



مسئله‌های تنظیم مقررات
رسانه‌های صوت و تصویر فراگیر

ساترا

اداره کل فناوری و سامانه‌های نوین

رهنمودهای بلندای صدا برای محتوای رسانه‌ها
بر اساس سند فنی

AESTD1006.1.17-10



پایش و همسان سازی بلندای صدا

Based on
Loudness Guidelines for OTT and OVD Content

(۱/۳/۰)

رهنمودهای بلندای صدا برای محتوای رسانه‌ها

پایش و همسان سازی بلندای صدا

Loudness Guidelines for OTT and OVD Content

بر اساس سند فنی

AESTD1006.1.17-10

(۱/۳/۰)

سازمان تنظیم مقررات صوت و تصویر فراگیر
اداره کل فناوری و سامانه‌های نوین

پاییز ۱۴۰۰

رهنمودهای بلندای صدا برای محتوای رسانه‌ها

پیشینه سند:

گروه فنی^۱ AGOTTVS در اوایل سال ۲۰۱۶ تشکیل شد تا بسیاری از مسائل مربوط به تغییرات صوتی آنلاین بلندای صدا را بررسی کند. هدف آن تهیه توصیه‌نامه‌های جامع و ارائه رهنمودهای مؤثر برای مدیریت بلندای صدا در موسیقی متن تلویزیون و محتوای ویدیویی است که توسط OTT^۲ و توزیع‌کنندگان آنلاین ویدیو OVD^۳ در اختیار مصرف‌کنندگان قرار می‌گیرند.

در پاییز سال ۲۰۱۶، این گروه نیاز فوری به انتشار رهنمودهای مقدماتی بلندای صدا را تشخیص داد که پاسخ به نگرانی اساسی بلندای صدا در بخش در حال توسعه تلویزیون آنلاین و ارائه محتوای ویدئویی از طریق پخش و تجربه مصرف‌کننده بود. با انتشار رهنمودهای مقدماتی بلندای صدا، این گروه دانش خود را در مورد کارهای پیش رو، به‌طور جامع و مداوم افزایش داد و از همه علاقه‌مندان دعوت کرد تا در این مهم مشارکت کنند.

عنوان سند مرجع:

Loudness Guidelines for OTT and OVD Content

رهنمودهای بلندای صدا برای محتوای OTT و OVD^۴

ارائه‌دهنده:

Audio Guidelines for Over the Top Television and Video Streaming (AGOTTVS) Study Group

دبیر و ویراستار:

Jim Starzynski

زمان انتشار سند مرجع:

تیر ۱۴۰۰، July 2021

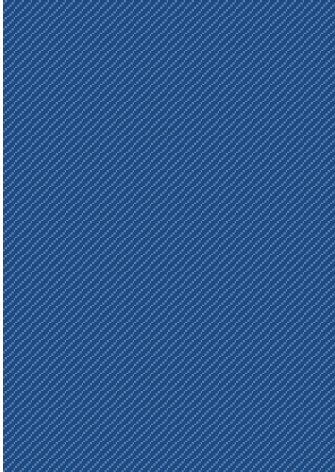
^۱ Audio Guidelines for Over the Top Television and Video Streaming (AGOTTVS) Study Group

^۲ Over the Top Television (OTT) For the purpose of this document OTT is defined as: The means to deliver video Content via Streaming, VOD, Pay TV, IPTV and Download via IP mechanisms

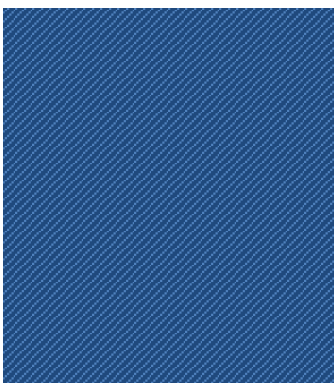
^۳ Online Video Distributor (OVD), defined as any entity that offers video Content by means of the Internet or other Internet Protocol (IP)-based transmission path provided by a person or entity other than the OVD

^۴ OTT and OVD do not include delivery of Content via means of traditional distribution e.g., broadcast TV, Cable TV, Satellite TV, and Telco Supplied TV Etc.

۶.....	مقدمه و معرفی.....	۱.
۶.....	مقدمه.....	۱.۱.
۶.....	اهداف:.....	۱.۲.
۶.....	نتایج پیروی.....	۱.۳.
۸.....	دستورالعمل ها.....	۲.
۸.....	تحويل محتوا بين توليدكننده و توزيع كننده.....	۲.۱.
۸.....	توزيع با فراداده ی مشخص.....	۲.۲.
۸.....	توزيع با فراداده ی نامشخص.....	۲.۳.
۹.....	توزيع با دامنه پويایی محدود.....	۲.۴.
۹.....	تحويل و توزيع با تنظيم قبلی.....	۲.۵.
۹.....	تحويل و توزيع محتوای غير منطبق.....	۲.۶.
۱۲.....	پيوست A ملاحظات محتوایی.....	۳.
۱۲.....	A.1 بلندای صدا.....	۳.۱.
۱۲.....	A.2 سطح اوج.....	۳.۲.
۱۳.....	A.3 دامنه پويایی.....	۳.۳.
۱۳.....	A.4 نوع محتوا.....	۳.۴.
۱۳.....	محتوای فرم کوتاه.....	۳.۴.۱.
۱۴.....	محتوای فرم بلند.....	۳.۴.۲.
۱۴.....	صداهاي توليد شده توسط سيستم:.....	۳.۴.۳.
۱۴.....	A.5 دامنه پويایی و کنترل بلندای صدا.....	۳.۵.
۱۴.....	محيط توليد و پخش.....	۳.۵.۱.
۱۵.....	محيط پخش.....	۳.۵.۲.
۱۷.....	پيوست B ملاحظات دستگاه OVD و OTT.....	۴.
۱۷.....	B.1 دستگاه های ثابت.....	۴.۱.
۱۷.....	کاربردها:.....	۴.۱.۱.
۱۷.....	قابلیت های عادی:.....	۴.۱.۲.
۱۸.....	B.2 دستگاه های پرتابل.....	۴.۲.
۱۸.....	کاربردها:.....	۴.۲.۱.
۱۹.....	قابلیت های معمول:.....	۴.۲.۲.
۲۱.....	پيوست C واژه نامه.....	۵.
۲۸.....	پيوست D جداول.....	۶.
۳۰.....	پيوست E خلاصه پيش آزمون بلندای صدای موبایل.....	۷.
۳۲.....	پيوست F نمودارهای اکوسيستم OVD و OTT.....	۸.
۳۴.....	مراجع.....	۹.



مقدمه و معرفی



۱. مقدمه و معرفی

۱.۱. مقدمه

محتوای تلویزیونی توزیع شده از طریق (OTT) یا به وسیله توزیع ویدیوی آنلاین (OVD) مستعد همان مشکلات مربوط به بلندای صداست که پیش از این در تلویزیون دیجیتال وجود داشت. این مشکل به دلیل افزایش قابل توجهی است که در دامنه پویایی^۵ صدا در مقایسه با اسلاف آنالوگ رخ داده است. همین موضوع باعث به وجود آمدن اختلاف صدای شدید بین برنامه ها و کانال هایی می شود که برای بینندگان فراهم است.

با انتشار توصیه های اندازه گیری بلندای صدا که در استاندارد ITU-R BS.1770 آمده است، سازمان های مختلف در سراسر دنیا دستورالعمل هایی را به طور مستقل جهت اجرا توسط مهندسين تخصصی تدوین نمودند. این دستورالعمل ها روی حفظ و بهبود کیفیت صوتی DTV و تجربه شنیداری به وسیله مدیریت بلندای صدای برنامه های گوناگون تمرکز کرده بودند.

به منظور امکان ارجاع به شیوه نامه و استاندارد واحد این سند در سازمان تنظیم مقررات صوت و تصویر فراگیر (ساترا) تنظیم شده است تا رسانه ها بر اساس آن نسبت به پایش، سنجش و کنترل و اصلاح بلندای صدا محتوای تولیدی و منتشره در رسانه فراگیر اقدام نمایند.

۱.۲. اهداف:

این سند، چالش های اساسی بلندای صدای OVD و OTT را با فراهم آوردن دستورالعمل های جدیدی برطرف می کند. این دستورات روی بلندای صدا و دامنه پویایی محتوای گیرنده ها و تجهیزات همراه تمرکز دارد.

۱.۳. نتایج پیروی

پیروی از این دستورالعمل ها نتایج زیر را به دنبال خواهد داشت.

۱. بلندای صدای ثابتی در برنامه های گوناگون، سرویس های مختلف و محتوای تبلیغاتی به وجود می آورد.
 ۲. دامنه بلندای پخش محتوا را برای تجهیزات مختلف و محیط های شنیداری گوناگون مناسب می کند.
 ۳. از محدود کردن شدید دامنه یا سایر پردازش ها که موجب کاهش کیفیت صدا می شود جلوگیری می کند.
 ۴. مقاصد هنری محتوا را حفظ می کند (فیلم ها، نمایش ها، موسیقی زنده ...)
- تجربه شنیداری را بهبود می بخشد.

^۵ Dynamic Range

دستور العمل

۲. دستورالعمل ها

برای حفظ صدای اصلی محتوای OTT و OVD، برای کاهش پرش صدای آزاردهنده هنگام سوئیچ محتوا، برای جلوگیری از پرش و جهت فراهم آوردن سطح شنیداری مناسب در دستگاه های مختلف، تولیدکنندگان و توزیع کنندگان لازم است تا از دستورالعمل های پایه ای پیروی کنند. این دستورات بسته به محتوای توزیع شده و کاربردها متفاوت هستند.

۲.۱. تحویل محتوا بین تولیدکننده و توزیع کننده

برای تحویل محتوا بین تولیدکنندگان و توزیع کنندگان توصیه می شود:

۱. وقتی اولویت و تنظیم قبلی جهت تحویل محتوا وجود ندارد، توصیه های برودکستر ملی (صداوسیما) باید مورد توجه قرار گیرد. (پیوست D، جداول ۱ و ۲)
۲. یک عنصر وزین^۶ برای مثال دیالوگ برای اندازه گیری بلندای صدا باید به کار گرفته شود. در یک برنامه کامل، اندازه گیری ها باید با توجه به دستورات برودکستر ملی (صداوسیما) باشد.
۳. فراداده ی کنترل دامنه پویایی^۷ و بلندای صدا که مطابق با محتوایست باید ضمیمه گردد.

۲.۲. توزیع با فراداده ی مشخص

برای توزیع محتوا در سیستم هایی با فراداده ی مشخص توصیه می گردد:

۱. وقتی تنظیم قبلی جهت توزیع محتوا وجود ندارد، توصیه های برودکستر ملی (صداوسیما) باید مورد توجه قرار گیرد. (پیوست D، جداول ۱ و ۲)
۲. یک عنصر وزین برای مثال دیالوگ برای اندازه گیری بلندای صدا باید به کار گرفته شود. در یک برنامه کامل، اندازه گیری ها باید با توجه به توصیه های برودکستر ملی (صداوسیما) باشد.
۳. فراداده ی کنترل دامنه پویایی و بلندای صدا که مطابق با محتوایست باید ضمیمه گردد.

۲.۳. توزیع با فراداده ی نامشخص

برای توزیع محتوا در سیستم هایی با فراداده ی نامشخص و ناشناخته توصیه می شود:

^۶ Anchor Element

^۷ Dynamic Range

۱. وقتی تنظیم قبلی جهت توزیع محتوا وجود ندارد، توصیه های برودکستر ملی (صداوسیما) باید مورد توجه قرار گیرد. (پیوست D، جداول ۱ و ۲)

۲. یک عنصر وزن دار^۸ برای مثال دیالوگ جهت اندازه گیری بلندای صدا باید به کار گرفته شود. در یک برنامه کامل، اندازه گیری ها باید با توجه به توصیه های برودکستر ملی (صداوسیما) باشد.

۲.۴. توزیع با دامنه پویایی محدود

برای توزیع محتوا برای تجهیزات یا شرایط شنیداری با دامنه پویایی محدود توصیه می شود:

۱. اگر سیستم توزیع، فراداده را پشتیبانی می کند، از وجود یک پروفایل^۹ DRC و فراداده جهت پخش در شرایط دامنه پویایی محدود اطمینان حاصل گردد.

۲. اگر سیستم توزیع از فراداده پشتیبانی نمی کند، درخواست بلندای و یا کنترل دامنه پویایی پیش از کدینگ ممکن است ضروری باشد تا صدا مشخص گردد.

۲.۵. تحویل و توزیع با تنظیم قبلی

برای تحویل و یا توزیع محتوا جایی که تنظیم قبلی و اولویت وجود دارد توصیه می گردد:

۱. حداکثر میزان بلندای و دامنه ماکزیمم از مقدار مشخص شده فراتر نرود. (پیوست D، جدول ۳)

۲. بلندای و دامنه ماکزیمم محتوا یا باید با تنظیمات پیشین منطبق باشد یا با توصیه های برودکستر محلی. (جدول ۱ و ۲ پیوست D)

۲.۶. تحویل و توزیع محتوای غیر منطبق^{۱۰}

برای توزیع یا تحویل محتوای تأیید نشده (محتوای تولید شده بدون توجه به توصیه ها یا استانداردها)

۱. برای توزیع محتوا در سیستم های با فراداده مشخص، بلندای و حداکثر دامنه باید طبق مقادیر توافق شده با تنظیمات پیشین باشد یا مطابق دستورات برودکستر ملی (صداوسیما). با توجه به ملاحظات قراردادی، هر محتوای نامنطبق باید:

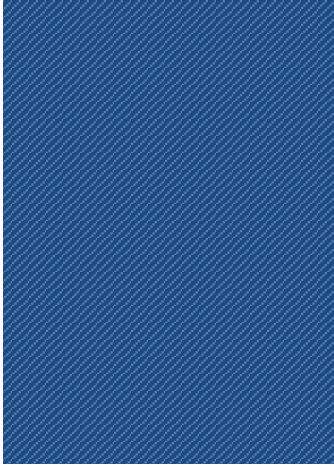
أ. به یک مقدار توافق شده با تنظیمات قبلی نرمالیزه شود.

^۸ Anchor Element

^۹ Dynamic Range Control

^{۱۰} non-compliant Content

- ب. به مقادیر توصیه شده توسط برودکستر ملی (صداوسیما) نرمالیزه شود.
- ج. به دلیل هماهنگی محتوا با فراداده ی DRC و بلندای صدا، تغییر نیافته باقی بماند.
۲. برای توزیع محتوا در سیستم های با فراداده ی نامشخص و نامعلوم، زمانی که تنظیمات قبلی وجود ندارد باید به توصیه های برودکستر محلی توجه شود و اندازه گیری ها طبق آن صورت پذیرد. (جدول ۱ و ۲ پیوست D). با توجه به ملاحظات قراردادی، هرگونه محتوای غیر منطبق، باید به توصیه های برودکستر نرمالیزه شود.



پوست A ملاحظات محتوایی



۳. پیوست A ملاحظات محتوایی

۳.۱. A.1 بلندای صدا

بلندای یک اندازه گیری ذهنی^{۱۱} از سطح صدای قابل درک از یک برنامه تلویزیونی یا یک محتوا است. (همان طور که توسط ITU-R BS1770 تعریف شده است) بلندای صدا یک برنامه در واقع همان بلندای صدای اندازه گیری شده در کل مدت آن برنامه و یا توسط عنصر وزین (مانند دیالوگ) اندازه گیری می شود.

نرمال سازی بلندای صدا^{۱۲} مقدار میانگین بلندای صدای برنامه ها را به دست می آورد در حالی که اجازه می دهد تا مقدار سطح اوج^{۱۳} با توجه به محتوا و نیازهای هنرمندانه و فنی آن متفاوت باشد. بالا بودن بلندای صدا برنامه و پایین بودن دامنه پویایی برنامه که ناشی از محدود بودن مقدار اوج و فشرده سازی دامنه پویایی است، برای متناسب کردن سیگنال صوتی در مقیاس کامل دیجیتال است.

برای دستیابی به بلندای صدای نرمال شده برای شنونده دو راه وجود دارد یک، نرمال سازی سیگنال صدا و دیگری استفاده از فراداده های بلندای صدا برای توصیف بلندای صدای یک برنامه. همه دستگاه هایی که افراد استفاده می کنند از راه حل فراداده پشتیبانی نمی کنند بنابراین روش قبلی توصیه می گردد، با داشتن فراداده های مناسب برای ایجاد بهترین سازگاری در بین دستگاه های مصرف کننده.

۳.۲. A.2 سطح اوج

سطح اوج در سیگنال صوتی معمولاً بر میزان بلندای صدا تأثیر نمی گذارد. با این حال، تأثیر قابل توجهی بر کیفیت صدای درک شده دارد. برنامه ای که مقدار سطح اوج آن بالا باشد نسبت به برنامه ای که سطح اوج آن محدود باشد معمولاً صدای واضح تر دارد و همچنین بینندگان را کمتر خسته می کند.

ITU-R BS.1770 با توجه به این واقعیت که پیک صدا واقعی ممکن است بین نمونه های دیجیتال رخ دهد و حتی از هر نمونه دیجیتالی در برنامه بالاتر باشد و ممکن است بالاتر از 0 dB FS باشد، اوج حقیقی^{۱۴} را به عنوان حداکثر سطوح اوج معرفی می کند. به تولیدکنندگان توصیه می شود که محتوا را با سطح اوج حقیقی پایین تر از 0 dB FS (به پیوست D مراجعه کنید) تولید کنند تا از قطع شدن در سیستم های رمزگشایی و پخش جلوگیری شود.

^{۱۱} Subjective^{۱۲} Loudness Normalization^{۱۳} Peak Level^{۱۴} True Peak

هرچه بلندای صدا برنامه کم تر باشد، کنترل اوج با احتمال کمتری مورد نیاز خواهد بود. برنامه هایی که نیاز دارند سطح صدا افزایش یابد تا بلندای صدای برنامه مورد نیاز را برآورده کنند باید به حداکثر سطوح اوج حقیقی توجه ویژه کنند.

۳,۳. A.3 دامنه پویایی

به شدت تشویق می شود که دامنه پویایی اصلی یک برنامه برای حفظ یکپارچگی صوتی و هدف هنری محتوا حفظ شود. با این حال موقعیت هایی وجود دارند که در آن ها ممکن است این امکان وجود نداشته باشد و یا مطلوب نباشد:

ا. بلندای مورد نیاز پخش شده بالاتر از بلندای صدای برنامه باشد.

ب. دامنه پویایی محتوای تحویل داده شده از دامنه پویایی سیستم مصرف کننده بیشتر باشد.

ج. وجود افراد در یک محیط پرسروصدا و نویزی مانند هواپیما.

دامنه پویایی سیستم یک نسبتی بین مقادیر کمینه و بیشینه سطح سیگنال است که یک سیستم می تواند آن را با دقت خاصی انتقال دهد. در بعضی از کشورها (به دلیل مقررات تنظیم گری دولت ها)، ممکن است خروجی تلفن های هوشمند نسبت به خروجی های لپ تاپ یا دستگاه های ثابت از دامنه پویایی پایین تری برخوردار باشند. برخی از برنامه ها ممکن است از دامنه پویایی محتوای پخش معمول فراتر روند. به عنوان مثال، محتوا تولید شده در سینما که در آن ها بلندای صدای پایین تر، رایج است. در این موارد، باید از عنصر وزین (به عنوان مثال، دیالوگ) برای اندازه گیری بلندای صدای یکپارچه، به جای اندازه گیری ترکیب کل برنامه برای شرایط ذکر شده در توصیه های ملی (صداوسیما) استفاده شود.

۳,۴. A.4 نوع محتوا

۳,۴,۱. محتوای فرم کوتاه^{۱۵}

اندازه گیری محتوای فرم کوتاه مانند تبلیغات باید با دقت لازم انجام شود. اندازه گیری طولانی مدت از همه کانال ها برای کل محتوا باید در نظر گرفته شود و یا به دستورالعمل های اندازه گیری ملی (صداوسیما) برای محتوای فرم کوتاه مراجعه شود.

برای تمامی موارد و همه مناطق، جدول ۱ رجوع شود.

^{۱۵} Short Form Content

۳.۴.۲. محتوای فرم بلند

اندازه گیری بلندای صدا محتوای فرم بلند مانند برنامه های معمول تلویزیونی مانند برنامه های ورزشی، اخبار و ... باید با دقت شود.

برای مناطق ATSC، عنصر Anchor (به طور معمول دیالوگ) و یا یک نمونه از عنصر Anchor باید با استفاده از BS.1770 اندازه گیری شود. اگر عنصر Anchor قابل جداسازی و اندازه گیری نباشد، متوسط بلندای صدا کل برنامه باید اندازه گیری و گزارش شود. [1]

برای مناطق EBU، اندازه گیری طولانی مدت از همه کانال ها برای کل محتوا رایج ترین روش اندازه گیری و مناسب ترین روش برای اکثر محتوای broadcast است. برای برنامه های صوتی که دامنه پویایی گسترده ای دارند که LRA آنها بیش از ۲۰ LU است (به EBU Tech 3343 sec 4.1 مراجعه کنید) یا در مواردی که سطح بلندای صدا از مقدار اندازه گیری شده برنامه بیش از ۳-۵ LU تفاوت داشته، ممکن است از نرمال سازی بلندای صدا بر پایه Anchor استفاده شود. [2]

برای تمامی موارد و همه مناطق، جدول ۲ رجوع شود.

۳.۴.۳. صداهای تولید شده توسط سیستم:

بلندای صدا منابع صوتی دیگر مانند دستگاه های پخش ثابت و تلفن همراه بخشی از تجربه شنیداری هستند. توسعه دهندگان برنامه ها باید تلاش کنند تا بلندای صدای تولید شده در این سیستم ها با محتوا تلویزیونی و یا ویدیوی مطابقت داشته باشد تا در صورت لزوم به تجربه دل پذیری دست پیدا کنند.

با وجود این، درک می شود که بلندای صدا نقش مهمی در تشخیص "حلقه های" ورودی بر روی محتوا، محلی سازی صدا در تصویر، هشدار و ارائه اطلاعات اضطراری مهم (به عنوان مثال هشدارهای اضطراری بی سیم Alerts-WEA) و غیره در صورت لزوم ایفا می کند. در این موارد و سایر شرایط مشابه، انتظار می رود که میزان بلندای صدا تولید شده توسط سیستم های تولید صدا و محتوای مربوط به آن بیشتر از سایر صداها باشد.

۳.۵. A.5 دامنه پویایی و کنترل بلندای صدا

۳.۵.۱. محیط تولید و پخش

اکیداً توصیه می شود که همیشه هنگام استفاده از سیستم پخش، بلندای صدا و فراداده های کنترل دامنه پویایی را با محتوای موجود مطابقت دهید. این روش اجازه می دهد تا محتوا برای دستگاه های مختلف و سناریوهای شنیداری هنگام پخش به صورت غیرمخرب طراحی شود. استفاده و پشتیبانی از بلندای صدا و فراداده کنترل دامنه پویایی، از جمله پشتیبانی در CODEC های صوتی، در ATSC A / 85 [۱] و EBU Tech Doc 3344 [7] بیان و مستند شده است.

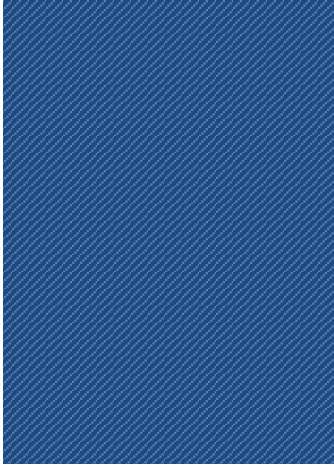
یک روش CODEC agnostic برای انتقال و حمل بلندای صدا و فراداده کنترل دامنه پویایی در MPEG-D DRC (ISO / IEC 23003-4) ثبت شده است. [5]

هنگام نوشتن فراداده محتوا برای پخش، استفاده از مشخصات یا نمایه های تعریف شده DRC توصیه می شود. این روش فرصتی را فراهم می کند تا در صورت لزوم بلندای صدا و دامنه پویایی DRC Profile را به درستی در دستگاه پخش تنظیم کند.

۳.۵.۲. محیط پخش^{۱۶}

در حال حاضر تعداد زیادی از دستگاه ها وجود دارند که از چندین CODEC استفاده می کنند. ممکن است یک یا چند مورد از این CODEC ها از فراداده پشتیبانی نکنند. در حالی که بسیاری از این دستگاه ها با پیروی از توصیه های CTA CEB11 [6] و EBU R-128 / Tech 3344 [7]، این مسئله را با موفقیت برطرف کرده اند.

^{۱۶} PlayBack



پوست B ملاحظات دستگاه



۴. پیوست B ملاحظات دستگاه OTT و ODD

دستگاه های قابل حمل در مک آن های مختلفی استفاده می شوند که شرایط محیطی متنوعی دارند. از میان شنوندگان انتخاب شده، بیشتر آنان از هدفون استفاده می کنند. در محیط های پرسروصدا می توان از DRC^{۱۷} برای شنیدن صدای آرام معابر استفاده کرد. در خانه، افراد همچنین می توانند دستگاه های خود را مستقیماً از طریق HDMI به تلویزیون یا AVR متصل کنند یا از طریق فناوری های بی سیم مانند cast استفاده کنند. همه این سناریوها چالش های عمده ای را پیرامون کیفیت صدای بازتولید شده ایجاد می کنند. محتوای توزیع شده به همراه CODEC با استفاده از فراداده صحیح از قابلیت های ارتباطی دستگاه از طریق اینترفیس های رسانه ای سود می برد. HDMI و حوزه داده های EDID می توانند سازگاری CODEC را بین دستگاه ها شناسایی کنند که به عنوان مثال تنظیمات خاصی که شنوندگان برای میزان بلندای صدا و DRC انجام می دهند را تسهیل کند. محتوای توزیع شده همراه با CODEC که از فراداده استفاده نمی کند، زمانی که به دستگاه دیگری متصل گردد، پردازش صوتی که منجر به نمایش غیرقابل پیش بینی آن می شود به نرم افزار، سخت افزار یا ترکیب هر دو آن وابسته خواهد بود.

۴.۱. B.1 دستگاه های ثابت

محیط های پخش خانگی بسیار متفاوت است. سطح سروصدا از منابع خارجی و داخلی مانند ترافیک اطراف وسایل خانگی، سایر اعضای خانواده، فاصله از منبع صدا، بازتاب صدای اتاق، کیفیت و جایگیری بلندگوها تأثیر زیادی بر صدایی که شنوندگان دریافت می کنند دارد. در همه سناریوها، قابل درک بودن صدای ترکیب شده توسط شنونده عامل مهمی محسوب می شود.

۴.۱.۱. کاربردها:

دستگاه های ثابت معمولاً برای پخش رسانه در خانه یا سایر محیط های داخلی در نظر گرفته شده اند مانند اتاق های نشیمن، اتاق های هتل، فضاهای اداری، کلاس ها و غیره

۴.۱.۲. قابلیت های عادی:

دستگاه های ثابت در نظر گرفته شده اند تا به تلویزیون ها، ساندمبار^{۱۸} ها یا گیرنده های صوتی و تصویری متصل گردند. محتواها از طریق موارد زیر استریم می شوند:

- Over The Top Television (OTT) applications
- (OVD) Online Video Distribution سرویس پخش زنده یا
- the Air traditional broadcast

^{۱۷} Dynamic Range Control

^{۱۸} Soundbars

▪ MVPD (Cable, Satellite, Telco) Transport Stream or IP services

برخی از دستگاه های ثابت از هارد دیسک های داخلی برای ذخیره سازی (از طریق بارگیری^{۱۹} یا بارگیری جانبی^{۲۰}) یا ضبط برنامه ها برای پخش در مدت زمان طولانی استفاده می کنند. بسیاری از دستگاه ها از پخش Blu-ray Disc یا DVD نیز پشتیبانی می کنند. این دستگاه ها توانایی تولید دامنه پویایی صوتی کل محتوا را دارند که در استریم صدا ورودی ارائه می شود. دستگاه ها همچنین حاوی سخت افزار و نرم افزار پردازش سیگنال دیجیتال^{۲۱} و رمزگشایی هستند که می توانند صدا را با توجه به نیاز شنونده تراز کنند. نمونه هایی از این موارد عبارت اند از:

- Speaker/room equalization systems
- added room reverberation
- independent dynamic range control compression
- Metadata enabled decoder Loudness Normalization and dynamic range control

این ویژگی ها می توانند به طور مستقل یا به صورت ترکیبی مورد استفاده قرار گیرند و به افراد در تنظیم دقیق تجربه شنیداری خود کمک می کند.

۴.۲. B.2 دستگاه های پرتابل

کاربردهای بلندای صدا و کنترل بلندای صدا^{۲۲} در محیط های معمول پخش

مشابه ویژگی های بیان شده در قسمت دستگاه های ثابت، دستگاه های قابل حمل می تواند بلندای صدا های مختلف و پیکربندی های مربوط به پردازش دامنه پویایی را ارائه دهند. این سازگاری کلیدی برای بهینه سازی صدای پخش شده از پخش کننده جریان^{۲۳} ها برای محدودیت های مطرح شده توسط محیط های شنیداری آکوستیک است.

۴.۲.۱. کاربردها:

این دستگاه ها در خانه و فضای باز استفاده می شوند. اکثر گوشی های دستی و تبلت ها به عنوان پخش کننده های شخصی صوتی شروع به کار کردند و به دستگاه های ارتباطی کاملاً پیشرفته با قابلیت های پخش صوت و تصویر تبدیل شدند. در این گروه تلفن هوشمند دستگاه اصلی محسوب می شود.

^{۱۹} Download

^{۲۰} Side-Load

^{۲۱} DSP: Digital Signal Processing

^{۲۲} DRC

^{۲۳} Streaming Players

۴.۲.۲. قابلیت های معمول:

تلفن های همراه و تبلت ها را می توان به عنوان پخش کننده های شخصی که به صورت محلی اطلاعات را بر روی خود ذخیره می کند استفاده کرد و یا می توان از محتواهایی که از طریق Wi-Fi یا شبکه های سلولی به اینترنت دسترسی دارند، استفاده کرد.

محتواها می توانند:

- ابتدا Download شوند و سپس بعداً پخش شوند.
- ابتدا Side-load شوند و سپس بعداً پخش شوند.
- با استفاده از بیت ریت های مناسب استریم شوند.

عوامل مختلفی که بر روی کیفیت صدا دستگاه های پرتابل تأثیر می گذارد:

- پهنای باند موجود هنگام دستیابی به محتوا
- حافظه محلی موجود در دستگاه پرتابل
- پشتیبانی از DSP و یا CODEC در دستگاه پرتابل
- پردازش قدرت دستگاه پرتابل برای رمزگشایی مبتنی بر برنامه، پردازش قدرت مربوط به DSP
- حداکثر کیفیت موجود در adaptive bitrate stream

عوامل دیگری نیز در پیکربندی های خروجی دستگاه پرتابل وجود دارد. بیشتر آن ها می توانند دو یا چند

مورد زیر را پشتیبانی کنند:

- بلندگو (های) داخلی استریو یا مونو با دامنه پویایی و پاسخ فرکانسی محدود
- دستگاه های متصل (به عنوان مثال: هدفون) با دامنه محدود (در مقایسه با دستگاه های متصل در حالت دامنه کامل) و پاسخ فرکانسی متنوع
- دستگاه های بلوتوث از جمله هدفون، بلندگو و AVR ها. دامنه پویایی و پاسخ فرکانسی محتوا می تواند توسط عواملی مانند نسخه های بلوتوث، (CODEC) و پهنای باند موجود محدود گردد.
- پخش کننده های Wi-Fi و بلوتوث (مانند: AirPlay، ChromeCast) بسته به پهنای باند موجود می تواند از پخش با کیفیت پشتیبانی کند.
- HDMI (اتصال مستقیم باسیم) می تواند یک استریم صوتی کد شده بدون تغییر ارائه دهد که به وسیله دستگاهی با قابلیت صوتی پیشرفته رمزگشایی^{۲۴} شود؛ یا یک نسخه رمزگشایی شده صوتی PCM، مشروط بر پردازش صوتی پیشرفته فراداده یا پردازش داخلی

^{۲۴} Decode

پوست C و اثره نامه

۵. پیوست C واژه نامه

Anchor Element

The perceptual Loudness reference point or element around which other elements are balanced in producing the final mix of the Content, or that a reasonable viewer would focus on when setting the volume control. [3] It is typically speech or dialog in broadcast or movie Content.

AVR**Audio Video Receiver**

A combination audio amplifier and audio/video switching device for a home theater. It contains inputs for all the audio and video sources and outputs to one or more sets of speakers and one or more monitors (without a tuner) or TVs [۱۰].

Blu-ray

A digital optical disc data storage format supporting HD and UHD (2160p) resolution.

BS.1770**ITU-R BS.1770**

Specifies an algorithm that provides a numerical value that estimates the perceived Loudness of the Content that is measured. Loudness meters and measurement tools which have implemented the

BS.1770 algorithm will report Loudness in units of “LKFS” or “LUFS.”

CODEC

Coder-Decoder. A device (or software implementation thereof) which converts a signal into another format and converts the coded signal back to the original format.

Content

Material or essence used for distribution by an operator.

Content Distributor

An operator that distributes Content to other operators, Programmers or consumers.

Content Dynamic Range

The difference between the maximum and minimum Loudness level in an audio Program or piece of Content. It should be noted that a large difference between the Loudness of the dialog and the Loudness of the Full Program Mix Measurement may be an indication of wide Content

Dynamic Range. Similarly, large values of LRA may be an indication of wide Content Dynamic Range.

dB FS

Amplitude expressed as a level in decibels relative to 100% full-scale.

Dialog Level

The Loudness, in LKFS units, of the Anchor Element.

Download

The process of receiving data from a remote system, typically a server such as a web server , an FTP server, an email server, or other similar system.

DRC

Dynamic range control. The process of continually adjusting audio signal level to control the level difference between loud and soft passages according to some desired objective.

DRC Profile

A collection of parameters that describe how dynamic range control Metadata is calculated.

DVR**Digital Video Recorder**

An electronic device that records video in a digital format to a disk drive, USB flash drive, SD memory card, SSD or other local or networked mass storage device

EDID

Extended Display Identification Data (EDID) is a data structure provided by a display or other device to describe its capabilities to a video source device (e.g. graphics card or set-top box). [8]

File-Based-Scaling Device

A device used to apply an overall gain correction to audio Content stored as files.

Fixed Device

A set-top box type or plug-in stick type unit capable of supporting streaming media playback services via Ethernet and/or Wi-Fi Internet access.

Full Program Mix Measurement

The Integrated Loudness measurement using the BS.1770 algorithm, in units of LKFS or LUFS based on all channels and all elements of the audio Content, over the duration of the Program.

Game Console

A computer system that enables playing video games via a connection to a television or other display device with internal or external speakers. Most Game Consoles also support streaming internet video/audio over Ethernet/Wi-Fi or via internal disk player (Blu-ray, DVD) or removable digital media flash storage (SD card).

HDMI Stick

A plug-in device (via HDMI) for TV or AVR and supports streaming internet video/audio over Wi-Fi.

Integrated Loudness

An objective measurement of audio loudness of a program averaged over the entire length of the program.

Interstitial Content

Advertising, commercial, promotional or public service related material or essence. The typical duration is less than approximately two to three minutes. Synonymous with Short Form Content.

ITU

International Telecommunication Union. An organization which coordinates electronic communications across national boundaries and issues recommendations supporting those activities.

LKFS

Unit of Loudness, K-weighted, relative to full scale, measured with equipment that implements the algorithm specified by BS.1770. A unit of LKFS is equivalent to a decibel.

Long Form Content

Show or Program related material or essence. The typical duration is greater than approximately two to three minutes.

Loudness

A perceptual quantity; the magnitude of the physiological effect produced when a sound stimulates the ear.

Loudness Normalization

The process of applying a global offset, based on the measured Integrated Loudness of an audio Program, such that after the offset is applied, the audio Program will be at the desired target Loudness.

Loudness Range**LRA**

Quantifies the variation in a time-varying Loudness measurement.

LU**Loudness Unit**

Relative Loudness is expressed as a logarithmic ratio, 1 dB increase in level equals 1 LU increase in Loudness.

LUFS

Unit of Loudness, K-weighted, expressed as a level in LU relative to 100% full-scale.

Maximum True Peak Level

The maximum absolute value of the audio signal waveform of a Program in the continuous time domain.

Measured Loudness

The value reported, in units of LKFS or LUFS, when an audio signal is measured with equipment that implements the algorithm specified by BS.1770. It is an approximation of perceived Loudness.

Media Streamer**Streaming Media Set Top Box**

Modern set-top boxes that deliver Internet streaming Content directly to a T.V. Many media streamers are also capable of storing Content on a hard drive, computer or Network attached Storage.

Metadata**Audio Metadata**

Information attached to the audio bitstream, for example, to describe the audio essence or to support DRC.

MVPD

Multichannel Video Programming Distributor (includes DBS service operators, local cable system operators, and cable multiple system operators).

OTT**Over-the Top-Television**

The means to deliver video Content via Streaming, VOD, Pay TV, IPTV and Download via IP mechanisms.

OVD**Online Video Distributor**

Any entity that offers video Content by means of the Internet or other Internet Protocol (IP)-based transmission path provided by a person or entity other than the OVD.

Peak Limiting - Limiting

A process allowing signals below a specified input power or level to pass unaffected, while attenuating the peaks of stronger signals that exceed this threshold. Limiting is a type of dynamic range compression. Clipping is an extreme version of limiting.

Personal Computer

A compact computer that uses a microprocessor and is designed for individual use, as by a person in an office or at home or school, for such applications as word processing, data management, financial analysis, streaming media or computer games.

PMP**Portable Media Player**

A portable consumer electronics device capable of storing and playing digital media such as audio, images, and video files. Some devices also support streaming internet video/audio over Wi-Fi.

Portable Device

Personal type units capable of supporting file and streaming media playback services via Wi-Fi and/or cellular data internet access.

PPM

Peak Program meter. A device which measures and displays or records the maximum excursion of an audio program.

Program

An individual, self-contained audio-visual or audio-only item to be presented in Radio, Television or other electronic media. An advertisement (commercial), trailer, promotional item ('promo'), interstitial or similar item shall be considered to be a Program in this context.

Program Loudness

The Integrated Loudness over the duration of a Program.

Set Top Box

A hardware device that allows a digital signal to be received, decoded and displayed on a television. The signal can be a television signal or Internet data and is received via an MVPD.

Short Form Content

Advertising, commercial, promotional or public service related material or essence. The typical duration is less than approximately two to three minutes. Synonymous with interstitial Content.

Sideload

Media file transfer to a mobile device via USB, Bluetooth, Wi-Fi or by writing to a memory card for insertion into the mobile device.

Smart Phone

A device that combines a cell phone with a hand-held computer and Portable Media Player capabilities.

Smart T.V.

A digital television that has Fixed Device capabilities built in.

SPL

Sound pressure level in decibels referenced to 20 $\mu\text{N/m}^2$.

Tablet Computer

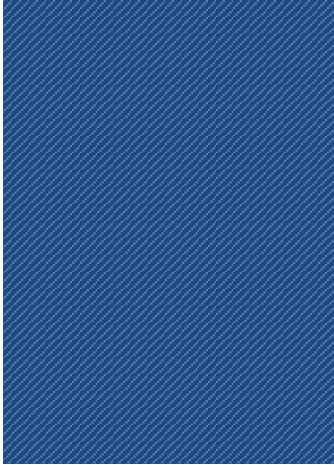
A very thin, portable computer, usually battery-powered, having a touchscreen as the primary interface and input device. Usually has Portable Media Player capabilities and may also support cellular data.

Target Program Loudness

A specified Loudness value for the Anchor Element, Dialog Level or Program Loudness, established to facilitate Content exchange from a supplier to an operator.

True Peak

The maximum absolute level of the signal waveform in the continuous time domain, measured per BS.1770. Its units are dB TP.



پوست D جداول



۶. پیوست D جداول

Table 1 - Recommendations for Short Form Content:

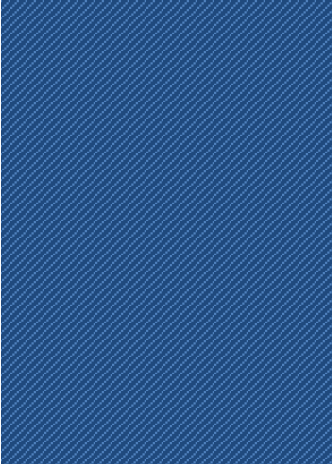
Broadcast Region	Integrated Loudness	Maximum Short Term Loudness	Maximum True Peak	Anchor Based Measurement	Full Program Mix Measurement
North America [1]	-24 ± 2 LKFS	N/A	-2 dB TP	Not permitted	Recommended
Europe [2]	-23 ± 0.5 LUFS	-18 LUFS and + 5 LU relative to IL	-1 dB TP	Not permitted	Recommended
Japan [3]	-24 ± 1 LKFS	N/A	-1 dB TP	Not permitted	Recommended
Australia [4]	-24 ± 1 LKFS	N/A	-2 dB TP	Not permitted	Recommended

Table 2 - Recommendations for Long Form Content:

Broadcast Region	Integrated Loudness	Maximum Short Term Loudness	Maximum True Peak	Anchor Based Measurement	Full Program Mix Measurement
North America [1]	-24 ± 2 LKFS	N/A	-2 dB TP	Recommended	Conditionally Permitted ⁴
Europe [2]	-23 ± 0.5 LUFS	N/A	-1 dB TP	Conditionally Permitted ⁵	Recommended
Japan [3]	-24 ± 1 LKFS	N/A	-1 dB TP	Not permitted	Recommended
Australia [4]	-24 ± 1 LKFS	N/A	-2 dB TP	Permitted	Permitted

Table 3 - Recommendations for maximum Loudness: [23]

	Integrated Loudness	Maximum True Peak
Short Form Content	-16 ± 1 LKFS/LUFS ⁶	-1 dB TP
Long Form Content	-16 ± 1 LKFS/LUFS ⁵	-1 dB TP



پوست E خلاصه پیش آزمون بلندای

صدای موبایل



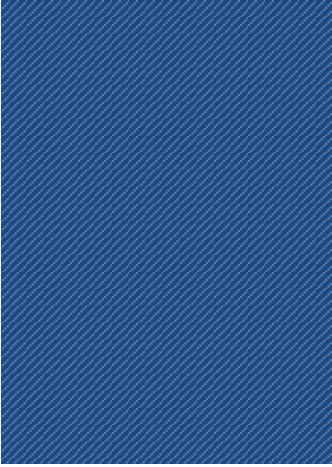
۷. پیوست E خلاصه پیش آزمون بلندای صدای موبایل

گروه AGOTTVS یک نظرسنجی غیررسمی بلندای صدا را در اوت ۲۰۱۷ با شنوندگان متخصص و غیرمتخصص از مناطق مختلف جهان انجام داد و شامل چندین عضو AGOTTVS بود.

این آزمایش با استفاده از هفت محتوای broadcast گزینش شده توسط اعضای AGOTTVS انجام شد و به سه هدف مختلف با بلندای صدا، ۲۴-، ۲۰- و ۱۶- LKFS نرمالایز شد تا مشخص بشود که آیا می توان میزان صدای دستگاه های پرتابل معمولی را برای تولید نتایج قابل فهم تنظیم کرد یا نه. شرکت کنندگان هم در محیط های آرام و هم در محیط های پرسروصدا کلیپ های صوتی را از طریق هدفون و بلندگوی داخلی دستگاه ها، گوش می کردند.

نتایج نشان داد که در محیط های آرام، همه کلیپ ها را می توان در یک سطح قابل فهم قرارداد، اگرچه بخشی از کلیپ ها به عنوان "بسیار بلند" گزارش شده بودند. در محیط های پرسروصدا، اکثر کلیپ ها را می توان در یک سطح قابل فهم قرار داد، اگرچه بخش های بیشتری به عنوان "بسیار بلند" گزارش شد.

به طور خلاصه، این کار مقدماتی این اطمینان را می دهد که سه سطح هدف، عملی هستند، اما کنترل دامنه پویایی به نوعی برای محیط های پرسروصدا مورد نیاز است. همچنین این موضوع تصریح می کند که به آزمایش دقیق تری نیاز است.



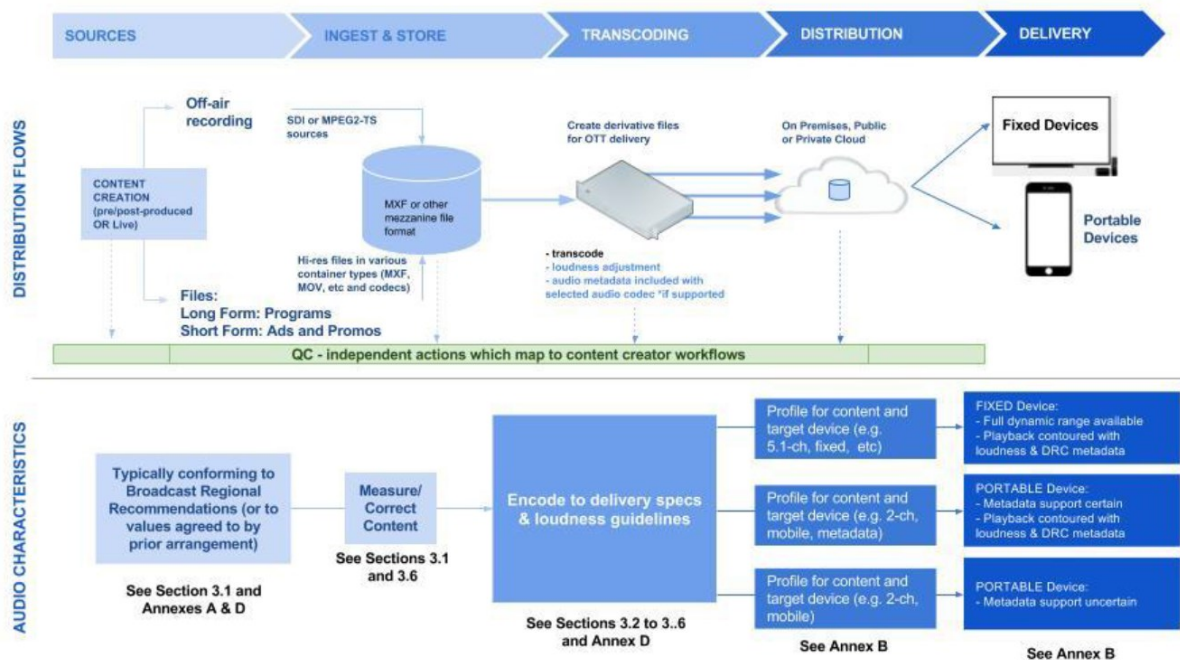
پوست F نمودارهای اکوسیستم OTT و

OVD

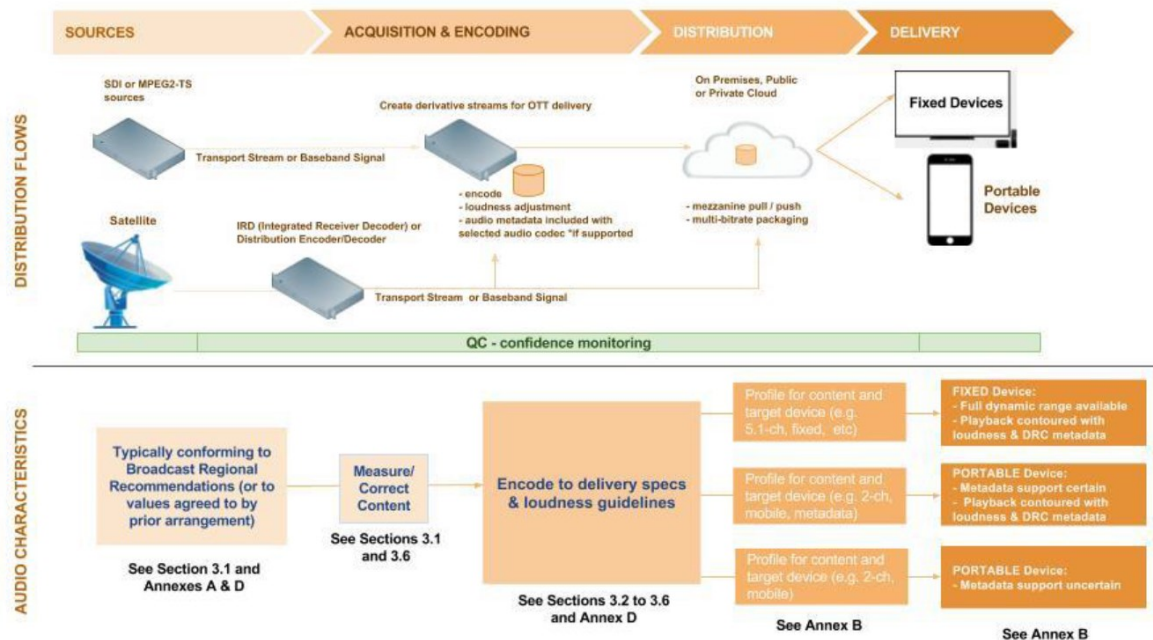


۸. پیوست F نمودارهای اکوسیستم OVD و OTT

OTT / OVD AUDIO WORKFLOW



LINEAR STREAMING AUDIO WORKFLOW



مراجع

۹. مراجع

Ref

[1] Advanced Television Systems Committee (ATSC) "Techniques for Establishing and Maintaining Audio Loudness for Digital Television" A/85-2013 (March, 2013)

<https://www.atsc.org/recommended-practice/a85-techniques-for-establishing-and-maintaining-audio-loudness-for-digital-television/>

[2] European Broadcast Union (EBU) "Loudness Normalisation and Permitted Maximum Level of Audio Signals" R 128 s1 ver. 2.0 (June 2014)

<https://tech.ebu.ch/docs/r/r128.pdf>

[3] Association of Radio Industries and Businesses (ARIB) "Operational Guidelines for Loudness of Digital Television Programs" TR - B32 (September, 2016)

https://www.arib.or.jp/english/std_tr/broadcasting/trb32.html

[4] Free TV Australia "Measurement and Management of Loudness in Soundtracks for Television Broadcasting" OP-59 (Feb 2013)

http://www.freetv.com.au/media/Engineering/Free_TV_OP59_Measurement_and_Management_of_Loudness_for_TV_Broadcasting_Issue_3_July_2016.pdf

[5] International Organization for Standardization ISO/IEC JTC1/SC29/WG11 Coding of Moving Pictures and Audio "White Paper on MPEG-D Dynamic Range Control" N15071 (Feb 2015)

http://www.mpeg-audio.org/docs/w15071_MPEG-D_DRC.pdf

[6] Consumer Technology Association (CTA) "NTSC/ATSC Loudness Matching" CTA-CEB11-C (January 2015)

<https://www.cta.tech/Research-Standards/Standards-Listing.aspx>

[7] European Broadcast Union (EBU) "Guidelines for Distribution and Reproduction in Accordance with EBU R 128" Tech 3344 version 2.1 (July 2016)

<https://tech.ebu.ch/docs/tech/tech3344.pdf>

[8] Wikipedia -The Free Encyclopedia "EDID" Definition

https://en.wikipedia.org/wiki/Extended_Display_Identification_Data