

DVB-T2
DVB-C2
DVB-S2
DVB-S2X
DVB-CI-Plus
DVB-RCS2
DVB-GEM

Audio Video Bridging (AVB)

IEEE-1722.1
IEEE-802.1Q
IEEE-1722
IEEE-1733
IEEE-802.1AS
IEEE-802.1BA
IEEE-802.1Qat
IEEE-802.1Qav
IEEE-1588

rfc3190
rfc3550
rfc3551
RTP Media Congestion Avoidance Techniques (rmcat)

今後の活動予定

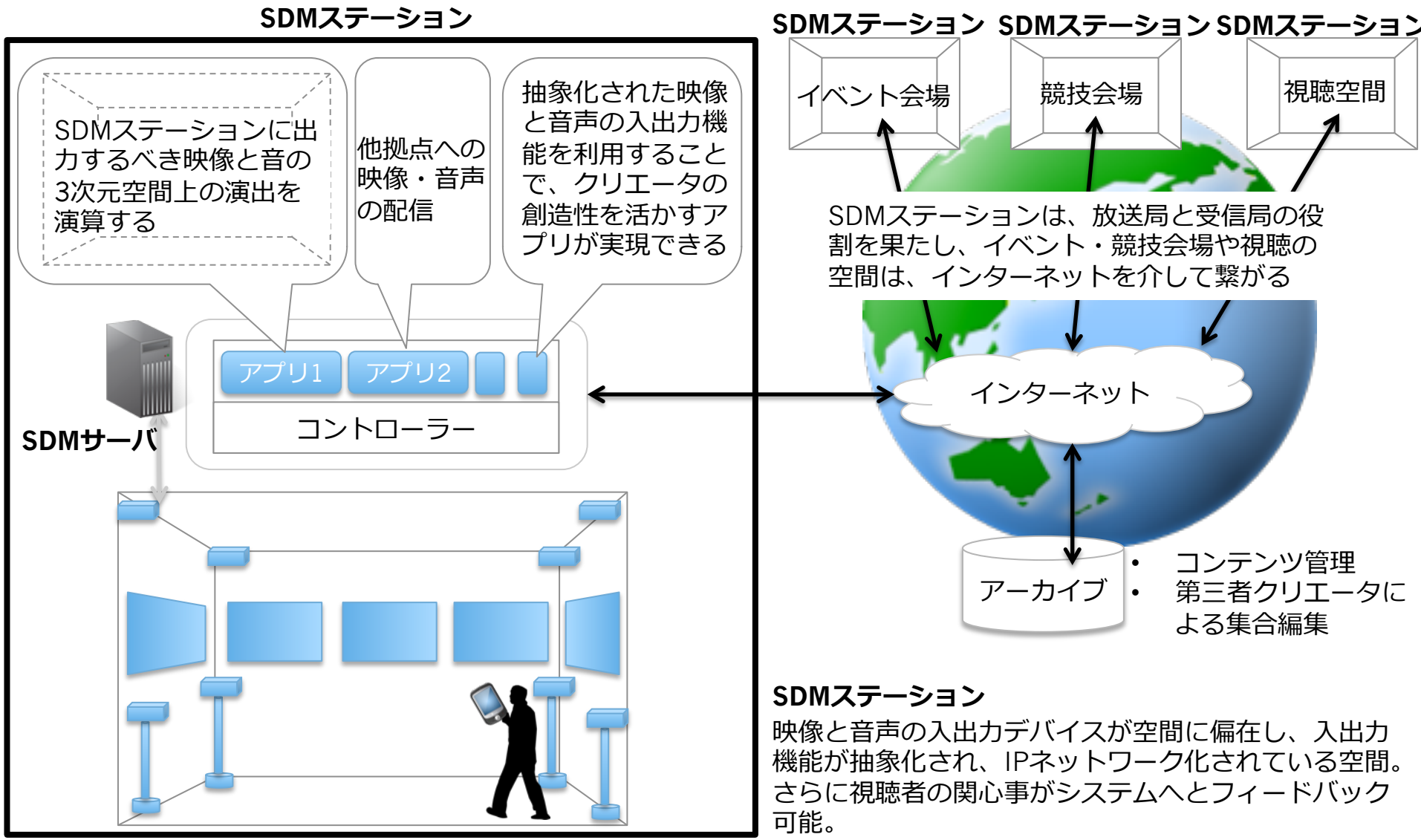
デモお披露目計画

- **2020年 8月 東京オリンピック**
- 2019年 9月 ラグビーワールドカップ@日本
- 2018年 6月 FIFAワールドカップ@モスクワ
- **2016年 8月 リオデジャネイロオリンピック**
- **2015年 11月 第49回 IETF@横浜**
- 2015年 9月 WIDE秋合宿
- 2015年 6月 Interop
- 2015年 3月 WIDE春合宿

デモの方針

- SDMのコンセプトを実証する
  - 3次元の映像・音声の演出をソフトウェアで制御
  - 映像/音声を複数ソースのミキシング可能
  - 視聴者の関心事 への集中可能性
- 以下を考慮した具体的なデモ・シナリオ案を考える
  1. SDMアーキテクチャの良さが伝わるデモ
  2. 新技術の意図が伝わりやすいもの
  3. 映像があるとよい
  4. 集音のデモもあるとよい
  5. スマホなどを使ったインタラクティブなデモも望ましい
- デモの対象
  - IETFのデモ：研究者/技術者が対象 （技術的側面）
  - パブリックビューイング：一般の人が対象（楽しい側面）

デモの大枠

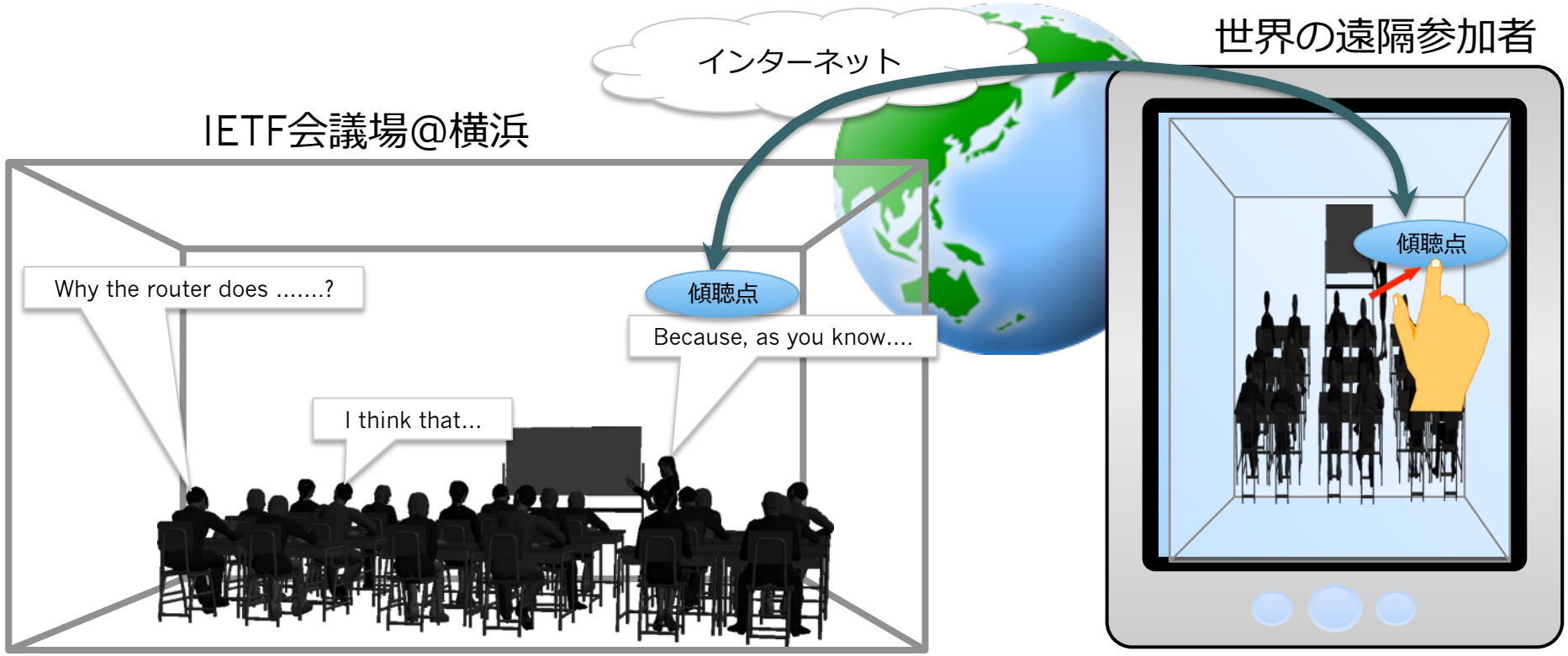


SDMステーション

映像と音声の入出力デバイスが空間に偏在し、入出力機能が抽象化され、IPネットワーク化されている空間。さらに視聴者の関心事がシステムへとフィードバック可能。

IETF遠隔参加者を支援する傾聴点指定システム

- IETFは音声での遠隔参加が可能であるが、マイクで発言が拾えない場合や、私語の騒音で発言が聞きにくい問題がある
- この問題を解決するため、遠隔参加者が音声が必要な場所を指定するシステムを開発する
- 時差に配慮しアーカイブで過去の音を再現

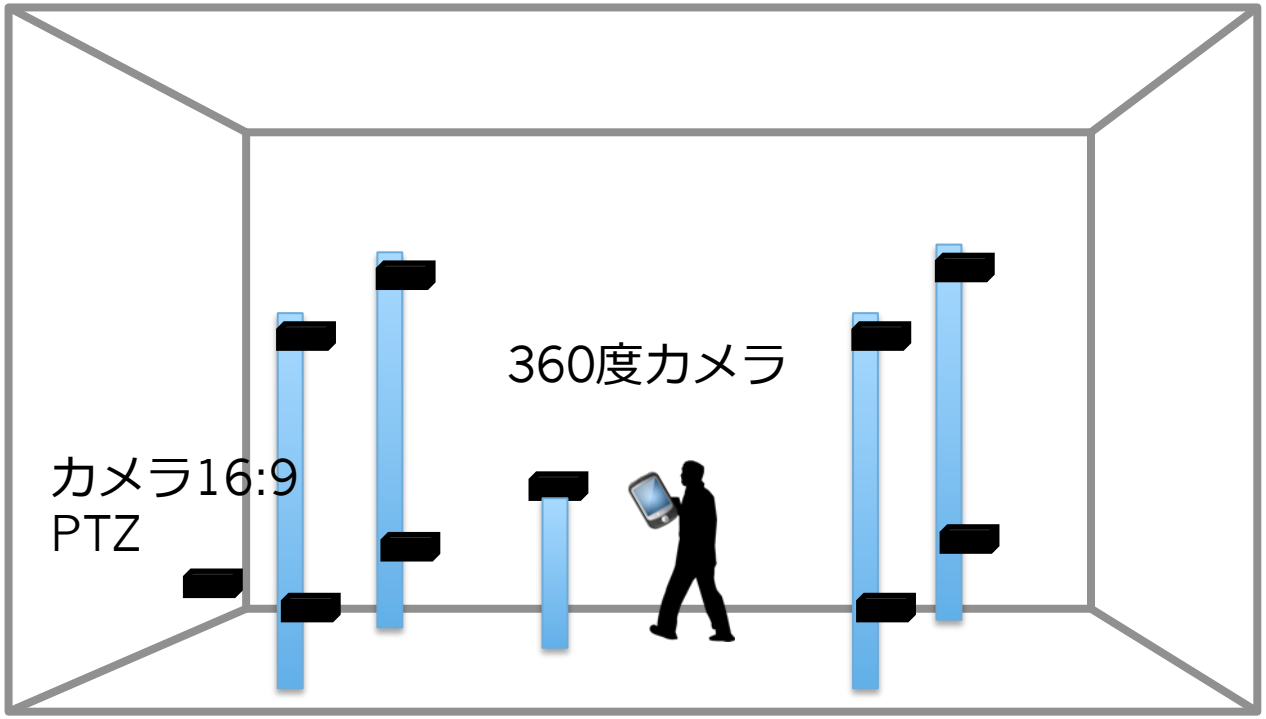


実験環境

- 東京大学 弥生キャンパス I-REF棟6Fヒロビー
  - スピーカー（最小１０） = 5.1 ch + 上空に4つ
  - カメラ　（360度カメラ、カメラ16:9, PTZカメラ）
  - 持ち運び可能なシステム

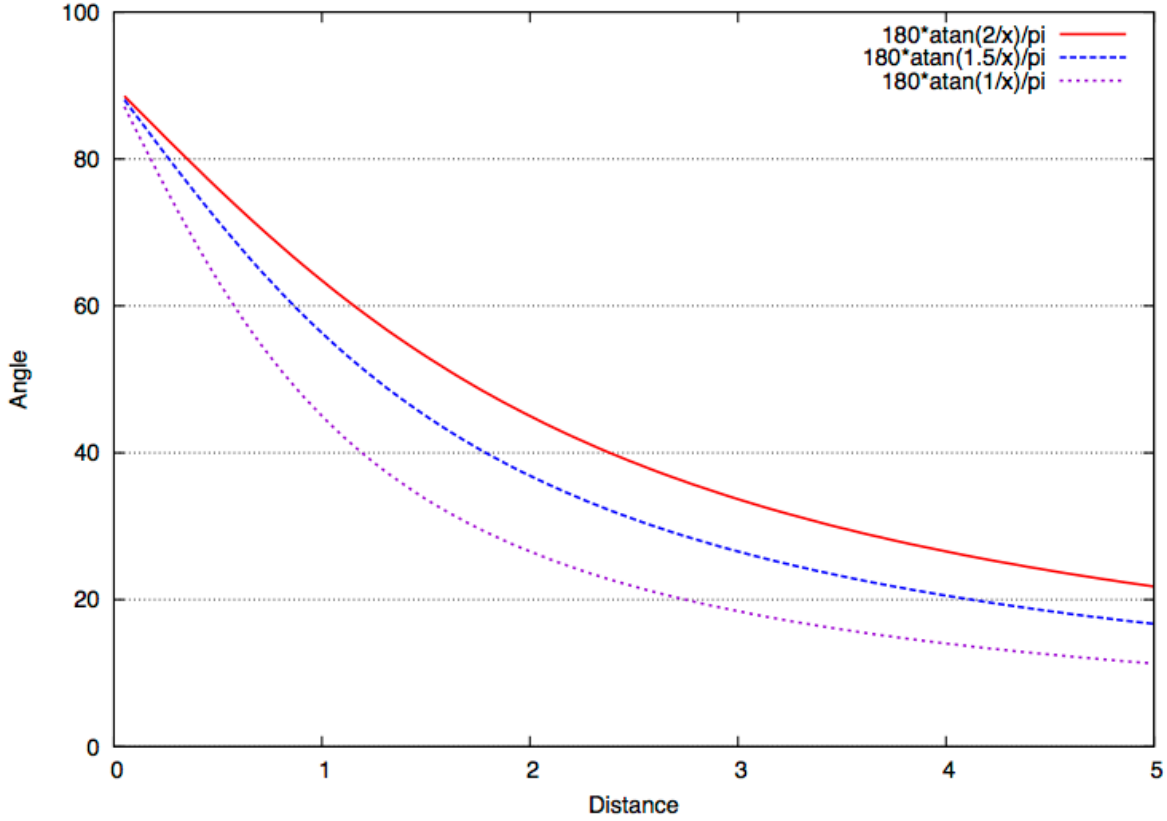
実験環境

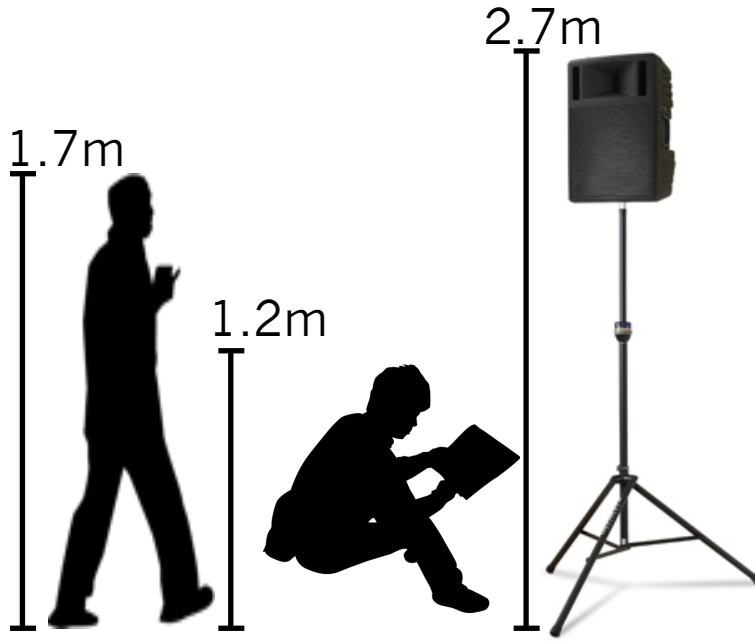
- 東京大学 弥生キャンパス I-REF棟6Fヒロビー
  - スピーカー（最小１０） = 5.1 ch + 上空に4つ
  - カメラ　（360度カメラ、カメラ16:9, PTZカメラ）
  - 持ち運び可能なシステム



高さとスピーカーの関係

- 着席中の耳の高さ1.2mでスピーカーの高さ2.7mの場合、距離が2mほどで角度が40度を下回る





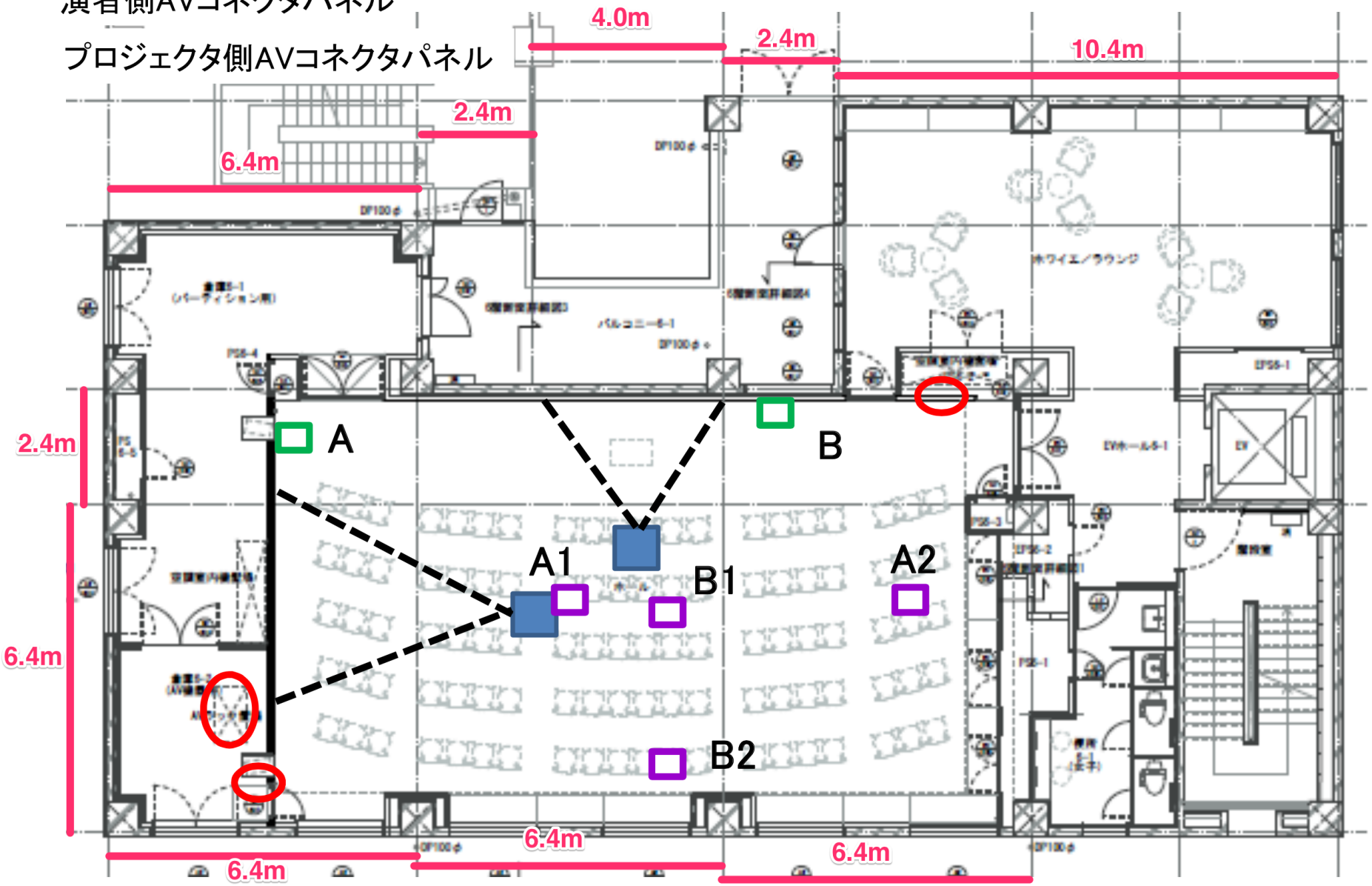
ヒロビー配置図

AV機器操作パネル

プロジェクタ設置場所

演者側AVコネクタパネル

プロジェクタ側AVコネクタパネル



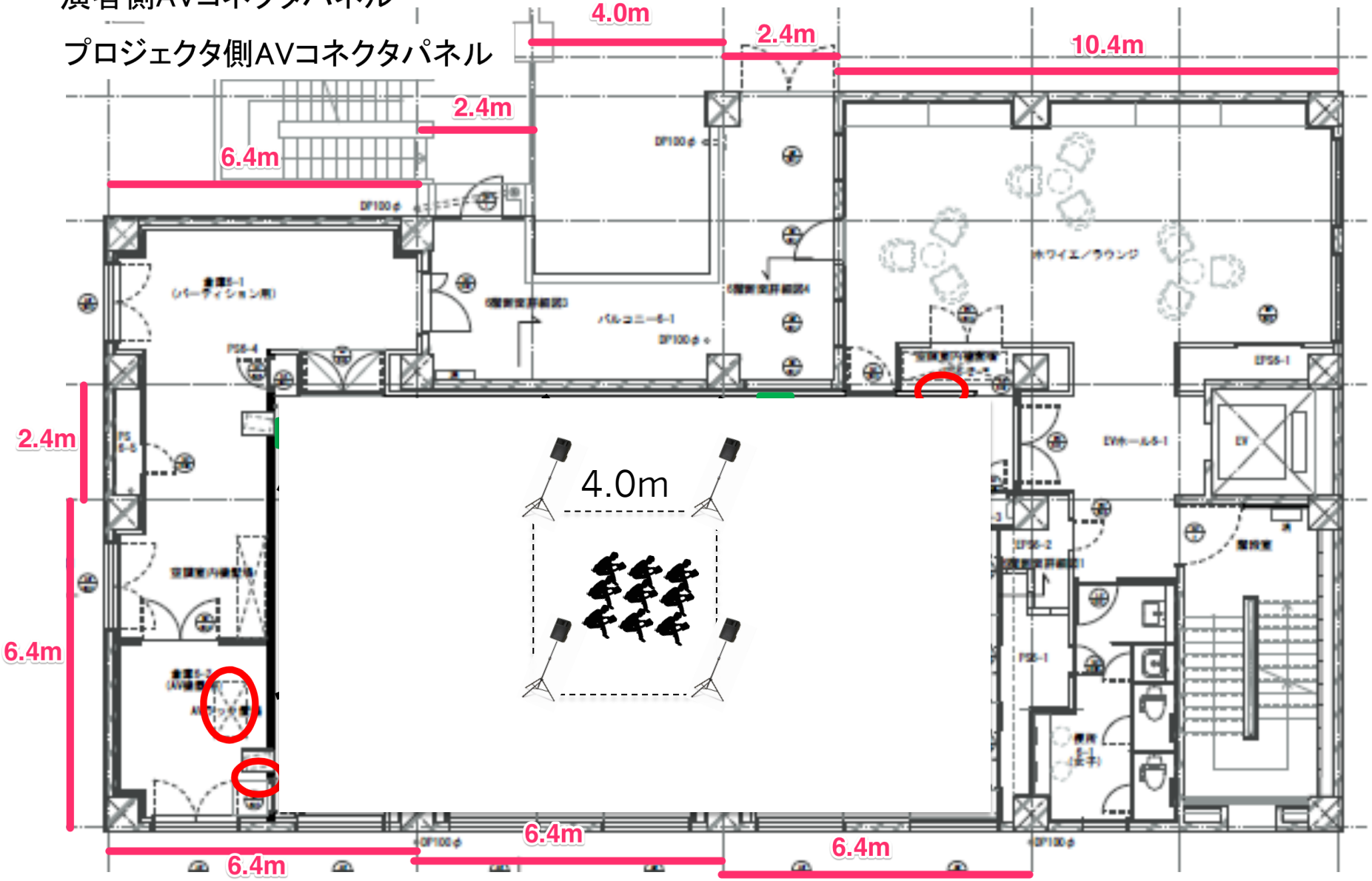
ヒロビー配置図

AV機器操作パネル

プロジェクタ設置場所

演者側AVコネクタパネル

プロジェクタ側AVコネクタパネル



まとめ

- サービスとしての映像音響を提供するためSDM WGを2014年はじめに開始
- 目指すシステムの方向性が見えてきた
  1. 3次元の映像・音声の演出をソフトウェアで制御
  2. 映像/音声を複数ソースのミキシング可能
  3. 視聴者の関心事 への集中可能性
- SDMのコンセプトを実証するデモンストレーション
  - 2015年11月 第49回 IETF@横浜 （技術的側面）
  - 2016年, 2020年　オリンピック （楽しい側面）
- 協力できるパートナーとの関係を構築

- 今後の予定
  - プレイグラウンド用の実験環境の構築
  - デモンストレーション開発
  - 関連技術の標準化動向をまとめる

SDM WGへの参加

- 興味がある人募集
  - SDMの環境構築
  - SDMで実現できるアプリケーション
  - SDMのデモと実験
  - SDM分野のサーベイ

- 月1回のミーティング
- sdm-wg-request@wide.ad.jp

以上