



Immersive Sound for Cinema

Standards Update Webcast



© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

SMPTE Standards Update Webcasts

- Series of quarterly 1-hour online, interactive webcasts covering advances in SMPTE's standards development work
- Free for everyone!
- Sessions are recorded for on-demand viewing convenience
 - <https://www.smpte.org/standards-webcasts-on-demand>

2

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Your Host

Joel E. Welch
Director of Education
SMPTE



3

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Today's Guest Speaker

Pete Ludé

Senior VP
RealD Inc.

Chair, SMPTE TC-25CSS10
Working Group on
Interoperable Immersive
Sound Systems for Digital
Cinema



4

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Standards Update Immersive Sound for Cinema

- Today's Goal:
 - Provide a briefing on the substantial progress being made in standards within the SMPTE Technology Committees
 - Brief tutorial on “what” and “why” of Immersive Cinema Standards
- Scope:
 - Digital Cinema
 - Not home content distribution

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Today's Experts



Dr. Ton Kalker

VP, R&D Emerging Technology
DTS, Inc.



Charles Q. Robinson

Researcher, Sound
Technology
Dolby Laboratories

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Today's Topics

- Some Definitions
- Background / History
- Decoding the SMPTE Standards Committees
- Why / How Immersive Sound Standards? - Dr. Ton Kalker
- Spatial Audio Scene Description - Charles Q. Robinson
- Summary and Q&A

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Definitions

Thanks to Brian Vessa!

- **SOUND**
 - A pressure disturbance in air that can be perceived by hearing, and thus is “audible”
- **AUDIO**
 - An electrical representation of sound. Sound is converted to audio via a transducer (such as a microphone) and audio is converted into sound by a transducer such as a loudspeaker
 - Audio itself cannot be perceived by hearing, and thus is not audible

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Definitions

- **SOUNDFIELD**
 - The acoustical space created by simultaneously reproducing one or more audio sources
 - Traditional soundfields are two dimensional in the horizontal plane
- **SOUNDFIELD CONFIGURATION**
 - Defined arrangements of configuration of loudspeakers that convey the intended Soundfield

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Definitions

- **IMMERSIVE SOUNDFIELD**
 - An acoustical space where sound can be reproduced from all three dimensions
- **IMMERSIVE SOUNDFIELD CONFIGURATION**
 - A defined arrangements of configuration of loudspeakers that conveys an Immersive Soundfield

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Definitions

- **IMMERSIVE SOUND**
 - Sound that is experienced within an Immersive Soundfield
- **IMMERSIVE SOUND SYSTEM**
 - A sound system that is capable of producing Immersive Sound
- **IMMERSIVE AUDIO**
 - Audio that is designed to feed an Immersive Sound System
 - Audio that is created with the intent of being reproduced as Immersive Sound

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Definitions

- **CHANNEL**
 - Distinct collection of sequenced audio samples that is intended for delivery to a single loudspeaker or an array of loudspeakers
 - May contain encoding that will allow enhanced reproduction by a number of additional loudspeakers in a designated manner

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Definitions

- **OBJECT:**

- Audio of any duration that has associated metadata that describes how it is to be reproduced within a soundfield
- This metadata describes the position, spread, motion characteristics and other rendering information
- Objects can move within the soundfield, be reproduced in a single position, or by a specific loudspeaker
- An object may be defined to behave very much like a channel

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Definitions

- **CHANNEL-BASED AUDIO ESSENCE (CBAE)**

- A collection of audio channels that conveys the information necessary to create an intended soundfield
 - Current Digital Cinema standards specify up to 16 channels in the Main Audio track file
 - CBAE is delivered via AES3

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Definitions

- **OBJECT-BASED AUDIO ESSENCE (OBAE)**
 - A collection of audio objects that conveys the information necessary to create an intended soundfile and to place particular sounds within it
 - There are currently no Digital Cinema standards in place for utilizing OBAE in the Dcinema architecture

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Background

- Immersive Sound has become popular in Cinema
 - Over 1,000 auditoriums
 - Over 200 titles
- Studios and Exhibitors want “single inventory”
 - One size fits all

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Digital cinema object-audio addendum

Digital Cinema Initiatives, LLC (DCI)

Digital Cinema Object-Based Audio Addendum

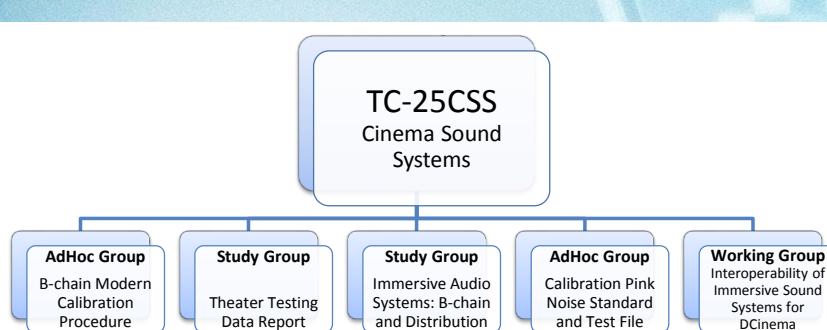
1. Introduction

Object-based audio formats that utilize new methods for delivering audio essence to the theater have come to market. This document is a specification for packaging, distribution and theatrical playback of object-based motion picture D-Cinema audio content that exceeds the delivery capability of the Digital Cinema Package (DCP) audio track file as defined in DCI's *Digital Cinema System Specification* (DCSS)¹.

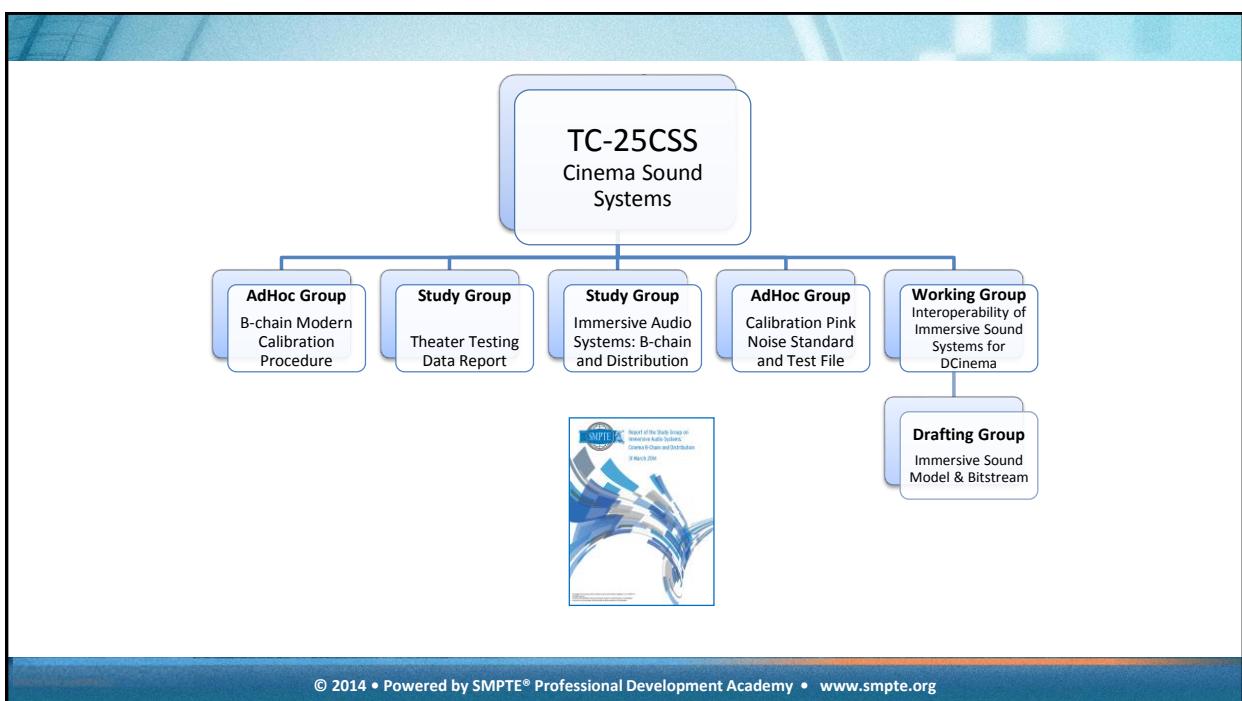
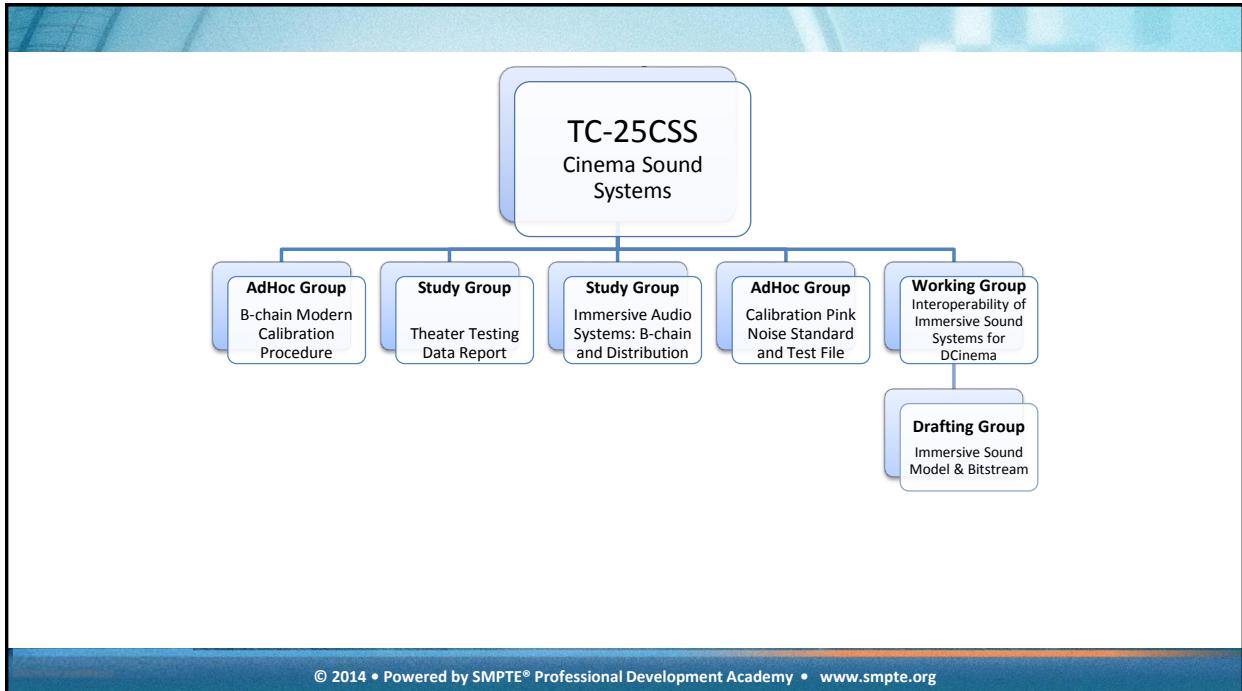
This specification provides requirements that will assure interoperability of object-based audio content while maintaining the current DCSS compliant architecture and KDM structure, recognizing that the D-Cinema industry is necessarily constrained by an existing installed base of equipment.

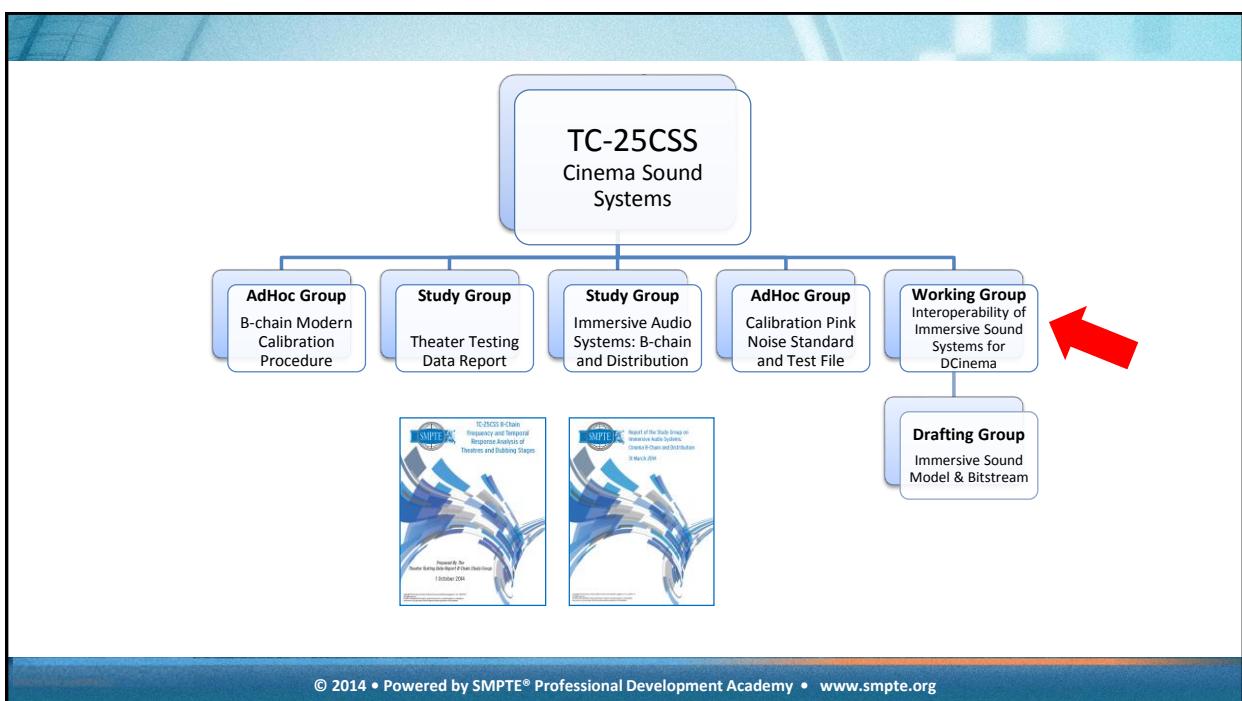
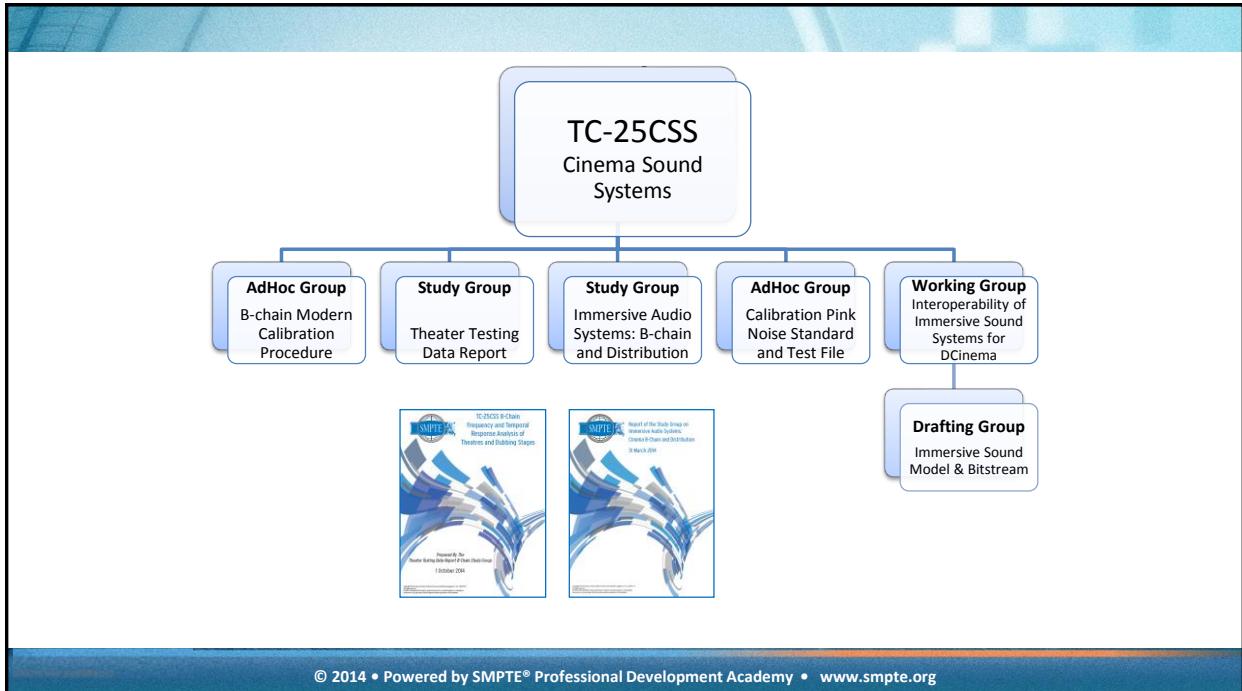
The rendering of object-based audio into a specific sound reproduction format is proprietary to manufacturing companies and is not addressed in this document. *Though object-based audio rendering may be carried out differently by individual systems, it is required that said audio in the DCP be interoperable within all DCSS compliant architectures that support object-based audio.*

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

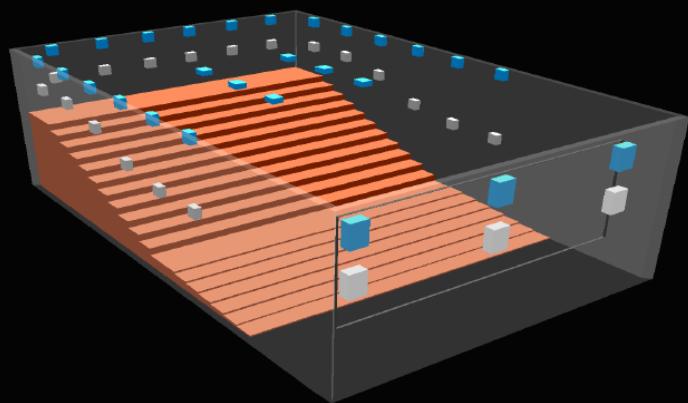


© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org



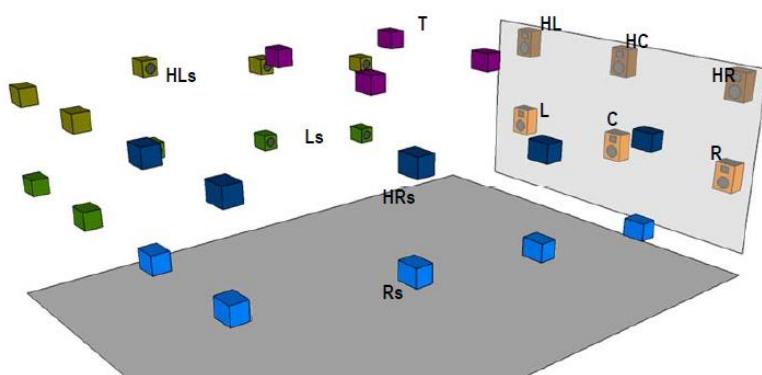


Auro 3D



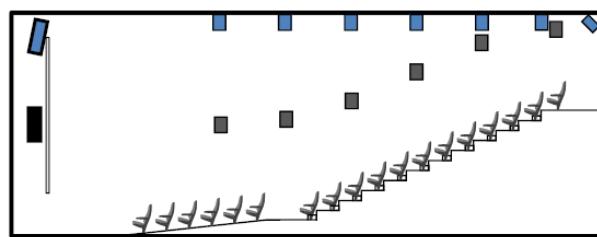
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Auro 11.1 - Alternative view



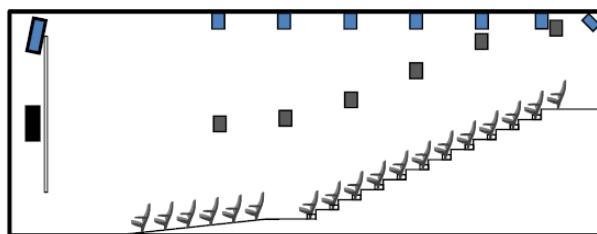
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Auro 11.1 - Side view



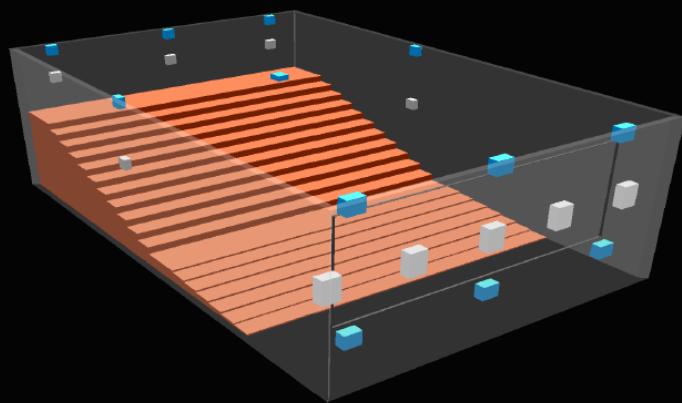
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Auro-Max - Side view



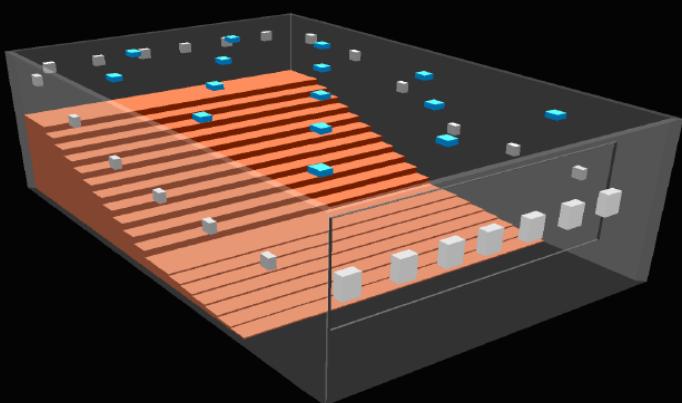
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

NHK 22.2



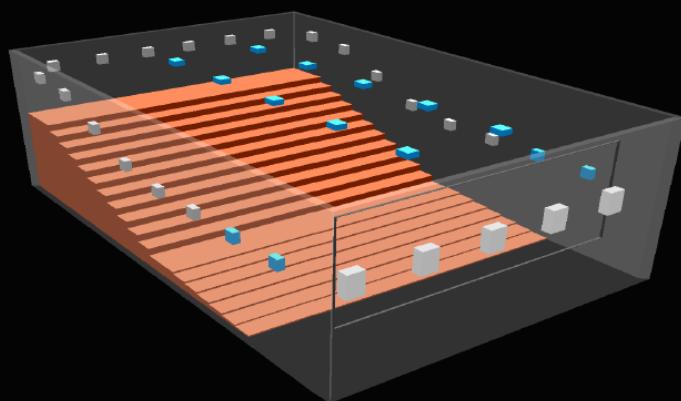
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Iosono



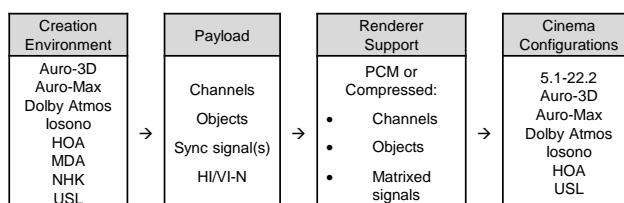
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Dolby Atmos



© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

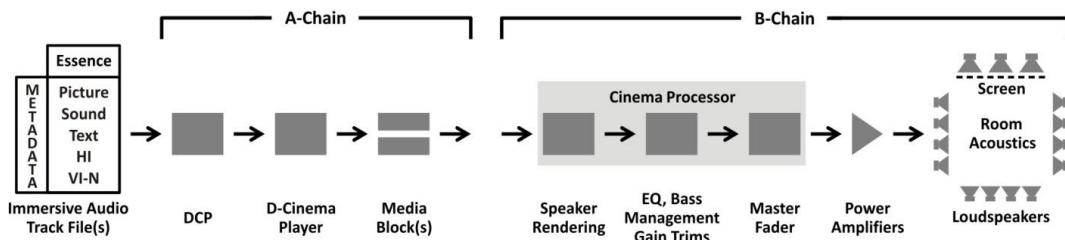
Overview of Immersive Audio Process



Key stages in proposed immersive audio process
(from Report from Study Group on Audio Systems)

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

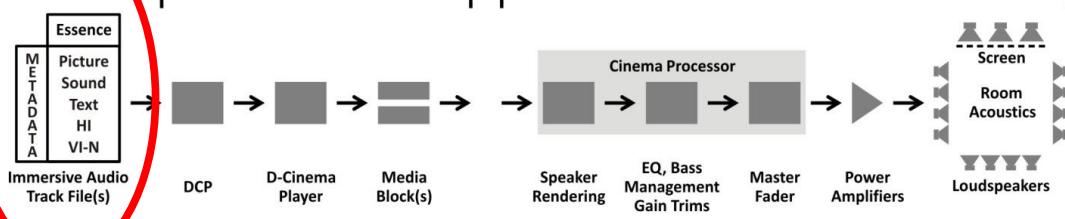
Ecosystem



© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Interoperable
Interchange

Ecosystem



© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Guest Speaker

Dr. Ton Kalker

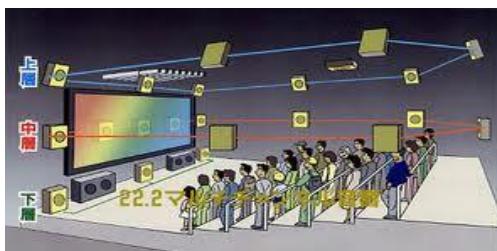
VP, R&D Emerging
Technology
DTS, Inc.



33

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Immersive audio in cinema is Very Young



- Do ticket buyers care? Is there ROI for exhibitors ?
- Yet to prove tangible value
- Success will require ease of use, affordability and inclusion

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Why object audio?

- Brings value to the cinema experience
 - Change the way we tell stories
- Liberates cinema audio creation
 - Focus on artistic intent
- Simplifies production/delivery chain and playback
 - Single package
- Flexible playback
 - In any environment
 - With any speaker configuration

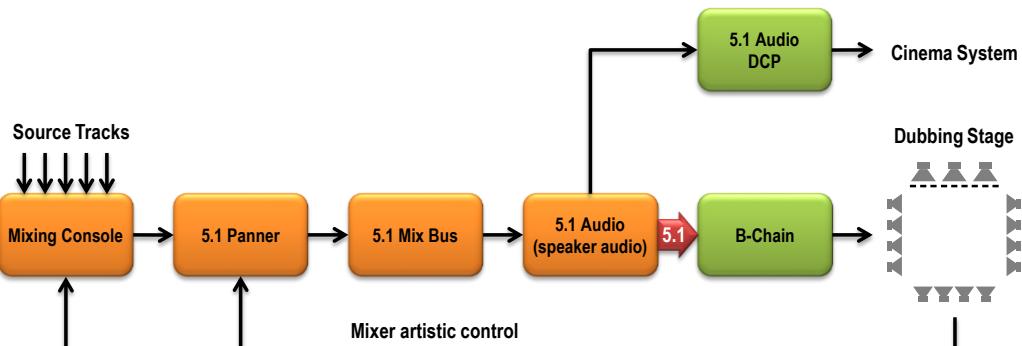
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Ideal workflow

- Create without speaker layout in mind
 - A single DCP as output
- Render in any layout configuration
 - Using single DCP
 - Preserving artistic intent

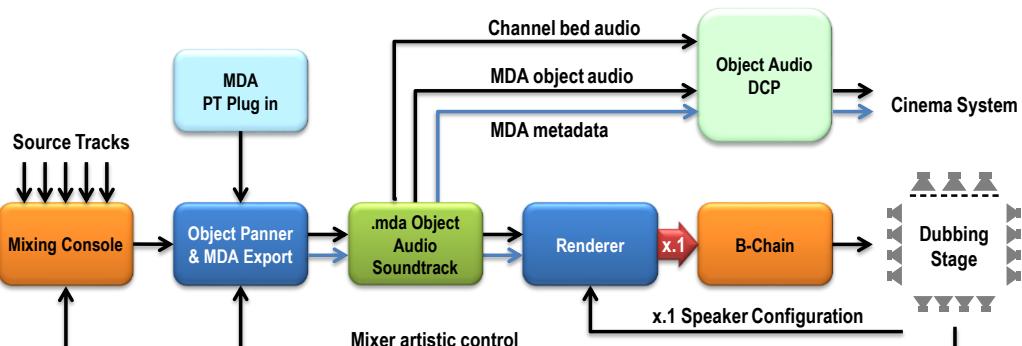
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Traditional Audio workflow (x.1, x=5)



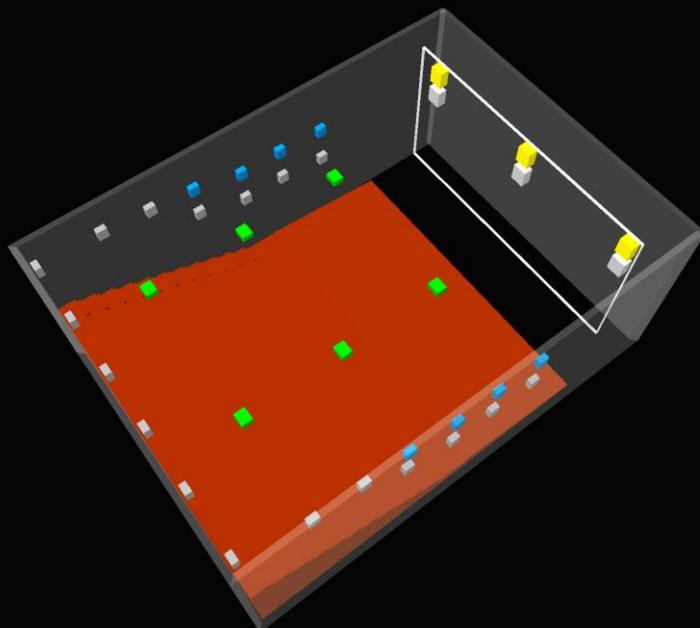
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Object audio workflow (MDA)

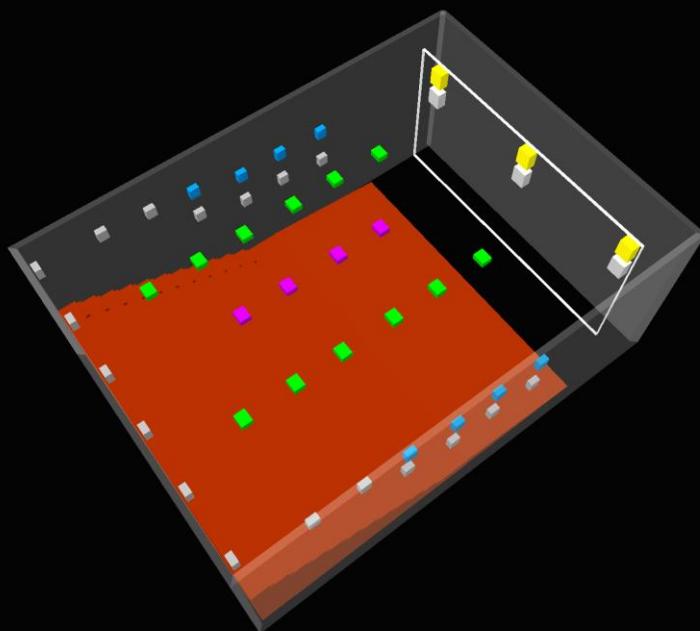


© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

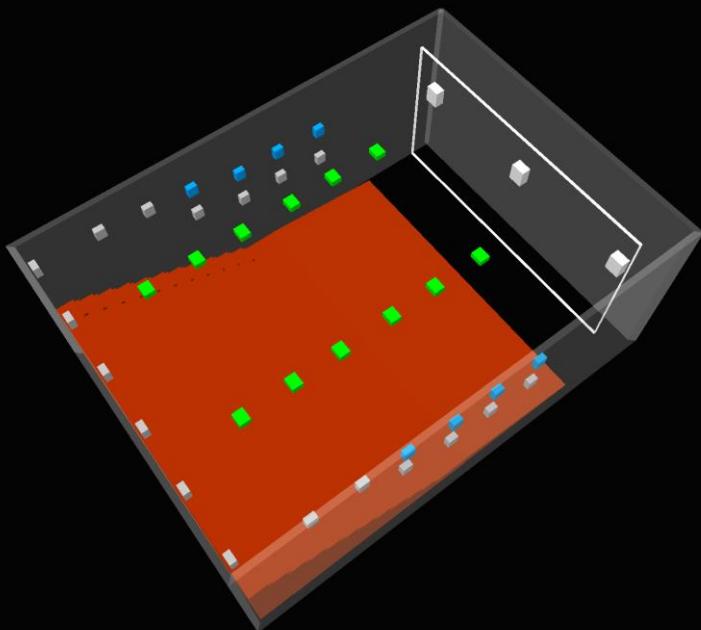
35.1



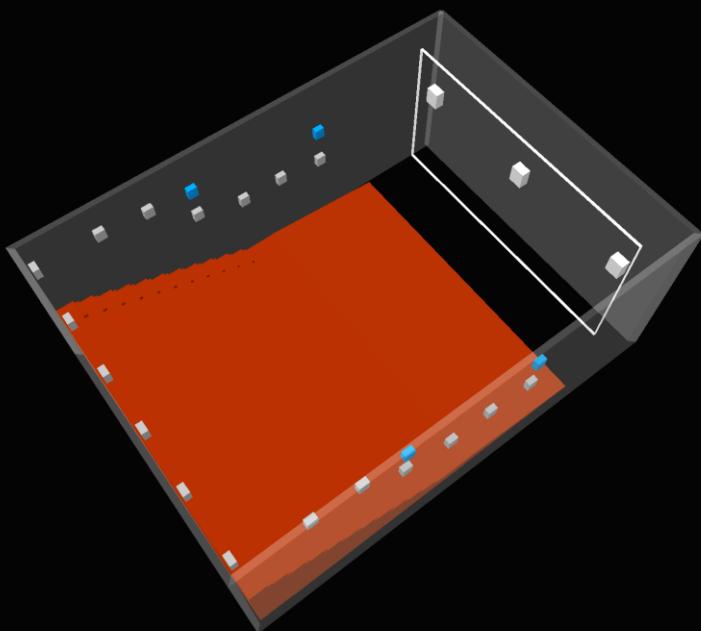
48.1



40.1



25.1



What stands in the way?

- Interoperability
 - Across industries: gaming, broadcast, cinema
 - Within cinema: multiple non-interoperable systems
- Tools
 - Early days ...
- Confidence
 - Will it play back as intended?
- Infrastructure
 - Upgrade? Costs?

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Guest Speaker

Charles Q. Robinson

Researcher, Sound Technology
Dolby Laboratories



44

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

Spatial Audio Scene Description

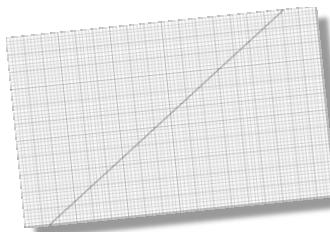
Charles Q Robinson



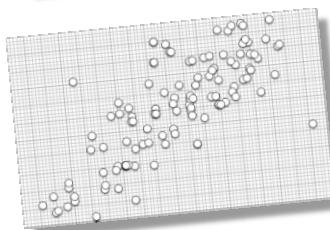
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

45

Principles



Practice



© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

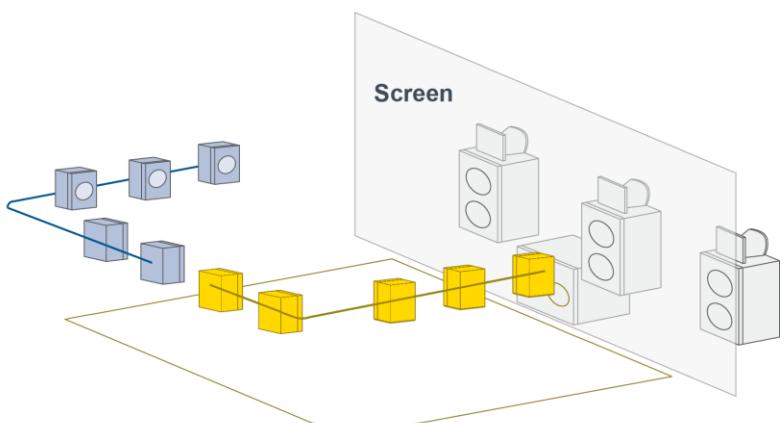
46

Scene Description

- The Scene Description specifies the physical attributes of the sound events
 - Audio waveform
 - Location
 - Spatial Extent (size)

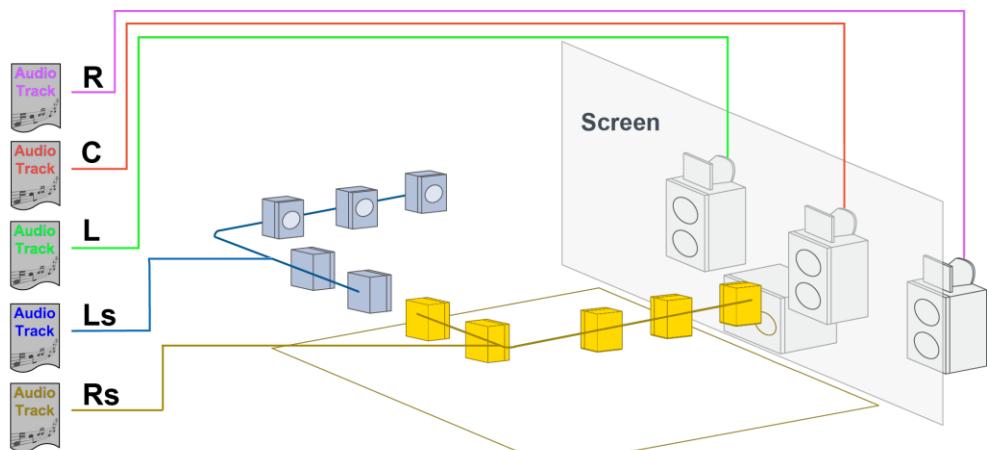
Scene Description

Channel-Based



Scene Description

Channel-Based

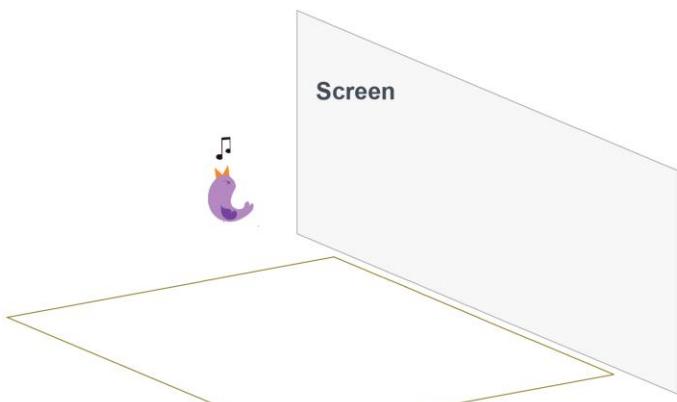


© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

49

Scene Description

Object-Based



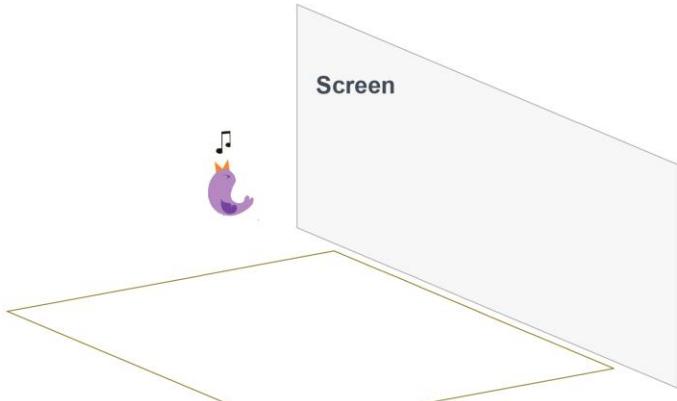
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

50

Scene Description

Object-Based

- Audio Signal



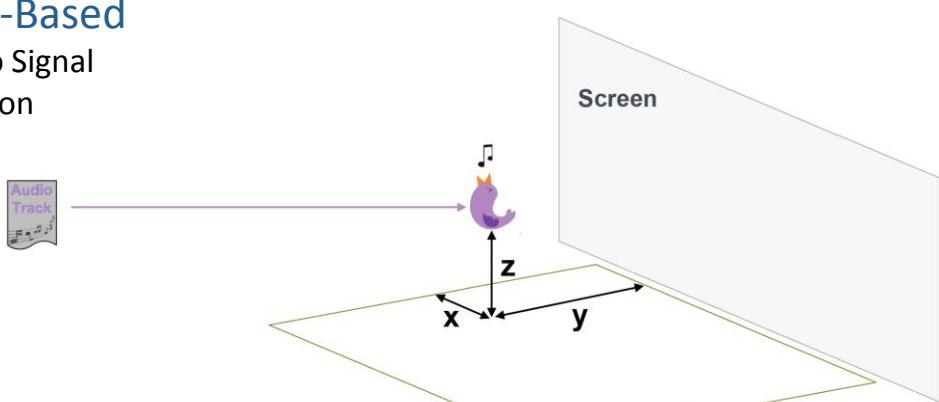
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

51

Scene Description

Object-Based

- Audio Signal
- Position



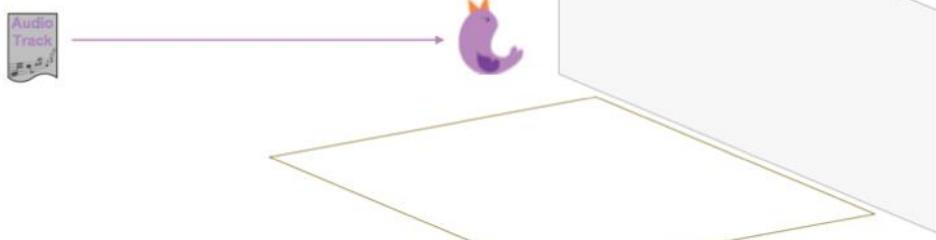
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

52

Scene Description

Object-Based

- Audio Signal
- Position
- Spatial Extent (apparent size)



© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

53

Scene Description

• Channel-based, Benefits

- Established work-flow, tools and techniques
- Easy to render and monitor
- Direct artistic control of loudspeakers (esp. Center, LFE)
- Efficient storage (many events in a small number of channels)

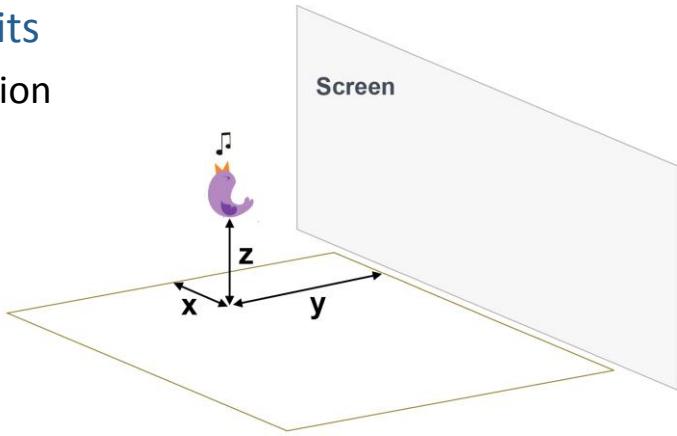
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

54

Scene Description

Object-Based, Benefits

- High spatial resolution



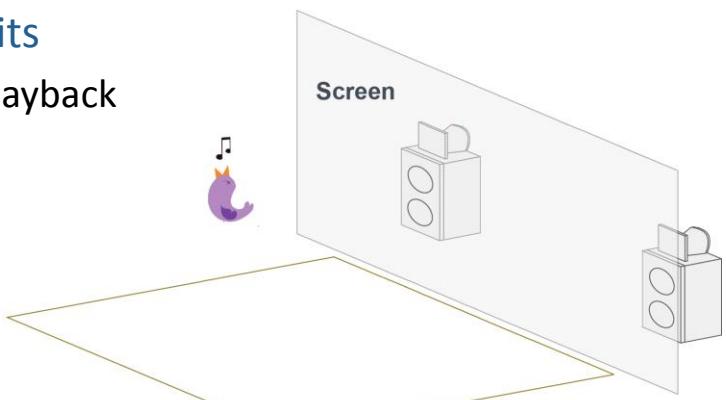
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

55

Scene Description

Object-Based, Benefits

- Flexible, Scalable Playback



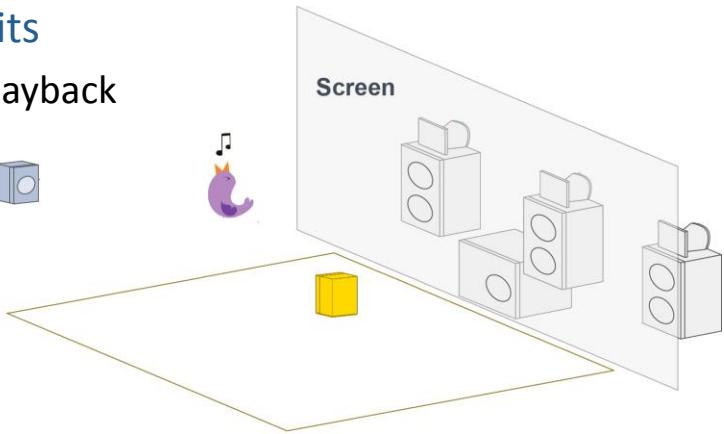
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

56

Scene Description

Object-Based, Benefits

- Flexible, Scalable Playback



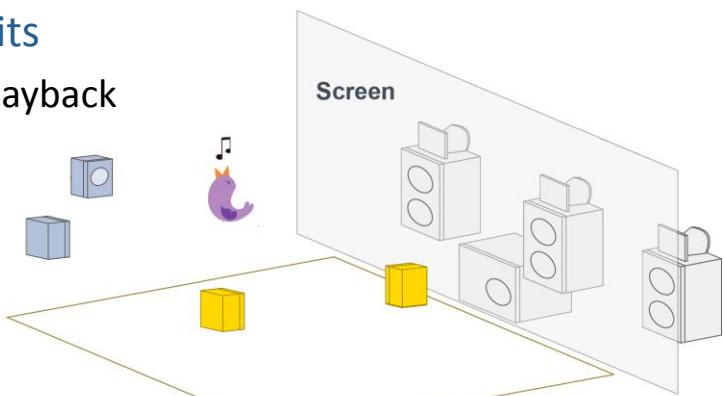
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

57

Scene Description

Object-Based, Benefits

- Flexible, Scalable Playback



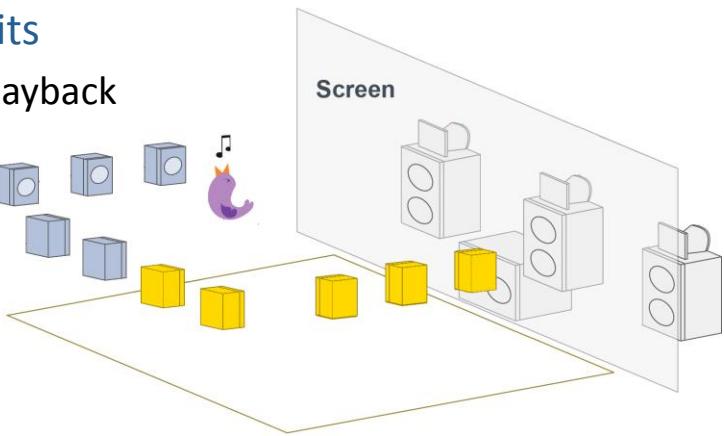
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

58

Scene Description

Object-Based, Benefits

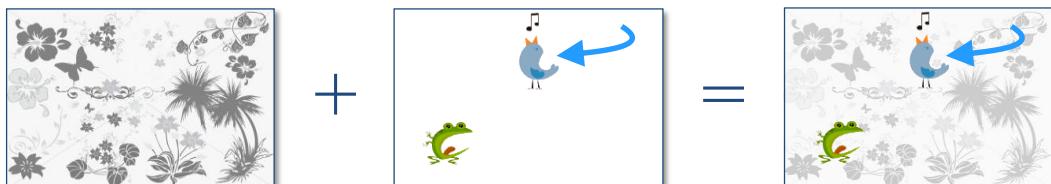
- Flexible, Scalable Playback



© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

59

Scene Description



Channels

- Direct control of key loudspeakers
- Complex Audio Textures
- Leverage existing tools & know-how

Objects

- Distinct, Dynamic Sounds
- Inherently scalable
- New artistic potential

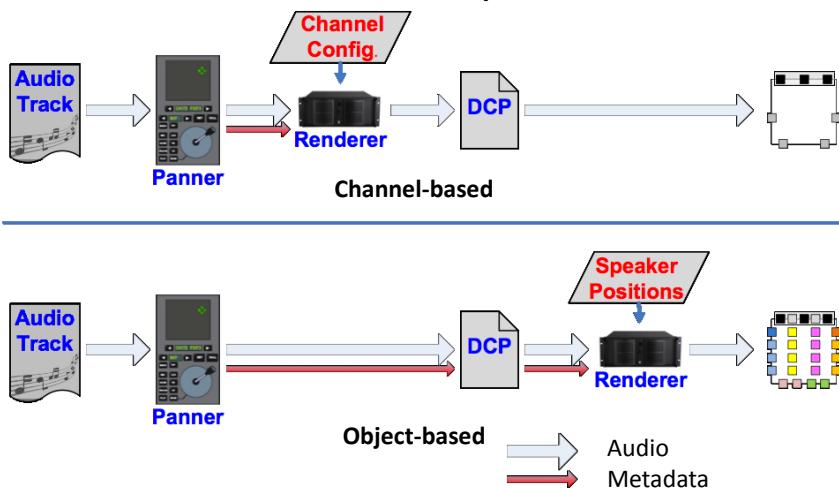
Channels + Objects

- Best of both:
- Envelopment and imagery**
- Mix once, play everywhere:**
- 5.1, 7.1, and beyond**

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

60

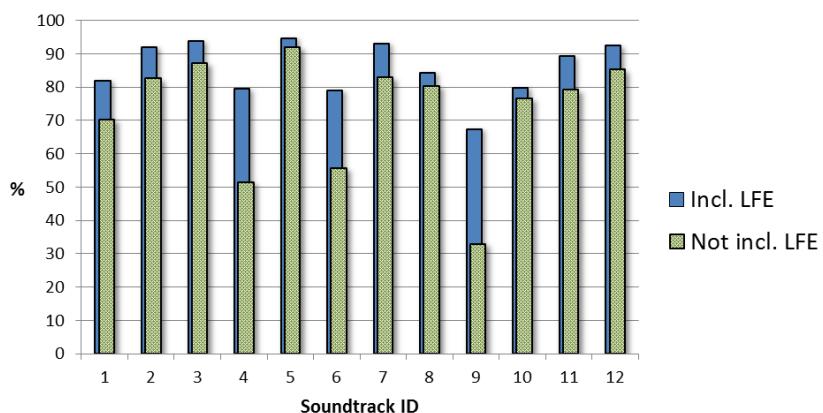
Scene Description

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

61

Scene Description, in Practice

Channel energy / total program energy

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

62

Location Specification

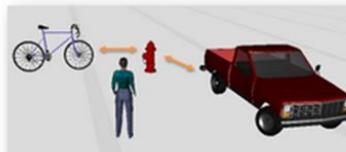
Egocentric (self-to-object)

Represents the location of objects in space relative to the body axes of the self (left-right, front-back, up-down).



Allocentric (object-to-object)

Encodes information about the location of one object or its parts with respect to other objects. The location of one object is defined relative to the location of other objects.

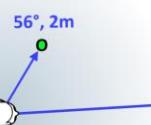


Harvard / NUS Mental Imagery and Human-Computer Interaction Lab

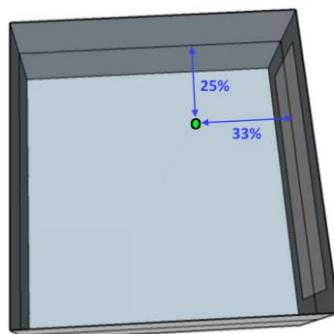
© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

63

Location Specification



Egocentric
Listener-based



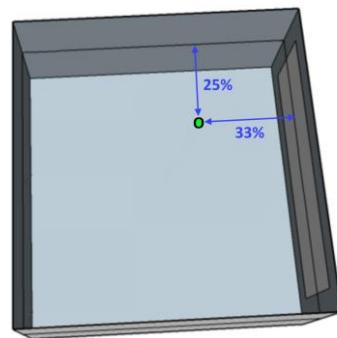
Allocentric
Auditorium-based

© 2014 • Powered by SMPTE® Professional Development Academy • www.smpte.org

*64

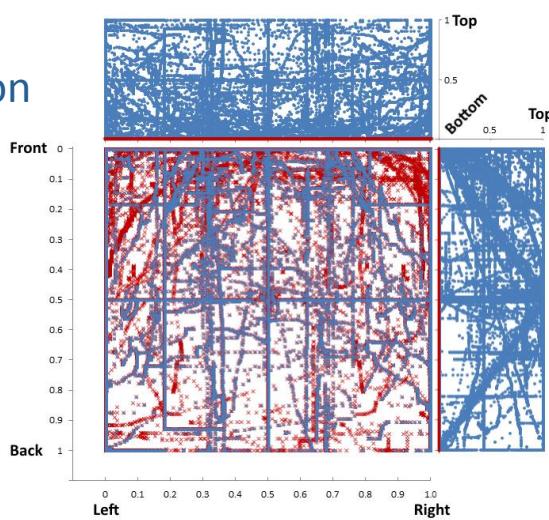
Location Specification

- SMPTE Immersive Sound Model
 - Allocentric Reference
 - Cartesian Coordinates
- To be determined
 - Resolution
 - Origin



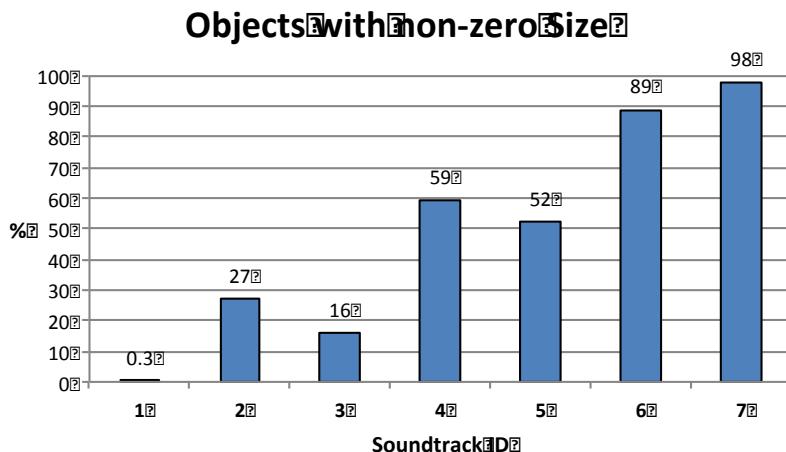
Scene Description, In Practice

Object Location



Scene Description, In Practice

Object Size



Summary

- Channels and Objects used extensively in ALL soundtracks analyzed
- Location:
 - Biased toward screen
 - Trajectories seem to be influenced by authoring tool automation and loudspeaker configuration
 - Best translation achieved by using the same system for authoring and playback.
- Spatial Extent (size) use was highly variable across titles analyzed

Q & A



Pete Ludé

Chair, SMPTE TC-25CSS10
WG on Interoperable
Immersive Sound Systems for
Digital Cinema



Dr. Ton Kalker

VP, R&D Emerging Technology
DTS, Inc.



Charles Q. Robinson

Researcher, Sound
Technology
Dolby Laboratories

Joel E. Welch

