

# CLUE Framework

IETF 82 Nov 14, 2011

Mark Duckworth

Allyn Romanow

Brian Baldino

Andy Pepperell

# Contents

- Summary of changes in current framework draft
- List of major issues
- Proposal for coordinate system
- Discussion of the other major issues
  - try to agree on a path to resolve them

# Changes since interim meeting

- These changes are included in draft-ietf-clue-framework-01
- Add some overview at beginning of section 5
- Clarify media capture is not just camera or microphone, it can be synthetic
- Capture set can have attributes for area of scene and millimeter scale
- Combine video composed and audio mixed attributes into a single more general attribute
- Clarify a provider can have more than one capture set
- Clarify left and right (i.e. camera left; audience left)
- Add new section on extensibility

# Issues

1. Coordinate system for spatial relationships
2. Source selection
3. Describing composed captures
  - Information from provider to consumer
4. Selecting composed captures
  - Requests from consumer to provider

# Issue 1 Coordinate System

- Proposal based on discussion in October meeting
- Define a coordinate system, area of capture, and point of capture

# What is this addressing

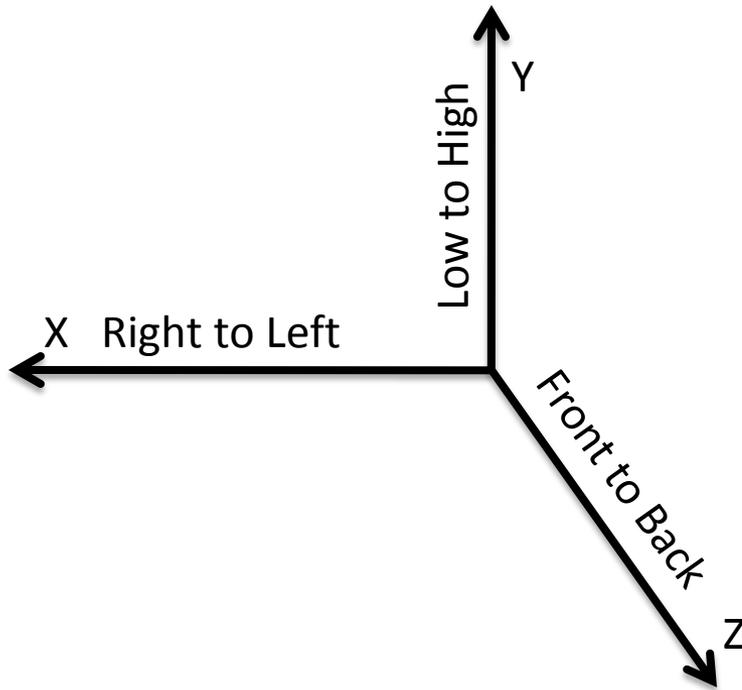
- Enable describing devices in 3-dimensional space
- Enable receivers to render correctly in terms of spatial placement
- Enable implementers to give a sense of real-world dimensions
- Enable multiple viewpoint use cases
- Enable complex devices to be described accurately but don't burden simple devices

# Proposal

- Coordinate system is Cartesian X, Y, Z
  - Origin (0, 0, 0) located at a spot of the implementer's choosing
- Coordinates can either be “virtual” or “real” units (mms).
- Each capture set has its own coordinate system, independent of those for other capture sets

# Directionality

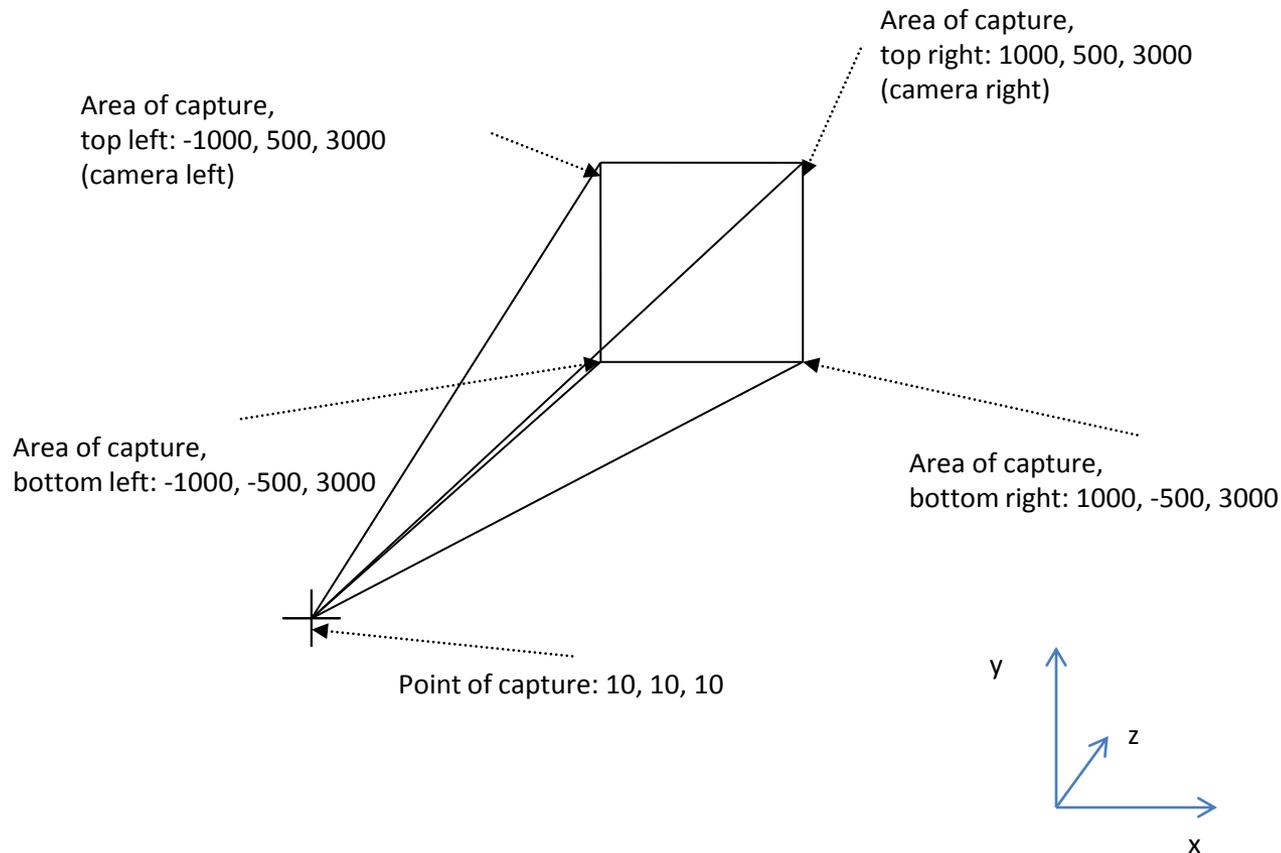
As numbers increase for the X axis, movement is from audience right to audience left  
As numbers increase for the Y axis, movement is from audience low to audience high  
As numbers increase for the Z axis, movement is from audience front to audience back



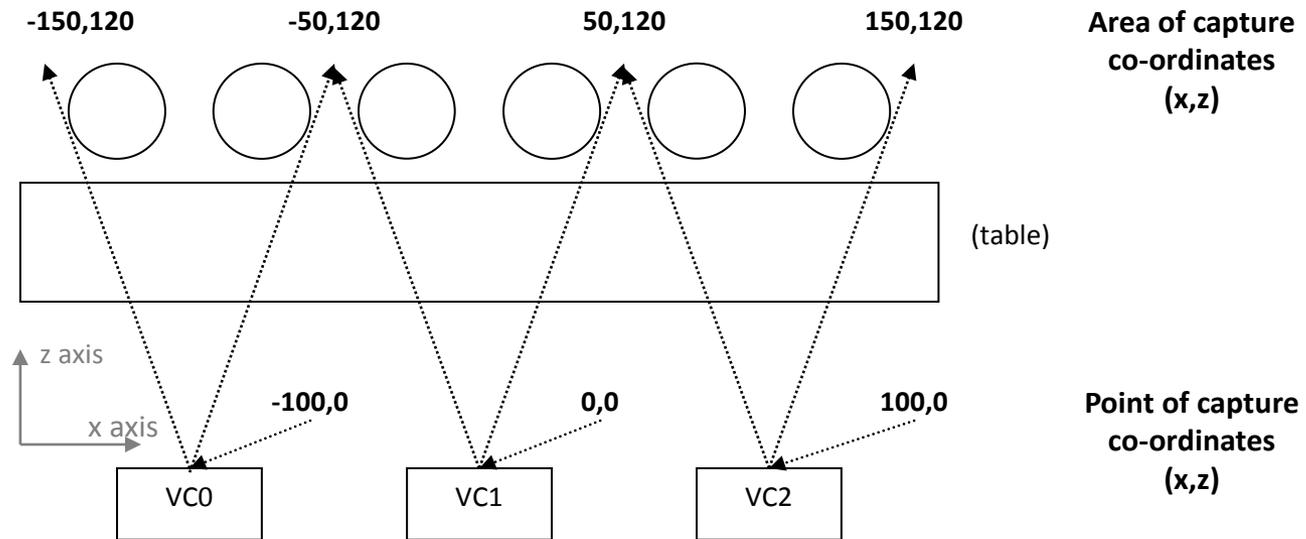
# Relevant Terms

- For each media capture there is a 'point of capture' and 'area of capture' defined
  - Point of capture is the location from which the media is captured and is defined by a single point
  - Area of capture is the extent captured by the media capture and is defined by 4 points

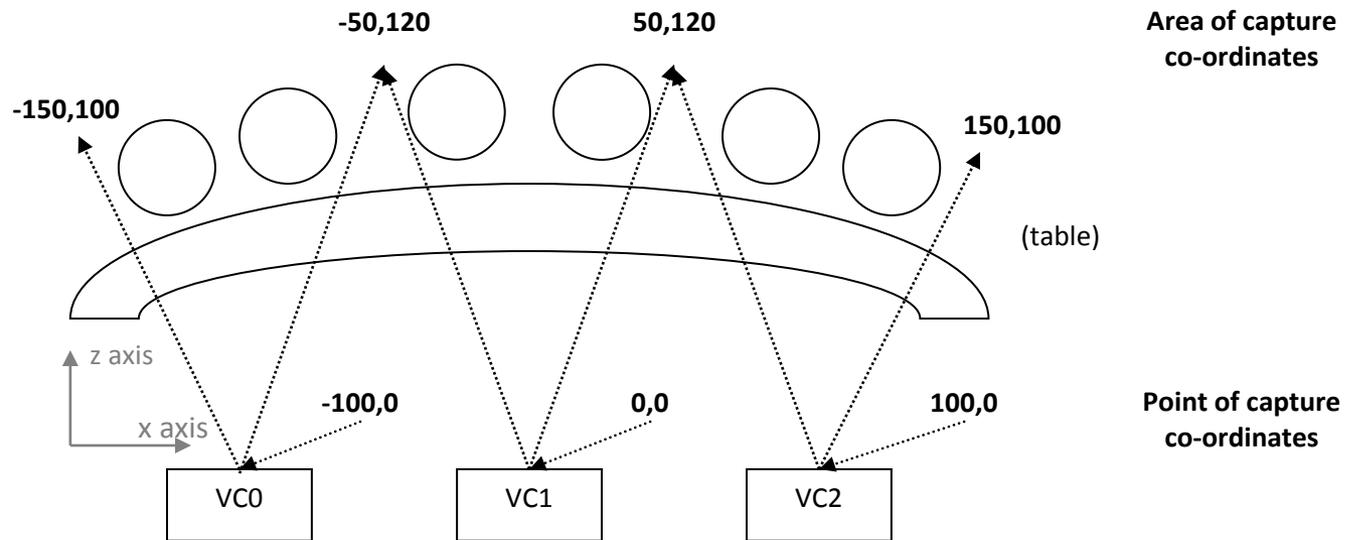
# Point of Capture/Area of Capture



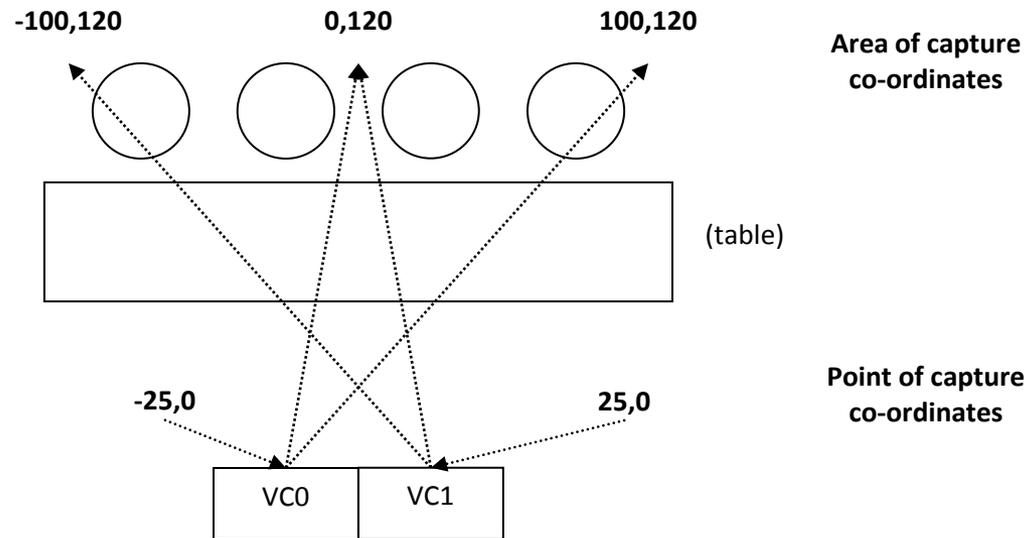
# 3 Camera Example



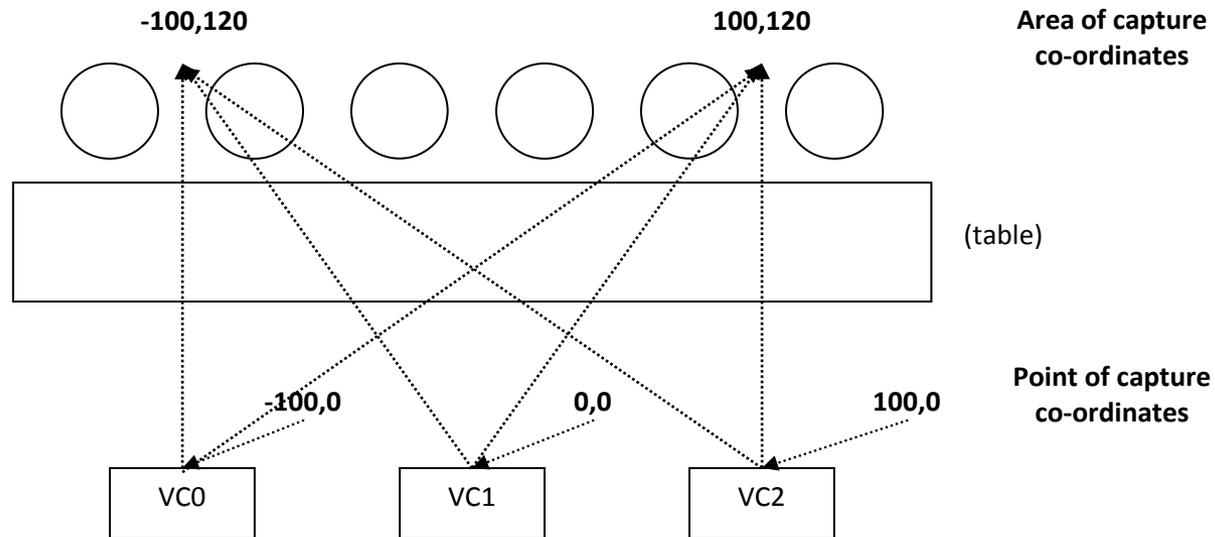
# Curved Table Example



# 2 Camera 'Crossed' Example



# Multiple Viewpoint Example



# Final thoughts

- Allows for modeling things like gaps between captures
- Providing coordinates for a capture set is optional
  - Devices that know their physical dimensions should provide them
  - Simple devices need not be burdened

# Issue 2 - Source Selection

- Do we need the ability for a consumer to select a particular source from multipoint conference? - yes
- Do we need the ability to advertise every media capture in the whole conference to every consumer?
- Many streams from middlebox to endpoint - does the middlebox capture set advertisement essentially become the union of all other endpoints capture sets?
  - Filtered to remove mutually exclusive choices?
- Should we relate this to draft-westerlund-dispatch-stream-selection?

# Issue 3 - Describing composed (including switched) captures

- What needs to be described about composed captures?
- Should the consumer be told which sources are included?
- Should the consumer be told where (x, y) a particular source is, in a composed image?
- Should the consumer be told all the capture set and attribute information about the original sources?
  - Recursively?
- Options:
  - Proposal - Just use ssrc/csrc - somehow map to separate roster list with additional information
  - Define additional provider -> consumer message with information about what is in the streams at any given time

# Issue 4 - selecting composed (including switched) captures

- Should a consumer be able to select/request a certain composition algorithm from the provider?
- Selecting sources to be mixed into a composed capture?
- Options
  - Do nothing
  - Define a few specific policies provider can advertise, and consumer can choose?
    - Separate from media capture attributes?