

MPEG for the past, present and future of television

MPEG用于电视的过去、当前和未来

Leonardo Chiariglione

Keynote Speech at ATSC Annual Meeting
Washington, DC – 2012/05/08

2012-05-08, MPEG召集人在ATSC年会的主题演说
徐孟侠翻译, 黄铁军校对; 2012-09-18



MPEG for the past of television

MPEG 用于电视的过去

- System 系统
 - Transport Stream 传送流
 - DSM-CC 数字存储媒体—命令与控制
 - Timed metadata 有定时的宏数据
- Video 视频
 - MPEG-2 High Profile **MPEG-2**高档次
- Audio 音频
 - MPEG-1 Audio Layer II **MPEG-1**音频第2层
 - MPEG-2 AAC **MPEG-2**高级音频编码

An assessment 评价

- MPEG has provided the means for the television business on air, cable and satellite to migrate from the analogue to the digital age **MPEG**对地面、有线和卫星的电视服务从模拟过渡到数字时代，提供各种方法
- It has been a collective efforts 它是集体努力的结果
 - Representatives from the entire spectrum of the broadcasting world provided their requirements 来自广播界的方方面面代表已提出各种需求
 - Packages of patent licences were created (outside of MPEG) 在**MPEG**之外，已创建成套的专利许可证
- Today there is virtually no broadcasting system that is not based, at least partially, on MPEG standards 今日的现实中，没有一个广播系统不是基于**MPEG**标准的(至少是部分基于)

MPEG for the present of television

MPEG 用于电视的当前

- Systems 系统
 - Carriage of MPEG-4 on MPEG-2 TS
在 **MPEG-2 TS**上承载**MPEG-4**
 - Dynamic Adaptive Streaming over HTTP (DASH)
HTTP(DASH)上的动态自适应媒体流
- Video 视频
 - MPEG-4 Advanced Video Coding (AVC) **MPEG-4**高级视频编码
 - 3D extensions 各种三维扩展
- Audio 音频
 - Various forms of Advanced Audio Coding (AAC)
高级音频编码的各种形式
 - Universal Speech and Audio (USAC)
通用的语音和音频编码

An assessment 评价

- The deployment of 2nd generation digital broadcasting systems continues successfully
第二代数字广播系统的部署继续顺利进行
- Use of the AVC on MPEG-2 TS allows packing twice as many programs as MPEG-2 Video in an analogue channel
在单个模拟电视频道中，采用**MPEG-2 TS**承载**AVC**视频，替代**MPEG-2**视频，可使节目数量翻一番
- With IPTV the notion of broadcasting is expanding
采用**IPTV**后，广播的概念正在拓展
- MPEG is helping the broadcasting industry
MPEG正在帮助广播工业界
 - To enrich the broadcasted user experience
丰富了广播用户的体验
 - Through standard technologies that serve multiple fields
通过各种标准技术服务于多类领域

MPEG for the future of television

MPEG 用于电视之未来

- Well, that's what I am supposed to talk about...
好，下面就是我今天准备讨论的...

Broadcasting is a great business to be in...

广播业在未来...仍然将是一个巨大的事业

- Broadcasting is a proven business with a constant revenue flow

广播业已被证实具有稳定的收益流之事业

- Many technologies appear that 很多技术表现出能够：
 - Can help new competing businesses but 有助于具有竞争力的新事业
 - Can also be used to extend the broadcasting business 而且还可用于扩展广播事业
- MPEG is a neutral body devoted to the development of standard technologies for user communities to exploit
MPEG是一个中立团体，致力于开发各类用户群体所需的各种标准化技术

MPEG-H: a video standard for future broadcasting

MPEG-H:用于未来广播的视频标准

- ISO/IEC 23008 High Efficiency Coding and Media Delivery in Heterogeneous Environments

ISO/IEC 23008: 在各种异构环境中的高效编码和媒体发送

- A suite of standards for media coding and delivery
用于媒体编码和发送的一整套标准
- Currently scheduled to appear in 3 parts

目前计划的时间表包括3部分:

- MPEG Media Transport **MPEG**媒体传送
- High Efficiency Video Coding 高效视频编码
- 3D Audio 三维音频

- Timeline

时间表

CD – 委员会草案

DIS – 国际标准草案

FDIS – 最终的国际标准草案

Part	CD	DIS	FDIS
MMT	12/07	13/01	13/07
HEVC	12/02	12/07	13/01
3D Audio	14/04	14/10	15/01

Service composition in MMT

MMT (MPEG 多媒体传送流) 中的服务组装



<subtitle>
<标题>
Bird is flying...
鸟在飞...



legacy broadcasting 传统广播

MPEG for the past, present and future of TV

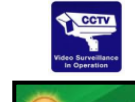
MMT Package MMT 信息包



Composition information 组装信息
friendly with html5
与 HTML5 友好

通用发送的各种内容





UD contents

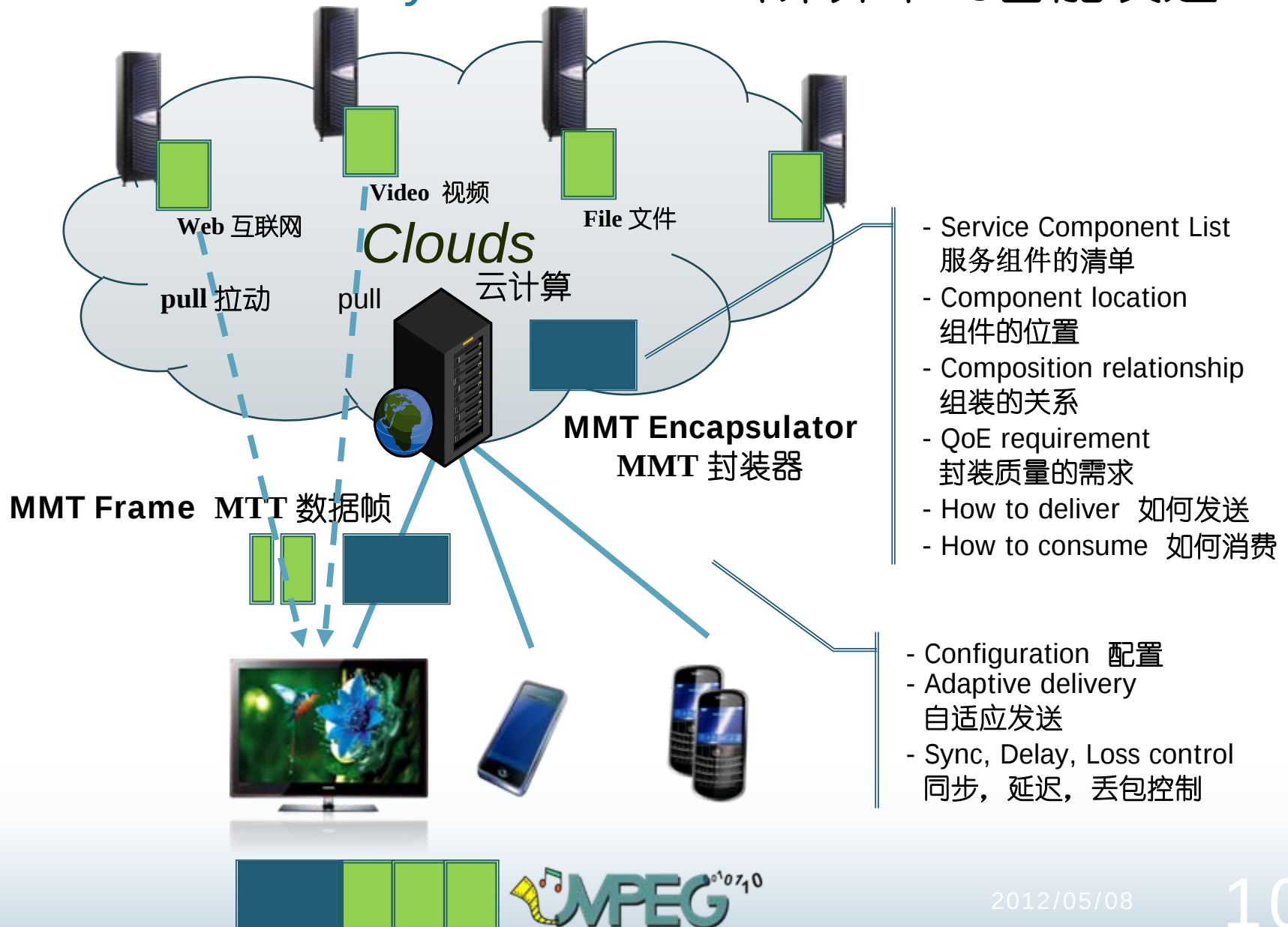
3D Earth 3D 地球

future broadcasting 未来广播

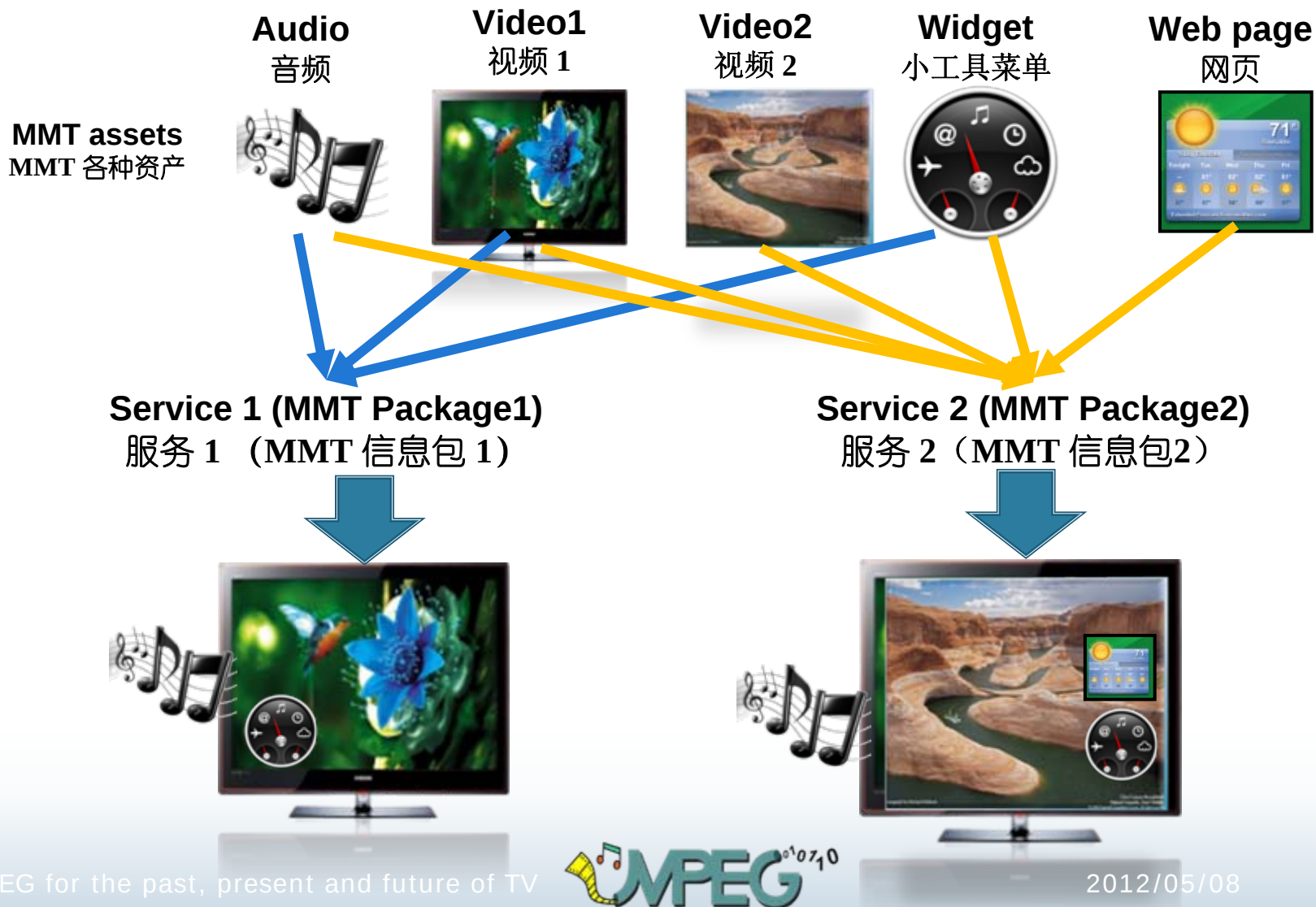


2012/05/02

Smart delivery in Clouds 云计算中的智能发送



Mash-up service 混成服务



Second screen management 第二屏幕的管理

MMT Package

MMT 信息包

Composition
information 组装信息

<device 1>...

<设备 1>...

<device2>...

<设备 2>...



Video content on 2nd device
在第二设备上的视频内容



Additional Information with Live Content (HTML)
具有实况内容的附加信息 (HTML)

MPEG-H part 2 High Efficiency Video Coding

MPEG-H 第 2 部分 高效视频编码

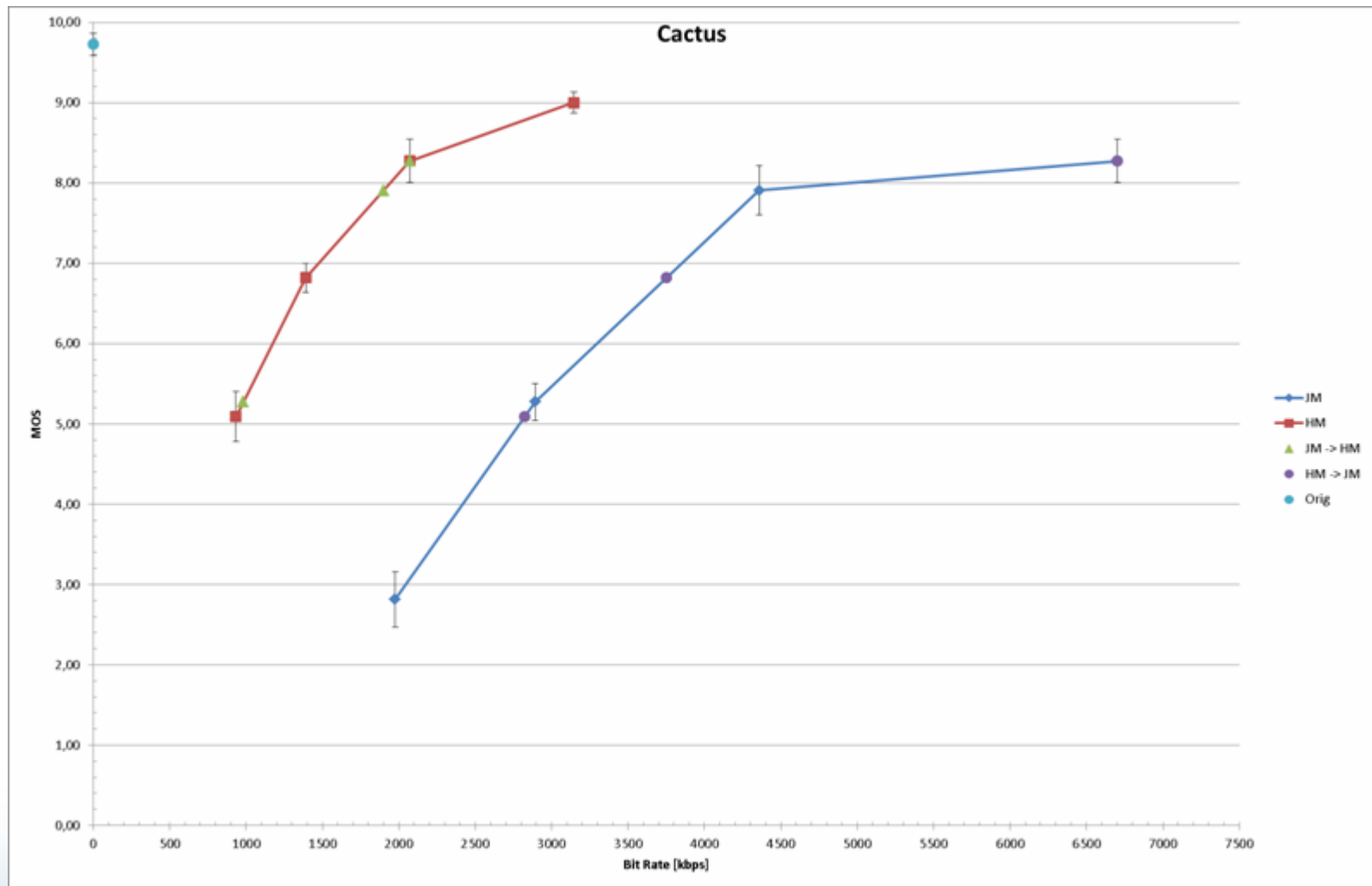
- Primary goal: $\geq 50\%$ better compression than AVC
最初目标：压缩效率比AVC提高 $\geq 50\%$
 - For large displays(e.g. 4kx2k) 用于大屏幕 (例如 4kx2k)
 - For mobile (e.g. WVGA) 用于移动服务 (例如, WVGA)
- Developed jointly with ITU-T by Joint Collaborative Team (JCT) on Video Coding
由MPEG与ITU-T共同成立的“视频编码联合协作组”来制定
- Technically in the same stream of past MPEG Video Coding standards (AVC and prior standards)
技术上使用与MPEG以前的视频编码标准(AVC 及以前各种标准)相同的传送流
- Currently at Committee Draft level
目前处于委员会草案(CD)的阶段

HEVC vs AVC Subjective Performance Measurements

HEVC 和 AVC 主观性能测试的对比

- *Preliminary* subjective tests carried out for HEVC Committee Draft
对HEVC的委员会草案(CD), 已进行初步的主观测试
- Reasonably small and non-overlapping confidence intervals
合理较小的和不重叠的信任度间隔
- Average rate savings at same (interpolated) MOS points
在相同的(内插的)**MOS**(评分平均)点, 得出可节省的平均比特率
 - HEVC with 67% lower rate in class B (Full HD)
在B类(全高清)中HEVC的比特率要低67%
 - HEVC with 49% lower rate in class C (WVGA)
在C类(WVGA)中HEVC的比特率要低49%
 - HEVC with 58% lower rate overall
HEVC整体比特率要低58%

A performance sample 某个性能样本



For those technology aware...

涉及的有关技术...

- Block-based/Variable block sizes 基于块的/可变块尺寸
- Block motion compensation 块的运动补偿
- Fractional-pel motion vectors 分像素的运动矢量
- Spatial intra prediction 空间域的帧内预测
- Spatial transform of residual difference
残差值的空间域变换
- Integer-based transform designs
基于整数的各种变换设计
- Arithmetic or VLC-based entropy coding
基于算术的或变字长编码的熵编码
- In-loop filtering to form final decoded picture
环内的滤波处理，以形成最终的解码图像

Where we are (May 2012)

我们现处何处 (2012 年 5 月)

- Activity: level of 1,000 documents per meeting crossed
活动：每次会议审议约**1,000**个量级的文件
- Very diverse participation from industry/academia
工业界和学术界的极广泛参加
- Significant technical advance over prior standards
与过去的各种标准相比，已有显著的技术进步
- Challenge: computational/implementation complexity
挑战：计算/实现的复杂度
- Deliverables: 可发布的：
 - Spec 技术规范
 - Reference software 参考软件
 - Conformance 符合性
- Currently only one profile 目前只有一个档次

HEVC Scalability Extensions/1

HEVC 可伸缩性扩展/1

- Serve the needs of heterogeneous environment of future digital TV distribution

服务于未来数字电视分发的异构环境之各种需求

- Multiple devices: SHDTVs, set-top boxes, tablets, smart phones, PCs etc.
多类设备：超高清电视机，机顶盒，平板电脑，智能手机，个人电脑等
 - Wide range of processing powers, display sizes and power consumption needs 较宽范围的处理能力、显示器尺寸和功耗需求
- Multiple networks: Cable, Satellite, Terrestrial, CDN, Cable Modem, xDSL, WiFi, 4G
多类网络：有线、卫星、地面、有线数字网络、有线调制解调器，xDSL，WiFi，4G
 - Wide range of channel bandwidths and QoS 较宽范围的各种信道带宽和服务质量
- Multiple protocols (多类协议): MPEG-2 TS, DASH, HLS, RTP, UDP
- Multiple services: Broadcasting, On demand, Streaming, Over The Top 多类服务：广播，点播，媒体流、三网融合

HEVC Scalability Extensions/2

HEVC 可伸缩性的扩展/2

- Options for backward compatible migration paths
用于向后兼容的过渡途径之各种可选项
 - From 720p60 to 1080p60 从720p60到1080p60
 - From 1080p60 to 4kx2kp60 从1080p60到4kx2kp60
 - From AVC to HEVC 从AVC到HEVC
- Time Line 时间表
 - May 2012 Preliminary Call for Proposals (public)
2012年5月, 初次征求提案(公开)
 - July 2012 Final Call for Proposals (public)
2012年7月, 最终征求提案(公开)
 - October 2012 2012年10月
 - Evaluation of proposals & start of collaborative design phase
评估各种提案, 并开始协作设计阶段
 - Development of the time line associated with standardisation phases (WD, CD/PDAM, FDIS/FDAM, DIS/AMD)
按照标准化程序, 设定执行时间表

Service Transitions 服务的几个过渡

- 1950s: Color TV 1950年代：彩色电视
 - Analog, backward compatible 模拟的，向后兼容
- 1990s: Digital TV 1990年代：数字电视
 - New infrastructure required 需要新的基础设施
 - Transitions from SD to HD 从标清到高清的各种过渡
- 2010s: 3D 2010年代：三维电视
 - Mixed results from services introduction is 来自各种服务导入的各种混合结果
 - Not a single format across all services 已不再是涵盖各种服务的单一格式



Current Status of 3D Video

三维视频的现状

- Increasing investments in 3D production and services
对**3D**制作和服务的投资日益增长
- Increasingly capable 3D displays in the market (many competing and emerging technologies) 市场上的**3D**显示器日益增长(大量竞争的和融合的技术)
- Market is primarily stereo (mix of different formats being deployed right now) 市场上主要是双目立体的(不同格式的混合, 目前正在部署中)
- Focus of current MPEG work: development of a new 3D format and associated compression techniques that could facilitate generation of multiview output
当前**MPEG**工作的重点: 设定一个新的**3D**格式及相应的压缩技术, 以支持多视输出的生成

Existing 3DV Coding Formats

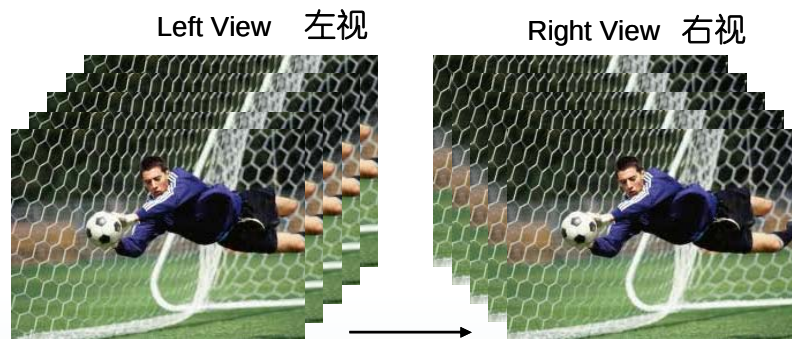
现有三维视频编码的各种格式



Left 左 Right 右

- Pack pixels from left and right views into a single frame (loss of resolution)
把来自左视和右视的各像素，拼接进单个图像帧（分辨率下降）
- Leverage existing infrastructure and equipment, only require additional signaling to de-interleave frame 权衡现有的基础设施，只需附加“去交织”图像帧的信令
- Embraced by broadcasters for first phase of 3D broadcast services
在 3D 电视广播的第一阶段受到广播业者欢迎

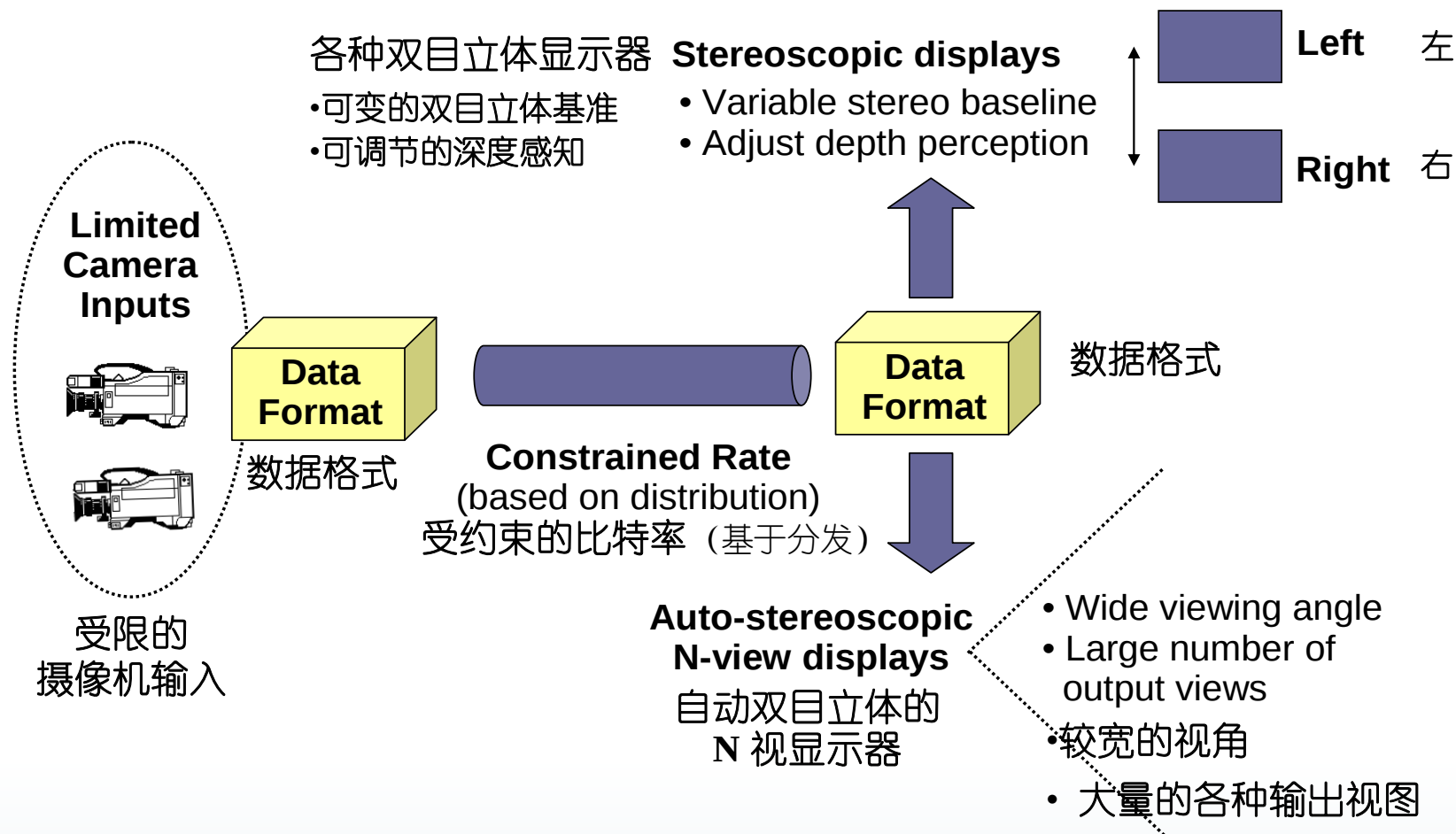
- Full-resolution coding of stereo and MV video as extension of AVC 双目立体和多视的视频之全分辨率编码，是AVC的扩展
- High coding efficiency achieved via inter-view prediction techniques 通过视间预测技术获得更高的编码效率
- 2D compatibility supported 能够兼容二维电视
- Adopted as format for 3D Blu-ray Disc
已采纳为 3D 蓝光盘的格式
- Being considered for second phase of broadcast standards 正在考虑用于各种广播标准的第 2 阶段



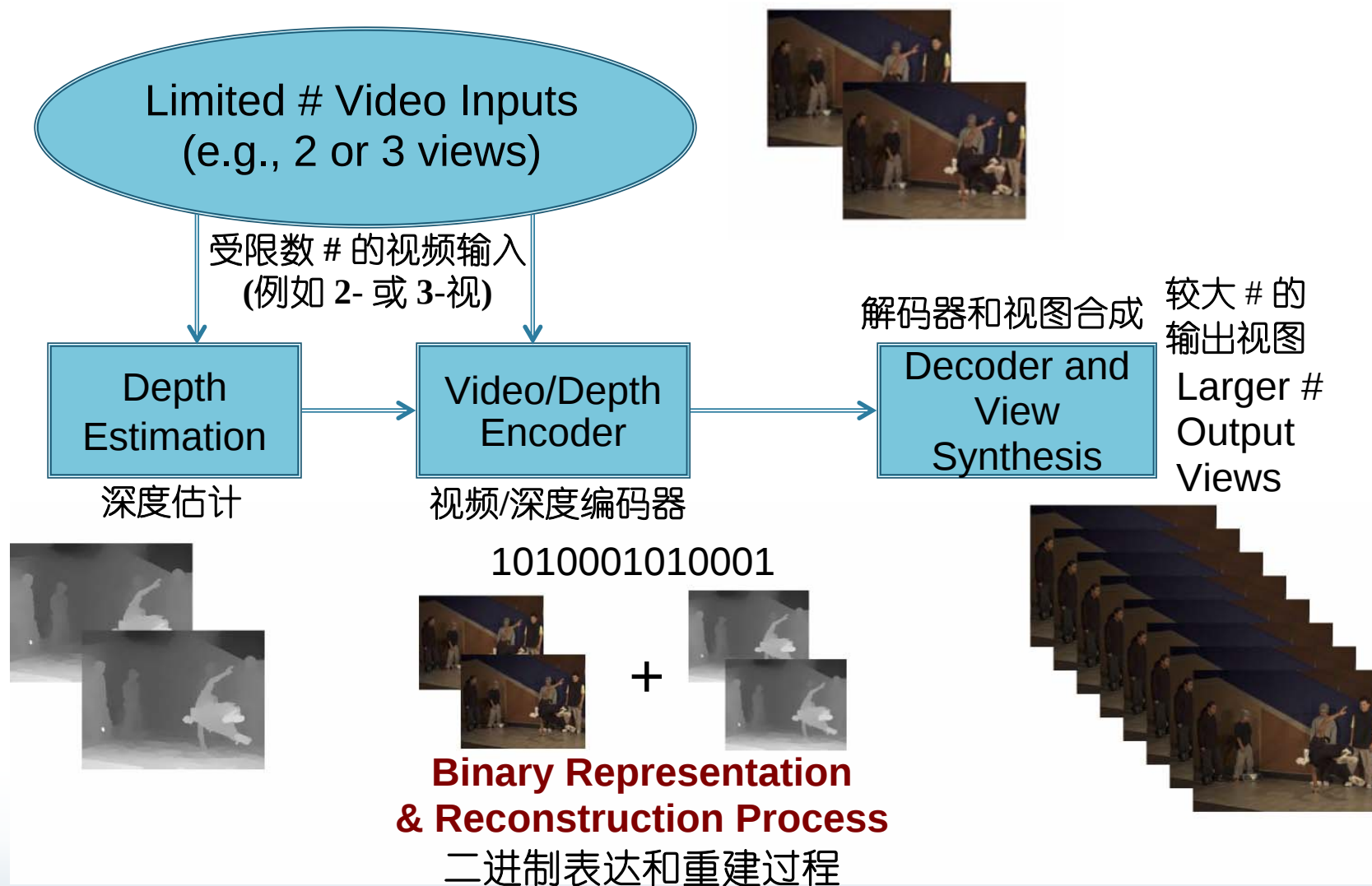
Effective prediction between views for high compression performance
各视之间的高效预测用于高压缩性能

Targets of Future 3DV Format

未来三维视频格式的各种目标



3DV Reference Framework 三维视频参考框架



Overview of Call For Proposals

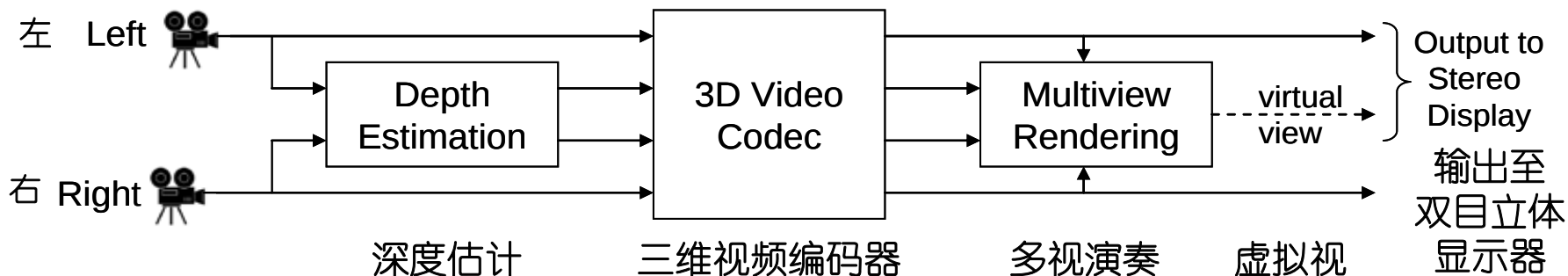
提案征集的概述

- Test Material & Conditions 各种测试素材及条件
 - 8 test sequences (1920x1080 & 1024x768) 8个测试序列
 - 4 target bit rates per sequence 每个序列有4个目标比特率
 - 2-view and 3-view test scenarios 2-视和 3-视的各种测试场景
 - AVC and HEVC compatible test categories
与AVC和HEVC兼容的各种测试类别
- 23 responses (12 AVC + 11 HEVC) 23个响应 [提案]
- Evaluation in Nov 2011 在2011年11月的评估
 - Objective quality (PSNR of synthesized views)
客观质量(各种合成视的PSNR)
 - Subjective assessment (stereo and auto-stereo)
主观评分(双目立体和自动立体)

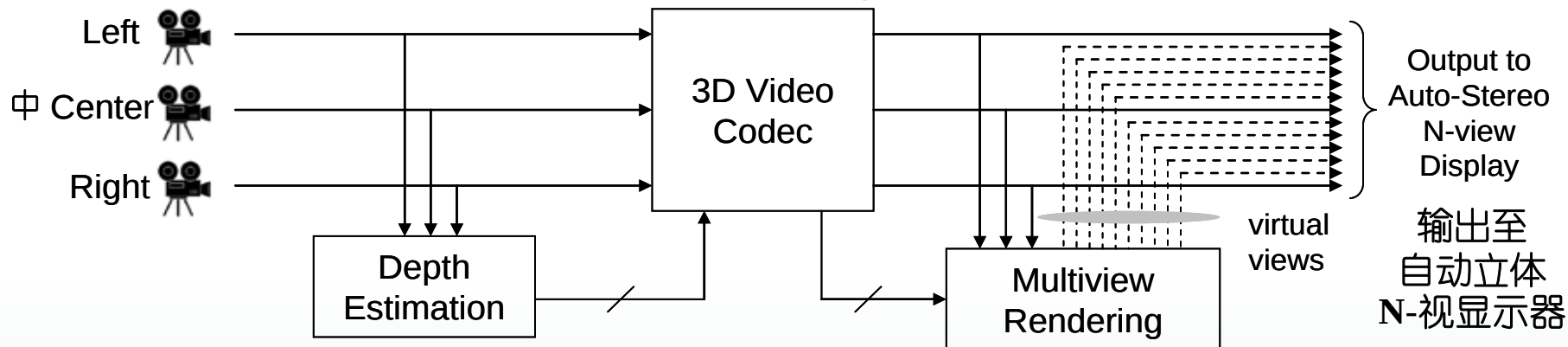
3DV Testing Scenarios

三维视频的各种测试场景

2-view Configuration 2-视配置



3-view Configuration 3-视配置



Tool Categories 工具的分类

- Texture coding 纹理编码
 - Independent of depth 不依赖于深度
 - E.g., inter-view prediction of color view, inter-view prediction of motion parameters and residual data
例如，彩色视的视间预测；运动参数和残差数据之间的视间预测
 - Using depth data 利用深度数据
 - E.g., view synthesis prediction, motion prediction
例如，视图合成的预测，运动预测
- Depth coding 深度编码
 - Independent of texture 不依赖于纹理
 - E.g., depth modeling modes, weighted prediction, reduced resolution coding
例如，深度建模的各种模式，加权预测，降低分辨率编码
 - Using texture data 利用纹理数据
 - E.g., motion parameter inheritance, intra prediction
例如，运动参数的继承，帧内预测

3DV Current Status 三维视频的现状

- Extensive subjective evaluation of 20+ proposals to the 3DV CfP

对3DV征集提案的20多个提案之广泛主观评估

- Considered both AVC and HEVC compatible

考虑同AVC和HEVC的兼容性

- Many new coding tools proposed for MV texture and depth coding

已提出多视纹理和深度的多种新编码工具

- Substantial rate savings compared to capabilities of existing standards

与现存标准的各种可能性相比, 可显著节省比特率

- AVC and HEVC extensions standardisation under way

AVC和HEVC的扩展之标准化正在进行

- Add support for depth

增加对深度的支持

- New AVC-based coding tools

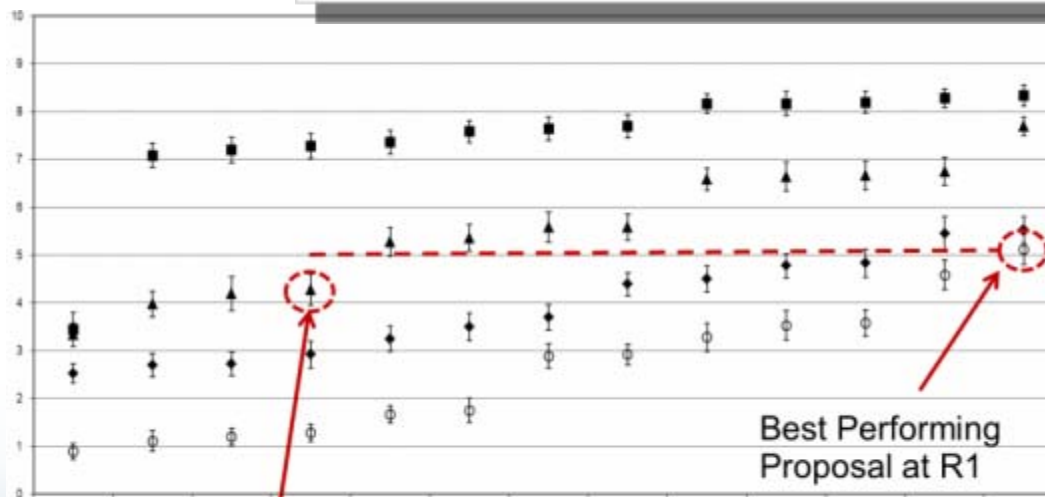
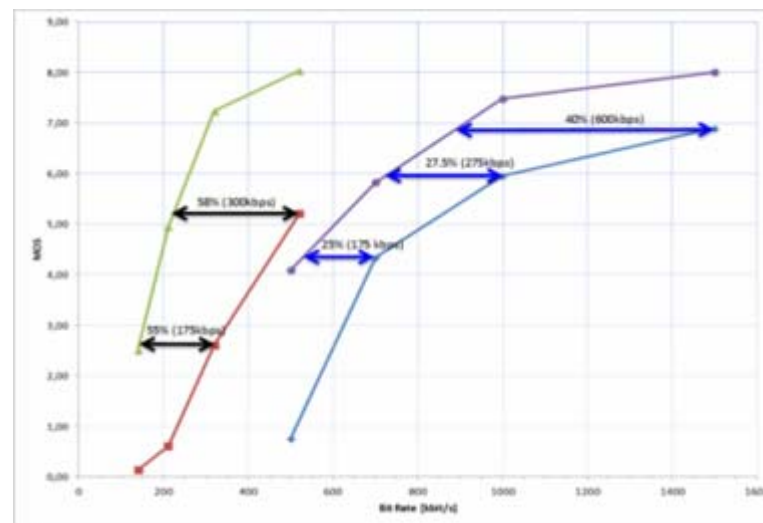
各种新的基于AVC的编码工具

- MV extensions of HEVC

HEVC对多视的各种扩展

- Hybrid architectures

各种混合结构



Anchor at R3 R3 的特例

R1 建议的最佳性能

Hybrid Architectures 各种混合结构

- From a pure compression efficiency point of view, it is always best to use the most advanced codec
单纯从压缩效率的观点来看，采用最先进的编码器总是最佳选择
- However, when introducing new services, providers must also consider capabilities of existing receivers and transition plan 然而，在引进各种新服务时，服务提供者还必须考虑现存接收机和过渡计划的各种可能性
- Use Case: 示例
 - Many terrestrial broadcast systems based on MPEG-2
很多地面广播系统基于**MPEG-2**
 - US cable systems based on mix of MPEG-2 and AVC
美国有线系统基于**MPEG-2**和**AVC**的混合
 - May not be easy to simply switch codecs in near term
在近期简单切换编码器可能不容易

Multiview Extensions of HEVC

HEVC 的多视扩展

- HEVC Stereo/MV extensions being considered
正在考虑HEVC双目立体/多视的各种扩展
 - Extensions expected to be completed approximately 12 months after the base specification, i.e., Jan 2014
这些扩展预期在基本规范确定后约12个月完成，即2014年1月
 - Reportedly simple extensions of HEVC can achieve 30-40% bit rate reduction compared to HEVC simulcast
已有报告：对HEVC的各种简单扩展与HEVC同播相比，可使比特率下降30-40%
 - Further gains expected if block-level changes to codec are considered
如果对编解码器考虑图像块级的各种变更，预期可得更多增益

MPEG-H part 3: 3D Audio

MPEG-H 第3部分：三维音频

- Well, there is no Systems and Video without Audio...
好，没有不带音频的各种系统和视频...

3D Audio requirements/1

对三维音频的需求/1

- **High quality:** quality of decoded sound perceptually transparent 高质量: 解码声的质量在感知上是透明的
- **Localization and Envelopment:** Accurate sound localization and very high sense of sound envelopment within a targeted listening area (“sweet spot”) 定位和组装: 在预定的收听区(“欣赏点”)内, 有准确的声学定位和极好体验的声学组装
- **Flexible Loudspeaker Placement:** the transmitted audio program to be mapped to a rendering setup in which loudspeakers are in alternate locations 灵活的扬声器摆放: 传输的音频节目将映射到一套演奏装置, 其中的扬声器则摆放在各种不同位置
- **Interactivity:** Interactive modification of the sound scene rendered from the coded representation 交互性: 从编码表示中演奏的声音场景, 可进行交互式修改

3D Audio requirements/2

对三维音频的需求/2

- **Rendering on setups with fewer loudspeakers (including headphones):** ability to derive a signal from the transmitted program material for reproduction with fewer loudspeakers

采用扬声器较少的配置(包括耳机)也可演奏: 扬声器较少时, 仍然能够从传输的节目素材中导出一个信号, 来满足重现的需求

- **Audio/visual alignment & consistency:** ability to adapt the rendered acoustic scene to be consistent with the visual

音频/视频的对准和一致性: 能够调节演奏的声学场景, 以便与视觉场景保持一致

- **Efficiency for decoding on different setups:** ability to be rendered on all reproduction loudspeaker setups or headphone

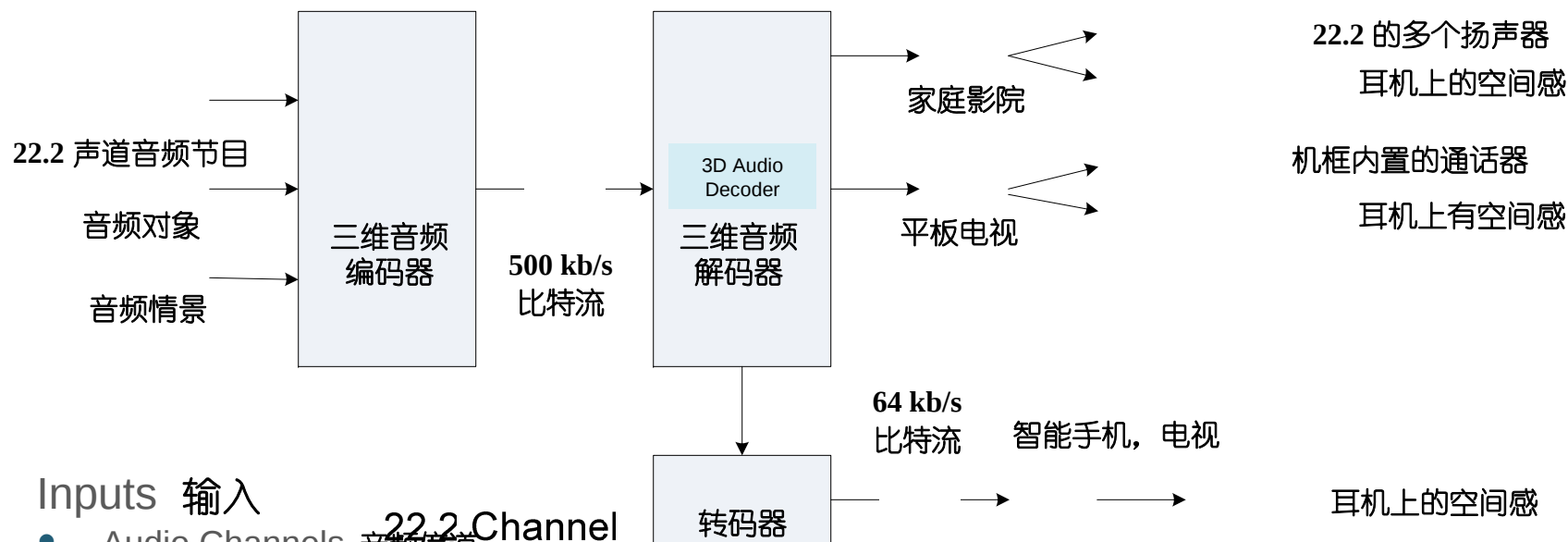
在不同装置上进行解码的有效性: 能够在所有重现的扬声器装置或耳机上进行演奏

3D Audio requirements/3

对三维音频的需求/3

- **Appropriate computational complexity and system latency** for the target application (Broadcasting, Spatial two-(n-)way communication / telepresence)
适当的计算复杂度和系统时延，以满足目标应用(如广播、空间的 2-(n-) 路通信/遥现)
- **Transcoding for low bandwidth devices:** A lower bandwidth signal can be extracted from the original program material
对窄带宽的设备之转码：可从源节目素材中提取窄带宽信号
- **Backward compatibility:** (e.g. to 5.1 channel programs and decoders or transmission of HRTF encoded signals)
向后兼容性：例如，兼容 5.1 声道节目和解码器，或 HRTF 编码的信号

Envisioned Architecture 设想的结构



Inputs 输入

- Audio Channels 音频信道
 - “normal” content “常规的” 内容
- Audio Objects 音频对象
 - “helicopter overflight” “直升机的飞过”
- Audio Scene 音频情景
 - Ambisonics 环境声效果

22.2 Channel Audio Program

Channel 信道

- 500 kb/s primary channel 500 kb/s 主信道
- 64 kb/s cellular channel 64 kb/s 蜂窝网信道

Audio Scene

Outputs 输出

- Loudspeakers 扬声器
- Headphones 耳机
 - Spatialized sound 有空间感的声音
- In-frame loudspeakers 机框内置的扬声器
 - “Sound Bar” “声吧”

Audio, part of an Audio-visual Experience

音频是音视频体验的一部分

- What is new in video? 视频中有哪些是新的？
 - Larger, high-resolution displays 更大的高分辨率显示器
 - Greater sense of envelopment 组装的更好感知
 - Possibly closer viewing distance 观看距离可能更靠近
 - More efficient compression 更高效的压缩
 - Wireless transmission to portable devices
无线传输至各种便携式设备
- Immersive and enveloping audio experience
具有浸润性的和组装的音频体验

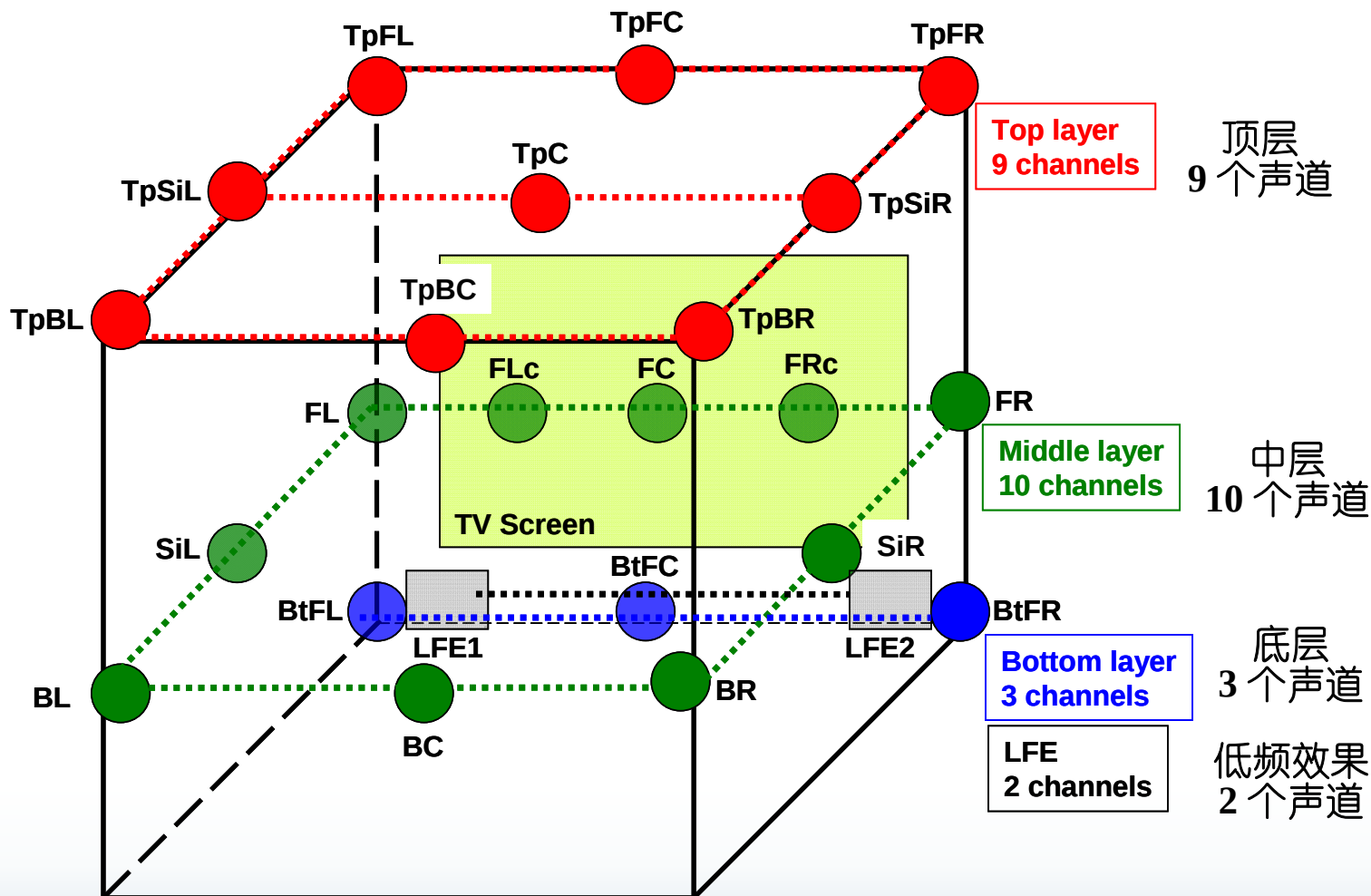
Use Case 1 Home Theatre

示例 1 家庭影院

- Many loudspeakers 多个扬声器
 - 10.1 [10 个扬声器加 1 个超低音扬声器]
 - 22.2
- “3D” Experience “三维” 体验
 - Height loudspeakers for greater envelopment
多个“高大音柱”用于更大的组装
 - Sense that audio objects are near or distant
音频对象为近或远距离时的感知
 - When near, they are consistent with video image
如果是近距离的，它们应该同视频图像保持一致性

22.2 multichannel sound system (NHK)

22.2 多声道声音系统 (日本广播协会)



Home Theater – Issues

家庭影院的有关问题

- Maintain compatibility with legacy systems: 5.1, stereo
维持同传统系统的兼容性：5.1, 立体声
- Will content providers really step up to 22.2?
内容提供者真的会升级到22.2 ?
- Will consumers adopt 22.2?: 2 front high
消费者会接受22.2 ? : 2个前置的“音柱”
- How to compress 22.2 for transmission?
如何压缩22.2, 以便传输 ?
 - MPEG AAC 64 kb/s/channel: -> 1.5 Mb/s
MPEG AAC 64kb/s/声道：增至1.5Mb/s
- How to render for legacy setups? 对传统装置如何演奏?
- Relatively slow-growth market 增长相对较慢的市场

Use Case 2 Personal TV

示例 2 个人电视机

- Small but high-resolution display
尺寸较小的、但高分辨率的显示器
 - “super-tablet” “超级平板[电脑]”
- Local wireless communications 本地的各种无线通信
 - To cable or fiber home hub 连接到有线或光纤的家庭集线器
- Possible audio presentation 可能的音频呈现方式
 - Headphones 耳机
 - Loudspeakers around perimeter of display
显示器周围的多个扬声器
- Issue: To what extent can the user get an enveloping experience from only “front” loudspeakers?
问题：用户仅仅从“前置的”多个扬声器，究竟可获得多大程度的组装体验？

Use Case 3 Mobile TV

示例 3 移动电视

- Hand-held display: Smartphone 手持显示器：智能手机
- Headphone listening: 耳机的收听
 - Stereo, perhaps with binauralisation
立体声；也许具有双耳效应化
- Fastest growing market: Quick rollout and adoption of standard 增长最快的市场：迅速推进和采纳标准
- What is new? 什么是新的？
 - USAC for stereo; MPEG Surround for binauralisation
USAC用于立体声；MPEG环绕声用于双耳效应化
- What is the “Wow” factor? 什么是“叫好”的因素？
 - To spur adoption of new technology 促进采纳新技术

User Experience 用户体验

Home Theatre
家庭影院



Spatialized on headphones
耳机上的空间感



Tablet TV
平板电视



Flexible Rendering 灵活的演奏

- Content providers create 22.2 program
内容提供者创建22.2节目
- Consumed on stereo, 5.1, 10.1, 22.2 layouts
在立体声, 5.1, 10.1, 22.2等配置上消费
- Consumed on “wrong” layouts 在“错误的”配置上消费
 - Mis-positioned surrounds 错误放置的环绕声
 - Missing surrounds 丢失环绕声
 - Non-standard layouts 非标准的配置
 - 2 front high, 5.1 mid 2个前置的“音柱”，而5.1在中间

3D Audio Test Platform 三维音频测试平台

- NHK “Loudspeaker Array Frame” (LAF)
日本广播协会的“扬声器阵列机架”(LAF)
- Suggestions for 建议作为：
 - Alternative platform for assessing “Home Theatre” use case
用于评估“家庭影院”示例的另一种平台
 - Model for assessing “Personal TV” use case
用于评估“个人电视”示例的模型

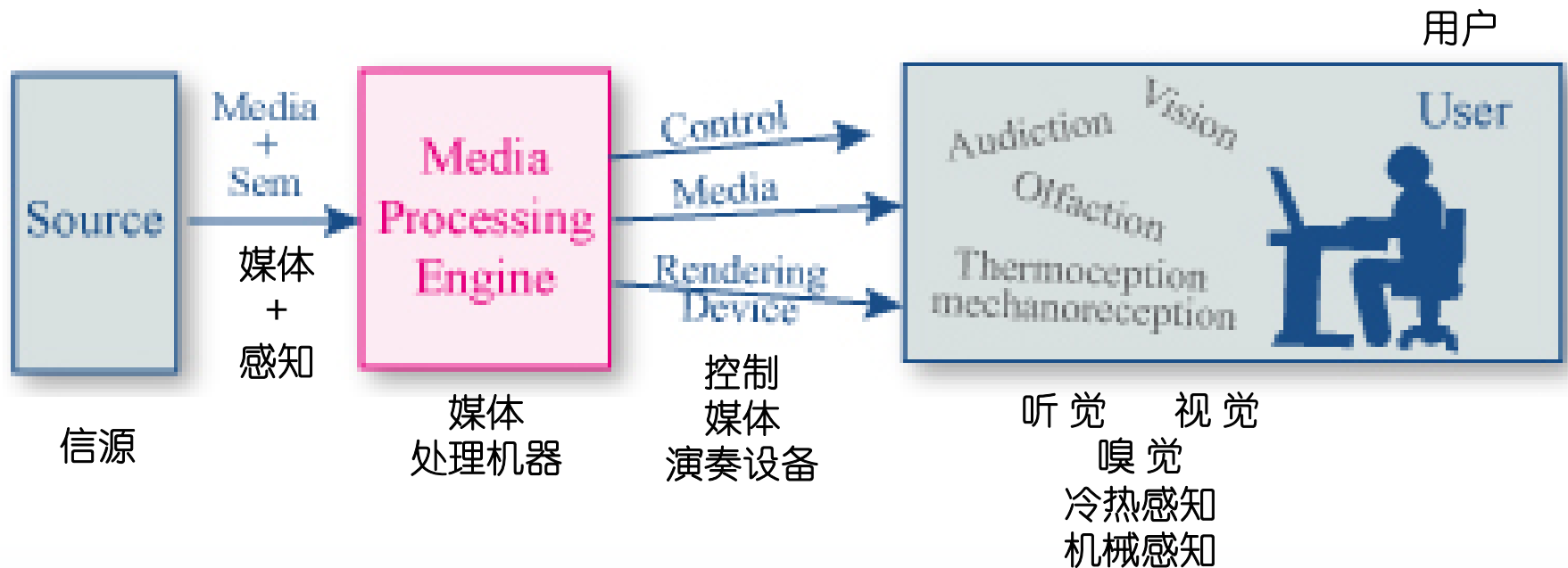


Summary 小结

- Provide compelling immersive audio experience for audio/visual programs
为各种音频/视频节目提供强烈浸润性的音频体验
- Create content once, present on many different loudspeaker layouts or on headphones
音频内容仅制作一次，就可以在多种不同的扬声器配置上或耳机上呈现
- Provide high compression and high-quality audio presentation
提供高压缩和高质量的音频呈现

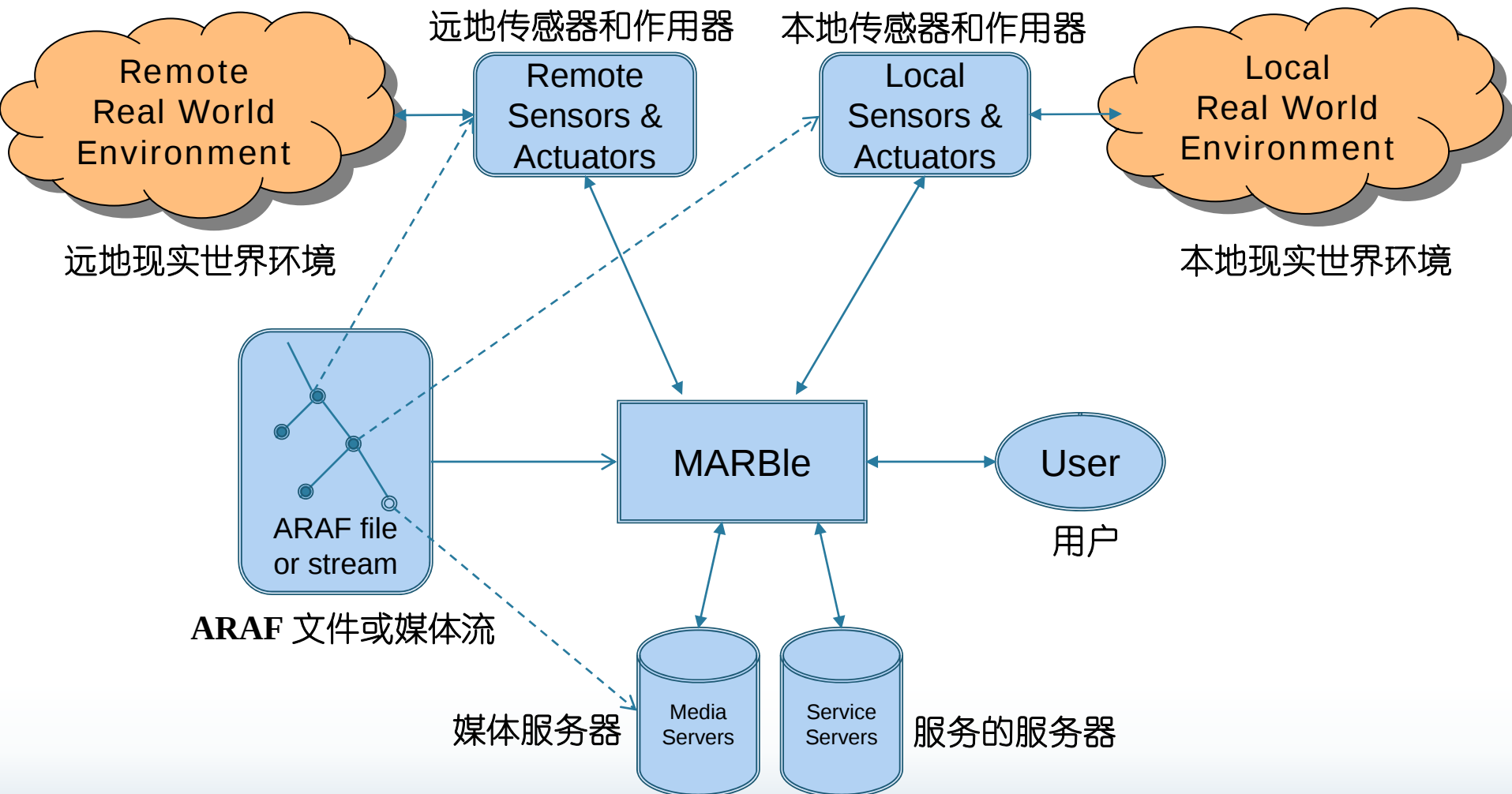
MPEG-V: Humans have more than two senses

MPEG-V: 人类有2种以上的感知



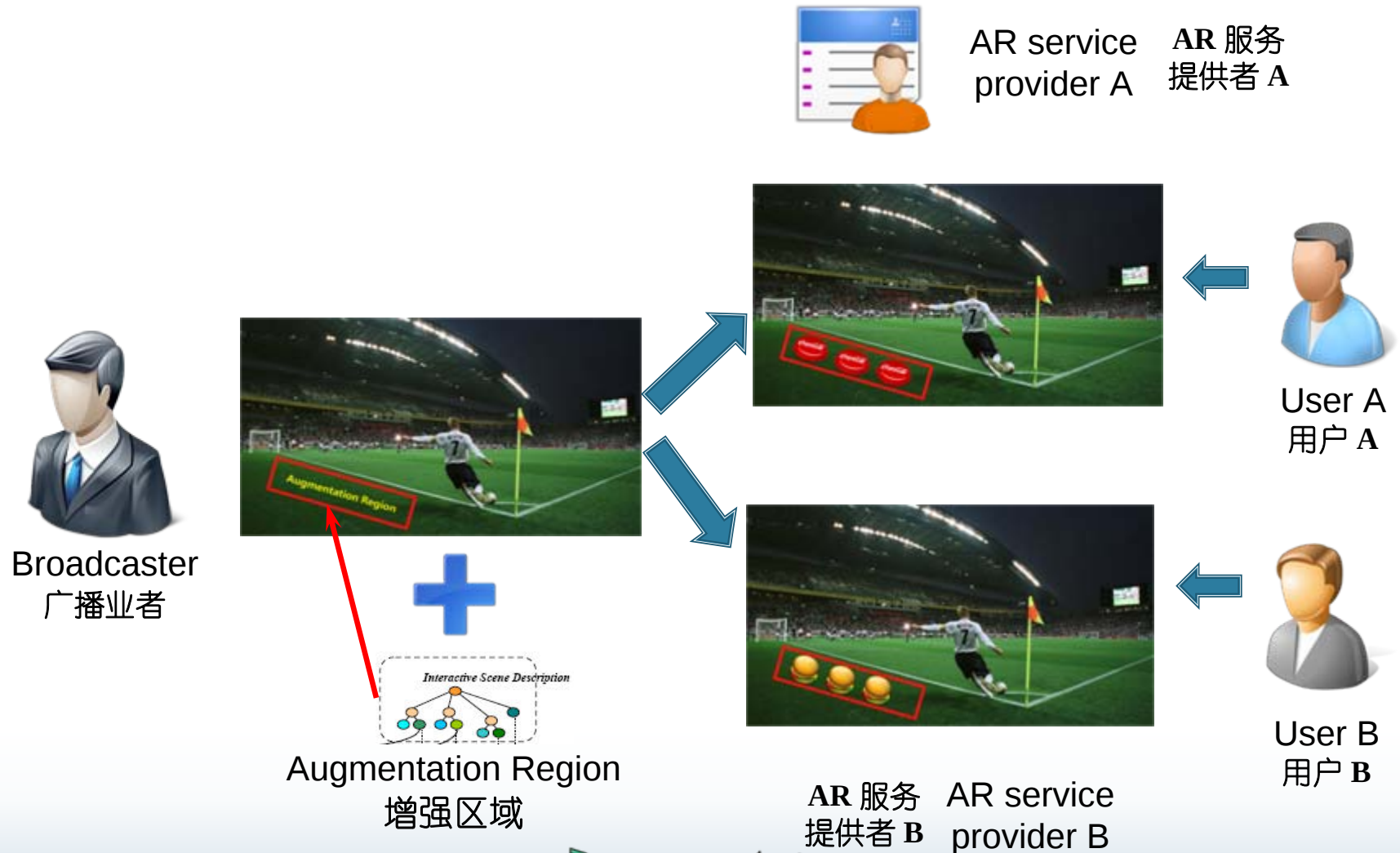
Augmented Reality Application Format (ARAF)

增强现实应用格式 (ARAF)



Augmentation Region in ARAF

ARAF 中的增强区域



The past, the present and the future, again

MPEG的过去、当前和未来；再次

- 25 years ago MPEG selected the digital media way bringing innovation and interoperability ...and thrived
25年以来，MPEG选择了带来创新和互操作性的数字媒体之路 ...因而MPEG已茁壮成长
- This year the broadcasting industry has started a move in future broadcast services with interoperability at its core ...and it is bound to thrive
今年，广播工业界对未来广播服务已启动一个活动 [指FoBTV]，以互操作性为其核心 ...因而MPEG必将继续茁壮成长
- MPEG looks forward to continuing a cooperation with the broadcasting industry providing the necessary standard technologies

MPEG今后寻求继续同广播工业界的合作，提供所需的标准技术



<http://mpeg.chiariglione.org/>