



**ESPE**  
**UNIVERSIDAD DE LAS FUERZAS ARMADAS**  
**INNOVACIÓN PARA LA EXCELENCIA**

**Desarrollo de un sistema de transcripción de audio a texto para generación y sincronización de subtítulos en transmisión de tv digital.**

Real Lalaleo, Melissa Elizabeth

Departamento de Eléctrica, Electrónica y Telecomunicaciones

Carrera de Ingeniería en Electrónica y Telecomunicaciones

Trabajo de titulación previo a la obtención del título de Ingeniera en Electrónica Y Telecomunicaciones.

Dr. Olmedo Cifuentes, Gonzalo Fernando

17 de junio del 2020



### Document Information

Analyzed document Real\_Melissa.docx (D75697750)  
Submitted 6/27/2020 2:29:00 AM  
Submitted by  
Submitter email mgutierrez@difusion.com.mx  
Similarity 5%  
Analysis address mgutierrez1.GDC@analysis.arkund.com

### Sources included in the report

URL: Tesis Final Ma Liliana Pinza.docx  
Fetched: 11/9/2016 11:06:00 PM

Firma:



---

**Dr. Olmedo Cifuentes, Gonzalo Fernando**

CC:1711696342

**DIRECTOR**



**DEPARTAMENTO DE ELÉCTRICA, ELECTRÓNICA Y  
TELECOMUNICACIONES**

**CARRERA DE INGENIERÍA EN ELECTRÓNICA Y  
TELECOMUNICACIONES**

**CERTIFICACIÓN**

Certifico que el trabajo de titulación, : **“Desarrollo de un sistema de transcripción de audio a texto para generación y sincronización de subtítulos en transmisión de tv digital.”**, fue realizado por la señorita **Real Lalaleo, Melissa Elizabeth**, el cual ha sido revisado y analizado en su totalidad por la herramienta de verificación de similitud de contenido; por lo tanto cumple con los requisitos legales, teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, razón por la cual me permito acreditar y autorizar para que lo sustente públicamente.

**Sangolquí, 2 de julio del 2020**

Firma:



---

Dr. Olmedo Cifuentes, Gonzalo Fernando

C.C. 171169634-2



# ESPE

UNIVERSIDAD DE LAS FUERZAS ARMADAS  
INNOVACIÓN PARA LA EXCELENCIA

DEPARTAMENTO DE ELÉCTRICA, ELECTRÓNICA Y  
TELECOMUNICACIONES

CARRERA DE INGENIERÍA EN ELECTRÓNICA Y  
TELECOMUNICACIONES

#### RESPONSABILIDAD DE AUTORÍA

Yo, **Real Lalaleo, Melissa Elizabeth**, con cedula de identidad N° 1803178910 declaro que el contenido, ideas y criterios del trabajo de titulación “**Desarrollo de un sistema de transcripción de audio a texto para generación y sincronización de subtítulos en transmisión de tv digital.**” es de mi autoría y responsabilidad, cumpliendo con los requisitos legales, teóricos, científicos, técnicos, y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

Sangolquí, 02 de julio del 2020

Firma

**Real Lalaleo, Melissa Elizabeth**  
C.C. 1803178910



# ESPE

UNIVERSIDAD DE LAS FUERZAS ARMADAS  
INNOVACIÓN PARA LA EXCELENCIA

DEPARTAMENTO DE ELÉCTRICA, ELECTRÓNICA Y  
TELECOMUNICACIONES

CARRERA DE INGENIERÍA EN ELECTRÓNICA Y  
TELECOMUNICACIONES

#### AUTORIZACIÓN DE PUBLICACIÓN

Yo **Real Lalaleo, Melissa Elizabeth**, con cédula de ciudadanía n°1803178910, autorizo a la Universidad de las Fuerzas Armadas ESPE publicar el trabajo de titulación “**Desarrollo de un sistema de transcripción de audio a texto para generación y sincronización de subtítulos en transmisión de tv digital.**” en el Repositorio Institucional, cuyo contenido, ideas y criterios son de mi responsabilidad.

Sangolquí, 02 de julio del 2020

Firma

Una firma manuscrita en tinta azul que parece decir 'Melissa Elizabeth Real Lalaleo'.

---

**Real Lalaleo, Melissa Elizabeth**  
C.C. 1803178910

## DEDICATORIA

*En memoria de mi abuelita quien siempre tuvo las palabras exactas  
para impulsarme a seguir adelante.*

## **AGRADECIMIENTO**

Agradezco a mis padres por ser mi mayor ejemplo, quienes con su apoyo, esfuerzo, dedicación y paciencia me han ayudado a cumplir todas las metas a lo largo de mi vida

A mis amigos que han permanecido a mi lado en este trayecto convirtiéndose en consejeros, maestros y compañeros, aportando a mi crecimiento profesional y personal.

A mi director de tesis el Dr. Gonzalo Olmedo que supo guiarme de la mejor manera con sus consejos y recomendaciones.

## ÍNDICE DE CONTENIDOS

<b>CERTIFICADO DEL DIRECTOR .....</b>	<b>2</b>
<b>RESPONSABILIDAD DE AUTORÍA.....</b>	<b>3</b>
<b>AUTORIZACIÓN DE PUBLICACIÓN.....</b>	<b>4</b>
<b>DEDICATORIA .....</b>	<b>6</b>
<b>AGRADECIMIENTO .....</b>	<b>7</b>
<b>ÍNDICE DE CONTENIDOS .....</b>	<b>8</b>
<b>ÍNDICE DE TABLAS .....</b>	<b>10</b>
<b>ÍNDICE DE FIGURAS .....</b>	<b>11</b>
<b>RESUMEN .....</b>	<b>13</b>
<b>ABSTRACT.....</b>	<b>14</b>
<b>CAPÍTULO I: INTRODUCCIÓN .....</b>	<b>15</b>
Antecedentes.....	15
Justificación .....	17
Alcance del Proyecto.....	20
Objetivos .....	21
General .....	21
Específicos .....	21
<b>CAPÍTULO II: MARCO TEÓRICO .....</b>	<b>23</b>
Televisión Digital Terrestre.....	23
TDT en Ecuador .....	25
ISDB-T .....	28
Transport Stream .....	30
Packet Elementary Stream .....	32
Tablas PSI/SI.....	33
Tablas PAT.....	35
Tablas PMT.....	36
Closed Caption .....	38
Caracteres usados en paquetes PES para Closed Caption.....	43
Código de Redundancia Cíclica .....	46



CRC 16 .....	46
Subtitulado .....	47
Tipos de Subtítulos .....	48
Reconocimiento de Voz .....	49
Transmisión .....	51
DTU-215 .....	51
<b>CAPITULO III: METODOLOGÍA.....</b>	<b>64</b>
Estructura del sistema.....	64
Reconocimiento de voz .....	66
IBM Watson Speech to Text API .....	68
Google Cloud Speech-to-Text API .....	69
Creación de paquetes.....	71
Paquetes PMT .....	72
Paquetes de Control de Subtítulos.....	74
Paquetes de Texto de Subtítulos .....	77
Multiplexación .....	84
Transmisión .....	86
Software de Control de Transmisión ESPE Player.....	86
<b>CAPITULO IV: RESULTADOS .....</b>	<b>89</b>
Transmisión con StreamXpress.....	90
Software de Control de Transmisión ESPE Player.....	95
<b>CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES.....</b>	<b>99</b>
Conclusiones.....	99
Recomendaciones.....	100
<b>BIBLIOGRAFÍA.....</b>	<b>102</b>

## ÍNDICE DE TABLAS

<b>Tabla 1.</b> Países Sudamericanos en TDT.....	24
<b>Tabla 2.</b> Indicadores TDT. ....	27
<b>Tabla 3.</b> Fases de Apagón Analógico.....	28
<b>Tabla 4.</b> Tablas PSI/SI .....	34
<b>Tabla 5.</b> Estructura de servicios en la PMT.....	37
<b>Tabla 6.</b> Parámetros del paquete PES de caption. ....	40
<b>Tabla 7.</b> Parámetros de grupos de datos. ....	41
<b>Tabla 8.</b> Parámetros de datos de gestión de subtítulos.....	42
<b>Tabla 9.</b> Set de Caracteres en Latín.....	43
<b>Tabla 10.</b> Códigos de control del set de caracteres, filas 0x y 1x. ....	44
<b>Tabla 11.</b> Códigos de control del set de caracteres, filas 8x y 9x. ....	45
<b>Tabla 12.</b> Características módulo DTU-215.....	52
<b>Tabla 13</b> Estructura de flujo de datos de paquetes PES .....	72
<b>Tabla 14</b> Estructura de descriptores.....	73
<b>Tabla 15</b> Valores iniciales de paquetes PES para subtítulos. ....	74
<b>Tabla 16</b> Valores de paquetes PES sincronizados. ....	74
<b>Tabla 17</b> Valor de parámetros de grupo de datos. ....	75
<b>Tabla 18</b> Valor de parámetros de datos de gestión.....	76
<b>Tabla 19</b> Caracteres de configuración para subtítulos en paquete PES.....	78
<b>Tabla 20</b> Caracteres de texto para subtítulos en paquete PES .....	80

## ÍNDICE DE FIGURAS

<b>Figura 1.</b>	Países que han evolucionado a la Televisión Digital Terrestre .....	24
<b>Figura 2.</b>	Estaciones que transmiten tv abierta analógica por provincia. ....	26
<b>Figura 3.</b>	Estaciones de televisión que transmiten TDT en cada provincia. ....	26
<b>Figura 4.</b>	Estructura del sistema ISDB-T.....	30
<b>Figura 5.</b>	Estructura de un Transport Stream .....	31
<b>Figura 6.</b>	Estructura de la cabecera TS. ....	32
<b>Figura 7.</b>	Estructura de un Paquete PES. ....	33
<b>Figura 8.</b>	Estructura de datos PAT .....	35
<b>Figura 9.</b>	Estructura de datos PMT.....	36
<b>Figura 10.</b>	Descriptor de identificación. ....	38
<b>Figura 11.</b>	Descriptor de componente de datos. ....	38
<b>Figura 12.</b>	Módulo DTU-215.....	53
<b>Figura 13.</b>	Interfaz de Software StreamXpress. ....	54
<b>Figura 14.</b>	Información de TS.....	55
<b>Figura 15.</b>	Información de PID. ....	55
<b>Figura 16.</b>	Información de archivo. ....	56
<b>Figura 17.</b>	Control de velocidad. ....	57
<b>Figura 18.</b>	Frecuencia de canal. ....	57
<b>Figura 19.</b>	Parámetros de Modulación. ....	58
<b>Figura 20.</b>	Parámetros ISDB-T.....	58
<b>Figura 21.</b>	Opciones de reproducción de video.....	60
<b>Figura 22.</b>	Interfaz ESPE Player.....	61
<b>Figura 23.</b>	Lectura de video TS.....	62
<b>Figura 24.</b>	Parámetros ISDB-T.....	62
<b>Figura 25.</b>	Parámetros de capa jerárquica.....	63
<b>Figura 26.</b>	Parámetros de reproducción.....	63
<b>Figura 27.</b>	Diagrama de bloques del sistema. ....	65
<b>Figura 28.</b>	Esquema de implementación del sistema. ....	66
<b>Figura 29.</b>	Diagrama de Flujos programa de reconocimiento de voz.....	67
<b>Figura 30.</b>	Credenciales de API Watson Speech to text.....	68
<b>Figura 31.</b>	Credenciales de Google Cloud.....	70
<b>Figura 32.</b>	Paquete de tabla PAT. ....	71
<b>Figura 33.</b>	Paquete de tabla PMT.....	73
<b>Figura 34.</b>	Parámetros de subtítulos en Paquete PES. ....	75
<b>Figura 35.</b>	Parámetros de grupo de datos en paquete PES. ....	76
<b>Figura 36.</b>	Parámetros de datos de gestión de subtítulos en paquete PES. ....	77
<b>Figura 37.</b>	Características del mensaje en paquete PES.....	79
<b>Figura 38.</b>	Mensaje en paquete PES.....	81
<b>Figura 39.</b>	Diagrama de flujo de programa de creación de subtítulos.....	81
<b>Figura 40.</b>	Diagrama de flujo de programa de multiplexación.....	85

<b>Figura 41.</b> Diagrama de Flujo de la función modificada. ....	87
<b>Figura 42.</b> Aplicación ESPE Player .....	88
<b>Figura 43.</b> Conversión a texto de video de prueba 1 con Google. ....	89
<b>Figura 44.</b> Conversión a texto de video de prueba 2 con IBM. ....	89
<b>Figura 45.</b> Panel de control de API de Google. ....	90
<b>Figura 46.</b> Configuración de software StreamXpress .....	91
<b>Figura 47.</b> Escenario de Transmisión de video. ....	92
<b>Figura 48.</b> Reproducción del video de prueba 1. ....	92
<b>Figura 49.</b> Video de prueba1.....	93
<b>Figura 50.</b> Análisis de paquete de subtítulos de video de prueba 1.....	93
<b>Figura 51.</b> Reproducción del video de prueba 2. ....	94
<b>Figura 52.</b> Video de prueba2.....	94
<b>Figura 53.</b> Análisis de paquete de subtítulos de video de prueba 2.....	95
<b>Figura 54.</b> Configuración de software ESPE Player. ....	96
<b>Figura 55.</b> Reproducción de video en software ESPE Player.....	97
<b>Figura 56.</b> Reproducción en tiempo real con ESPE Player.....	97
<b>Figura 57.</b> Video de reproducción en tiempo real. ....	98

## RESUMEN

En la actualidad el país se encuentra en una época de evolución hacia la TDT, cambiando la manera en la que llegará la televisión abierta a la población, el Plan Maestro de Transición a la Televisión Digital Terrestre aprobado por el MINTEL dispone las circunstancias en las que se realizará esta transición, después de un análisis técnico y socio económico del país se planifica un apagón digital gradual en los próximos años. En una transmisión de televisión los subtítulos facilitan la comprensión del mensaje no solo para personas con discapacidad auditiva sino también en circunstancias de difícil comprensión del audio. El presente proyecto se centró en la investigación, desarrollo y ejecución de un sistema de sincronización de subtítulos para transmisión en TDT, para lo cual se trabajó con el estándar ISDB-Tb y las normas ABNT NBR 15606-1 y ARIB SDT-B24 con las que se realizó la construcción de paquetes PES de subtítulos en los que se colocó los datos de control, sincronización y configuración del mensaje. El texto del mensaje se obtuvo mediante el reconocimiento de voz usando un servicio en la nube de herramientas computacionales que permiten procesar la voz humana y convertirla en información digital, después de esto se realizó la multiplexación y transmisión de los paquetes para lo que se utilizó un modulador ISDB-T con el software multiplexor que envió el video subtulado en un archivo de flujo de transporte tipo TS al decodificador para su recepción. Adicional se realizó la modificación del Software de Control de Transmisión ESPE Player para poder recibir los datos mientras están siendo generados y tener una transmisión en tiempo real.

### PALABRAS CLAVE

- TS
- API
- BTS
- MULTIPLEXACIÓN

## **ABSTRACT**

Currently the country is in an era of evolution towards Digital Television, changing the way in which open television will reach the population, the Master Plan for Transition to Digital Terrestrial Television approved by MINTEL establishes the conditions in which this transition will take place, after a technical and socio-economic analysis of the country, a gradual digital blackout is planned in the coming years. In a television broadcast, the subtitles facilitate the understanding of the message not only for people with hearing disabilities but also in circumstances where audio is difficult to understand. The present project focused on the research, development and execution of a subtitle synchronization system for broadcast digital television, for which it worked with the ISDB-Tb standard also ABNT NBR 15606-1 and ARIB SDT-B24 norms are used to construct the PES packets with the subtitle data in which the control, synchronization and configuration for message data was placed. The text of the message was obtained through voice recognition using a cloud service of computational tools that allow us to process the human voice and convert it into digital information, after which the multiplexing of the packets and the transmission for which the ISDB-T modulator with multiplexer software which sends subtitled video in a TS type transport stream file to the decoder for reception on digital television. Additionally, the modification of the ESPE Player Transmission Control Software was made in order to receive the data while it is being generated and have a real time transmission.

## **KEY WORDS**

- **TS**
- **API**
- **BTS**
- **MULTIPLEXING.**

## 1. CAPÍTULO I: INTRODUCCIÓN

### 2.1 Antecedentes

La televisión ha sido una de las mayores invenciones del siglo XX, después de varios intentos fallidos el inventor escocés John Logie Baird en 1925 usando dos discos, uno como transmisor y otro como receptor, logró transmitir una imagen, siendo este el primer paso para lo que llegó a convertirse en el medio de comunicación masivo con mayor influencia en la población mundial. En el Ecuador la televisión surgió en 1959 cuando el matrimonio Rosembaum-Zambrano importó equipos de televisión con los cuales se realizaron las primeras transmisiones y en 1960 se les otorga el primer permiso de frecuencia para el canal llamado "Primera Televisión Ecuatoriana". Desde entonces se ha ido mejorando la calidad de transmisión tanto en la señal como en la programación de cada estación de televisión; actualmente en el país operan 577 estaciones de televisión entre nacionales, regionales y locales.

En la década de los 90's empezó a surgir la idea de la televisión digital como una mejora en todos los aspectos de la transmisión, surgieron varios estándares en diferentes países siendo los principales el Europeo (DVB-T), Estadunidense (ATSC), Japonés (ISDB-T) y Chino (DTMB). Ecuador en el año 2010 acogió el estándar ISDB-T (*Integrated Services Digital Broadcasting Terrestrial*) para televisión digital, con sus adaptaciones brasileñas siendo el primer paso hacia la televisión digital. En el 2011 mediante Acuerdo Interministerial No. 170 se creó el Comité Interinstitucional Técnico para la Introducción a la Televisión Digital Terrestre en el Ecuador (CITDT) el cual en el 2012 aprobó el Plan Maestro de Transición a la TDT y se establecen fechas y fases para el apagón analógico en el país y autorizo la operación temporal de estaciones con

TDT en el 2013; ese mismo año se aprobó el reglamento en el que se establece que todos los televisores que ingresen al país deben contar con el sintonizador ISDB-T Internacional. En el 2015 se expidió la Norma Técnica de Televisión Digital Terrestre para Ecuador, la que establece las condiciones para la operación de estaciones de TDT y en el 2016 se dio inicio al Concurso Público para la Adjudicación de Frecuencias de Radiodifusión y Televisión en Señal Abierta. Actualmente aún no se cuenta con fechas para el apagón analógico pero ya se encuentran 30 estaciones de televisión brindando temporalmente servicio de TV digital.

Entre el contenido enriquecido que trae la TV digital, se encuentran los subtítulos que son una herramienta necesaria para la inclusión de la población con discapacidad auditiva o con poca audición, así como una ayuda en caso de difícil comprensión del audio por ruido externo. En la actualidad se cuenta con la tecnología de Closed Caption que realiza la transcripción del audio de un programa de televisión a texto en tiempo cuasi-real, esto implica una gran intervención humana, equipos, tiempo y costos de producción, además ya existen en el mercado transcritores “Speech to Text” con tecnología muy avanzada y eficiente, empresas como Microsoft han sacado al mercado transcritores de voz con un modelo personalizado que reconocen la voz a pesar del estilo de habla, el vocabulario y el ruido de fondo, conforme a sus necesidades y los datos disponibles. (Microsoft Azure, s.f.) IBM también cuenta con este servicio, con el software llamado Watson Speech to Text utiliza las capacidades de reconocimiento de voz para convertir el texto en árabe, inglés, español, francés, portugués brasileño, japonés, coreano, alemán y mandarín. (IBM, s.f.). También están aplicaciones como Dragon Home, Temi, Braina Pro que ofrecen varias características que ayudan a mejorar la transcripción de audio a texto, sin embargo para el uso de estas aplicaciones



se requiere de licencias pagadas esto conlleva a que se lo realice en su mayoría en programas pregrabados y en televisión pagada lo que hace que no sea accesible para los sectores vulnerables de la sociedad.

## 1.2 Justificación

La evolución hacia la TDT en el país ofrecerá muchos beneficios para la población que cuenta con televisión por señal abierta y además dará espacio para el desarrollo tecnológico en el mismo; para esta migración análogo-digital el Ecuador adoptó el estándar ISDB-T Internacional el mismo adoptado por la mayor parte de países Latinoamericanos, esto permitirá la implementación de la multiprogramación que permite el acceso a varias programaciones dentro de una misma estación, además la portabilidad y movilidad de la señal en dispositivos como teléfonos celulares o tablets, la recepción de alertas de emergencia, subtítulos y la opción del acceso a contenido interactivo. (Ecuador, s.f.)

La TDT abre un gran número de oportunidades de mejora para la televisión, no solo en el sentido de calidad sino de inclusión a sectores vulnerables que no tenían acceso a este medio de comunicación. Para las personas con discapacidad auditiva, con poca audición o adultos mayores los subtítulos son esenciales para una buena comprensión de la programación, actualmente en el Ecuador el 14.12% de la población discapacitada posee una discapacidad auditiva. La norma ISDB-T, con la cual se trabaja en el país, cuenta con el envío de texto síncrono (*Closed Caption*) el cual consiste en enviar una transcripción de lo hablado en el video como un componente de datos y transmitido en formato PES o "*Packetized Elementary Stream*"; para esto se opera con las normas ABNT NBR 15606-1 y ARIB SDT-B24.

La Ley para Estadounidenses con Discapacidades (ADA) considera a los subtítulos como un tipo de ayuda auxiliar, por ello los materiales de entretenimiento, informativos y de capacitación se subtitulan para las personas sordas en el momento en que se distribuyen y Ley de Telecomunicaciones de 1996 a obligar a la FCC (Comisión Federal de Comunicaciones) a adoptar las reglas que requieren los subtítulos cerrados en la mayoría de la programación de la televisión abierta, exigiendo además que todos los televisores incluyan un decodificador que permite a los televidentes el uso subtítulos cerrados en la programación habitual. ((NIDCD), 2017) En España la Ley 7/2010, de 31 de marzo, dispone que las personas con discapacidad auditiva tienen el derecho a que se subtitule el 75% de los programas de televisión abierta y cobertura estatal o autonómica. ("BOE", 2010) siendo estos países muestra del alcance de la tecnología de la TDT en la mejora de la comunicación.

En Ecuador la Ley Orgánica de Discapacidades en el Artículo 64 sobre la comunicación audiovisual dictamina que “La autoridad nacional encargada de las telecomunicaciones dictará las normas y regulará la implementación de herramientas humanas, técnicas y tecnológicas necesarias en los medios de comunicación audiovisual para que las personas con discapacidad auditiva ejerzan su derecho de acceso a la información. Dentro de las normas se establecerá la obligación de incorporar a un intérprete de lenguaje de señas ecuatoriana y/o la opción de subtitulado en los contenidos de programas educativos, noticias, campañas electorales y cultura general. Además, se establecerá la obligación a los medios de comunicación audiovisual y de radio para la emisión de un programa semanal en que las personas con discapacidad puedan interactuar.” (Ecuador, 2012)

Actualmente el Ministerio de Telecomunicaciones MINTEL se encuentra en el proceso de transición al servicio de radiodifusión de televisión de señal abierta denominado “Televisión Digital Terrestre” el cual trabaja en distintas líneas de acción como brindar facilidades a la población menos favorecida para la adquisición de un equipo receptor, apto para la señal de TDT y fomentar el desarrollo de contenidos digitales y plataformas tecnológicas. (MINTEL, Libro Blanco de la Sociedad de la información y del Conocimiento, 2018). Se ha presentado el Plan Maestro de transición a la Televisión Digital Terrestre el cual indica las acciones para el proceso de transición al servicio de radiodifusión de televisión digital, que garantice el acceso de la información, así como su universalización social y geográfica, implementando nuevos servicios y soluciones inherentes. Además como una de las estrategias del plan maestro se pretende “Fomentar la implementación de nuevos servicios, soluciones propias del estándar ISDB-T y sus innovaciones tecnológicas” para lo cual se analizará la factibilidad técnica, económica y social, en la implementación de los desarrollos del estándar ISDB-T Internacional. Los equipos tales como televisores y decodificadores para TDT, deberán incorporar el sistema de alerta de emergencia y propenderán a que tengan embebido el middleware para interactividad, de acuerdo con las especificaciones técnicas definidas por los organismos competentes. Esto incrementará el acceso a la información, como principal medio masivo, contribuyendo así a la consolidación de la Sociedad de la Información. (MINTEL, PLAN MAESTRO DE TRANSICIÓN A LA TELEVISIÓN DIGITAL TERRESTRE, 2018)

En el cronograma de evaluación del cese de señales de tv analógicas en el Ecuador, las fechas y localidades determinadas para el apagón surgen de un análisis técnico y socio-económico de la realidad del país, experiencias internacionales, y el

tiempo promedio para la migración de las estaciones por parte de los operadores, con lo que se ha decidido que el inicio de la Fase 1 será el 17 de mayo 2020 para el área de operación zonal de las estaciones cuya cobertura incluya a la ciudad de Quito y terminara con la Fase 4 el 01 de diciembre de 2023 con las áreas que tengan una población menor a 200.000 habitantes

El país se encuentra en una época de transición que brinda la posibilidad de lograr una difusión inclusiva en medios de comunicación como lo son la televisión abierta, proporcionando las herramientas necesarias para una comprensión total de personas con discapacidades, es por ello que una tecnología eficiente y accesible que genere subtítulos cerrados en la programación habitual es necesaria para dar un paso más hacia la inclusión social y el avance tecnológico.

### **1.3 Alcance del Proyecto**

Este proyecto se divide en cuatro etapas, en la primera se realiza una investigación sobre las normas y estándares (ABNT NBR 15606-1 y ARIB SDT-B24) previamente mencionadas para la implementación y transmisión, poniéndolas en práctica para verificar la configuración que se usará; además se analizan los trabajos ya existentes en este tema.

En la segunda etapa se realiza el análisis de videos que contengan *Closed Caption*, con los cuales se hace una verificación, manipulación y recomposición de la información existente en los paquetes de datos, se revisan la tablas de flujo de transporte y los descriptores de cada una de ellas, para posteriormente realizar la multiplexación y transmisión. Se realizaron varias pruebas en las cuales se varía la configuración y características del texto para obtener el más adecuado al proyecto.

En la tercera etapa se realiza el estudio de las técnicas de codificación existentes para la transcripción de audio a texto, poniendo atención en la calidad, eficiencia y accesibilidad, para esto se realizan varias pruebas con diferente codificación para obtener el mejor método que se adapte a las necesidades del proyecto.

Finalmente, se realiza un programa general que utiliza el método de codificación elegido para la transcripción de audio a texto en tiempo cuasi-real y lo transmite en un paquete PES de datos. Estos paquetes serán multiplexados con el audio y video en formato TS para su transmisión sobre TDT.

## **1.4 Objetivos**

### **1.4.1 General**

Desarrollar de un sistema de transcripción de audio a texto para generación y sincronización de subtítulos en transmisión de TV digital

### **1.4.2 Específicos**

- Investigar las tecnologías actuales en torno a la TDT para sincronización de subtítulos en tiempo cuasi-real.
- Realizar una transmisión y recepción usando la tecnología actual de Closed Caption.
- Analizar los sistemas de codificación, evaluar su eficiencia y accesibilidad al momento de su implementación para elegir el más adecuado para la transcripción de audio a texto.

- Generar los paquetes de datos que contendrán la información de los subtítulos y construir el flujo de salida que contenga el audio, video y datos (subtítulos) sincronizados.

## 2. CAPÍTULO II: MARCO TEÓRICO

### 2.1 Televisión Digital Terrestre

La TDT es una mejora tecnológica de la televisión analógica que no solo mejora la calidad de imagen y sonido sino que abre un abanico de posibilidades. La televisión es el medio de comunicación masiva consumido por la mayor parte de la población mundial, es un medio que se encuentra en la vida cotidiana del espectador, la TDT además permitirá la interacción del espectador con la información transmitida uniendo la tecnología de la comunicación con la sociedad y potencializando los desarrollos de la primera para las necesidades de la segunda. (Ramírez, 2008)

La TDT es un sistema de transmisión que consiste en la codificación de video, audio y datos en un flujo de transporte binario, esto hace que se incremente en gran medida el espacio disponible para transmisión y la televisión tenga una oferta más amplia de servicios. Se han desarrollado estándares la transmisión de esta tecnología, los usados actualmente a nivel mundial son, ATSC el estándar Americano, DVB el estándar Europeo, ISDB-T el estándar Japonés y DTMB el estándar Chino. En la figura 1 se puede observar como a nivel mundial esta tecnología ya ha sido acogida por un gran número de países.

**Figura 1.**

*Países que han evolucionado a la Televisión Digital Terrestre*



*Nota:* Información obtenida de MINTEL

En Sudamérica la mayoría de los países han optado por el estándar japonés ISDB-T y han culminado o están en proceso de migración a esta tecnología como se observa en la Tabla 1. (Abel Suing, 2015)

**Tabla 1.**

*Países Sudamericanos en TDT.*

País	Estándar	Migración TDT
Ecuador	ISDB-Tb	Finalizara en 2023
Colombia	DvB- T	Finalizara en 2021
Perú	ISDB-T	Finalizara en 2024
Brasil	ISDB-Tb	Finalizado
Argentina	ISDB-Tb	Finalizara en 2021
Bolivia	ISDB-T	Finalizara en 2024
Chile	ISDBT	Finalizara en 2024

*Nota:* Información obtenida de Mintel



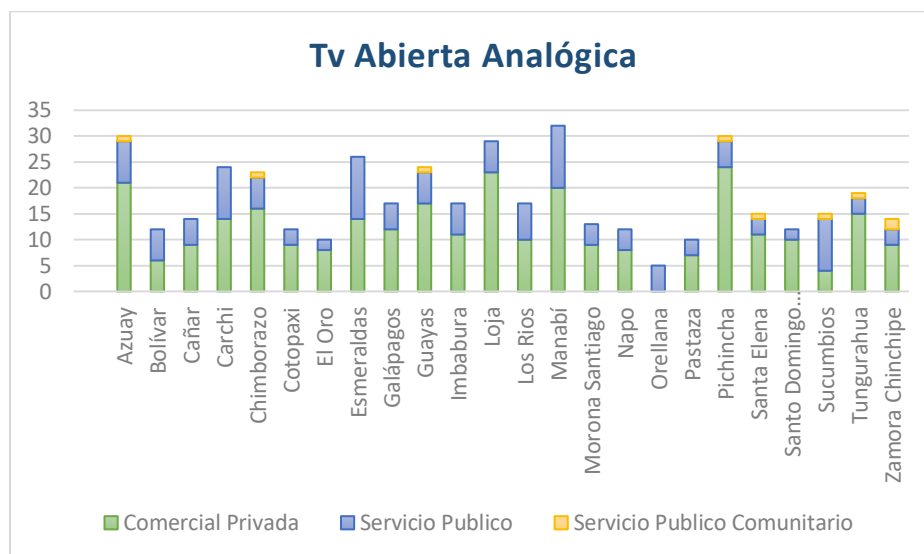
### **2.1.1 TDT en Ecuador**

En Ecuador se adopta el estándar ISDB-T Internacional el 25 de marzo de 2010, comenzando así el proceso de migración hacia la televisión digital; se delegó esta tarea al Ministerio de Telecomunicaciones y de la Sociedad de la Información (MINTEL) el cual el 9 de julio de 2018 presenta el Libro Blanco de la Sociedad de la Información y del Conocimiento (LBSIC) que cuenta con el Plan Maestro de Transición a la Televisión Digital Terrestre que establece que se incrementara la cobertura digital terrestre a nivel nacional mediante campañas de información y reparación para la transición desde la señal analógica, se facilitara la adquisición de equipos aptos para la recepción de señal de TDT para la población menos favorecida y se incentivara la investigación y desarrollo de contenido digital, interactivo e inclusivo. (MINTEL, PLAN MAESTRO DE TRANSICIÓN A LA TELEVISIÓN DIITAL TERRESTRE, 2018)

Actualmente a nivel nacional existen 5361 estaciones de televisión abierta que transmiten en formato analógico, de las cuales se les ha concedido un permiso temporal a 31 estaciones para que transmitan en formato digital abarcando un 54% de la cobertura del país, pero hay que tomar en cuenta que para recibir la señal digital el televisor debe contar con un sintonizador que cumpla el estándar ISDB-T o un decodificador adecuado que recepte la señal. En las Figuras 2 y 3 se pueden observar el número de estaciones por provincia que transmiten en Tv Abierta Analógica y TDT respectivamente.

**Figura 2.**

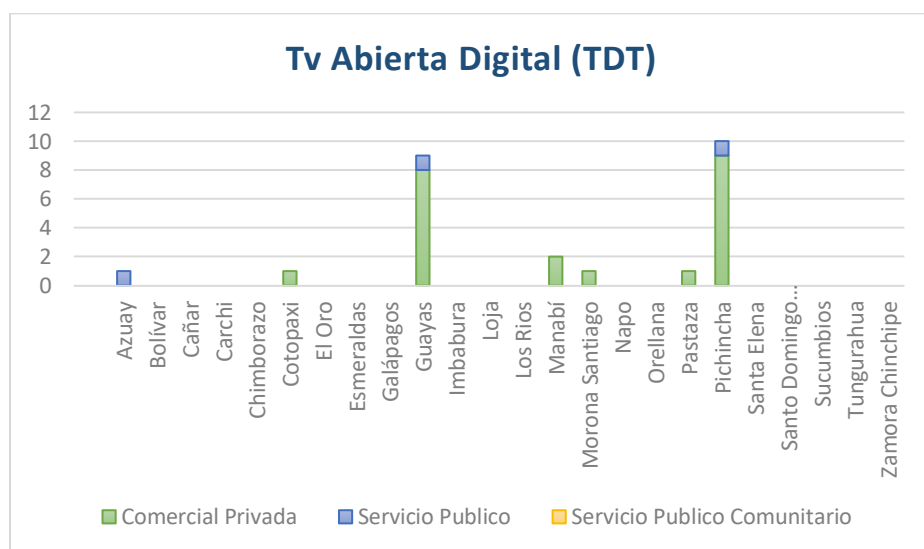
*Estaciones que transmiten tv abierta analógica por provincia.*



*Nota: Información obtenida de ARCOTEL*

**Figura 3.**

*Estaciones de televisión que transmiten TDT en cada provincia.*



*Nota: Información obtenida de ARCOTEL*

Para avanzar en la migración se debe incrementar en cobertura, y estaciones que brinden el servicio de TDT, además de informar sobre el proceso a realizarse a la mayor cantidad de población para que puedan estar preparados adecuadamente, a junio de 2017 únicamente el 16% de los hogares que ven televisión contaban con los equipos necesarios para recibir la señal digital, entre los lineamientos del plan maestro se plantean metas para mejorar estos porcentajes. Tabla 2. (MINTEL, Libro Blanco de la Sociedad de la información y del Conocimiento, 2018)

**Tabla 2.**

*Indicadores TDT.*

Indicador	Línea Base 2017	Meta 2021
Cobertura poblacional con señal TDT	54%	66%
Estaciones que brindan el servicio de TDT	5.7%	15%
Hogares que conocen el proceso de implementación de la TDT	26%	50%

*Nota:* Información obtenida del Libro Blanco de la Sociedad de la Información y del Conocimiento.

En el Plan Maestro de Transición a la TDT se han plantado tres estrategias necesarias para la migración.

1. Fortalecer el entorno regulatorio y mejorar las condiciones que favorezcan el despliegue de infraestructura necesaria para brindar un servicio de calidad.
2. Fomentar mecanismos que permitan incrementar el acceso a la televisión digital por parte de los hogares que se encuentran preparados e informar todos los beneficios del servicio.

3. Fomentar la implementación de nuevos servicios, soluciones propias del estándar y sus innovaciones tecnológicas. Recalcando la necesidad de la investigación y desarrollo de nuevas tecnologías, entre las cuales se encuentran el sistema de alertas de emergencia EWBS, aplicaciones interactivas en aplicativos GINGA y la posibilidad de brindar soluciones tecnológicas a problemas sociales que afectan al sector más vulnerable de la población.

Todo esto se pondrá en marcha mediante 4 fases previamente seleccionadas, las localidades y fechas de cada una han sido elegidas mediante un estudio técnico y socioeconómico, tomando en cuenta también las experiencias internacionales. Tabla 3.

**Tabla 3.**

*Fases de Apagón Analógico.*

<b>Fases</b>	<b>Localidad</b>	<b>Fecha de Apagón</b>
Fase 1	Quito y sus alrededores	17 de mayo 2020
Fase 2	Guayaquil y sus alrededores	09 de julio 2020
Fase 3	Poblados que tengan entre 1000000 y 200000 habitantes	03 de junio 2022
Fase 4	Poblados con menos de 200000 habitantes	01 de diciembre 2023

*Nota:* Información obtenida de MINTEL

## **2.2 ISDB-T**

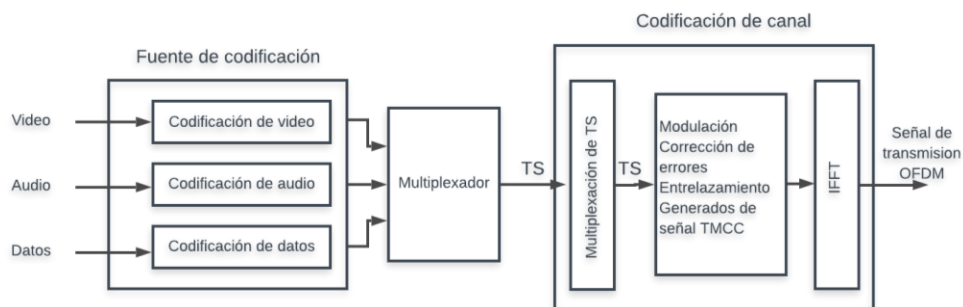
El estándar ISDB-T es un estándar internacional desarrollado en Japón a finales del siglo XX y que ha sido acogido en la mayoría de los países de Latinoamérica, es un sistema de emisión de audio, video y datos en alta calidad, que ofrece flexibilidad y capacidad de expansión. Está diseñado para trabajar con un ancho de banda de 6MHz

por canal y con una modulación OFDM (*Orthogonal Frequency Division Multiplexing*) segmentada BTS (*Band Segmented Transmission*), lo que permite que el ancho de banda se reparta en diferentes servicios teniendo una transmisión jerárquica hasta de 3 capas (A, B y C) y distintos parámetros de modulación dentro de la misma banda. El ISDB-T se desarrolló con ciertos requerimientos que aseguraran su calidad y eficiencia:

- Alta Calidad
- Robustez
- Movilidad y portabilidad
- Compatibilidad
- Uso para prevención de desastres/ EWBS
- One seg
- Mejor uso de recursos de frecuencias.

El ISDB-T ofrece multiprogramación debido a que cada canal se divide en 14 segmentos, pero uno no es usado y se deja como banda de guarda en los extremos del canal evitando así el entrelazado de frecuencia, los otros 13 segmentos se dividen entre las 3 capas para servicios de HDTV (*High Definition Television*) o televisión de alta definición, SDTV (*Standard Definition Television*) o televisión de definición estándar y one seg, de esta manera se pueden realizar transmisiones a dispositivos móviles y fijos con distinta definición simultáneamente. (ARIB, Reporte Técnico de ARIB.

Características del sistema ISDB-T)

**Figura 4.***Estructura del sistema ISDB-T*

La estructura de transmisión del sistema se muestra la en la Figura 4, se compone del bloque de codificación de fuente, el bloque de multiplexación y el bloque de codificación de canal para la codificación del video se utiliza MPEG-2 o MPEG-4 AVC/H.264, para la de audio MPEG-2 AAC y en el multiplexor MPEG-2, del multiplexador sale un solo flujo TS (*Transport Stream*) que ingresa para ser codificado nuevamente antes de su transmisión, ahí se añade una señal de control TMCC (transmission multiplexing configuration) que indica al receptor los parámetros de la transmisión; finalmente se genera un flujo llamado BTS (Flujo de Transporte Broadcasting) que contiene la información binaria total de la transmisión.

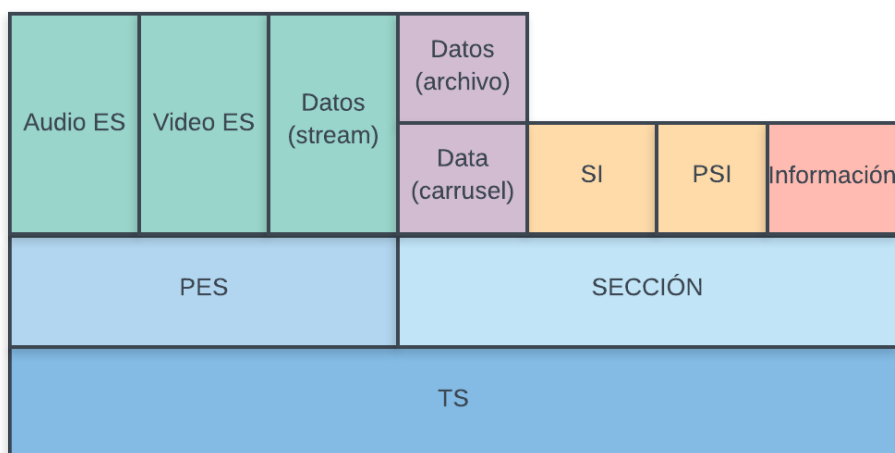
### 2.3 Transport Stream

En el sistema MPEG-2 el contenido de audio, video y datos son multiplexados en un flujo de datos denominado TS que usa paquetes pequeños y de tamaño fijo, compuesto por 188 bytes 4 de cabecera y 184 de carga útil. Para obtener un paquete TS primero el audio y video deben pasar por un proceso de codificación independiente para formar un flujo elemental denominado ES (*Elementary Stream*) que es la unidad básica de transporte de paquetes, después cada ES es paquetizado en formatos PES

(*Packet Elementary Stream*) para finalmente formar el TS, los TS de cada fuente deben ser re-multiplexados para un formar un único flujo de transporte, el cual al ser sometido a la codificación de Reed-Solomon pasa a formar una señal BTS (*Broadcast Transport Stream*) que aumenta 16 bytes al paquete y mantiene un tamaño fijo de 204 bytes.

**Figura 5.**

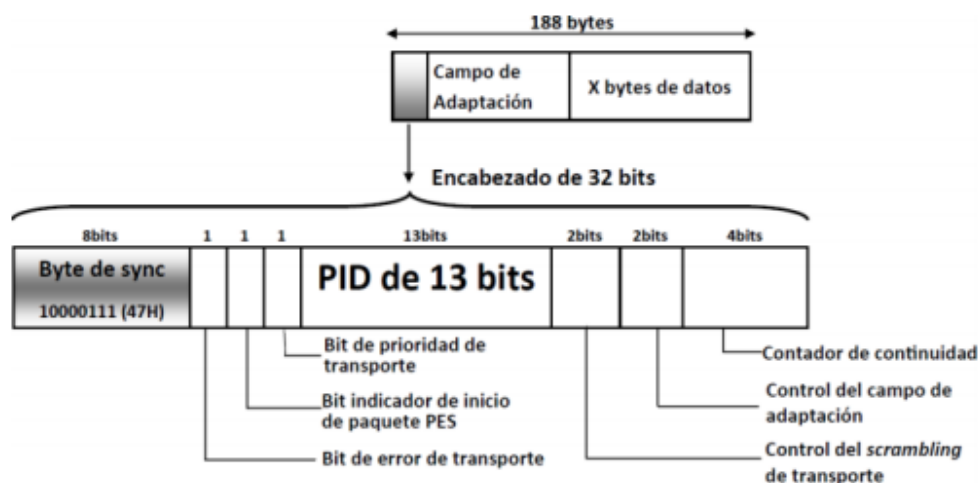
Estructura de un Transport Stream



La cabecera TS inicia con un byte de sincronización que siempre tienen el valor de 0x47 seguido de 3 bits indicadores, después se tienen 13 bits del PID (Packet Identifier) que es el identificador de paquete e indica el tipo de información que se encuentra en la carga útil, a continuación se tienen cuatro bits que indican el modo de codificación y por último, 8 bits del contador de continuidad que aumentan con cada paquete del mismo PID siempre y cuando exista información en el bloque de datos.

**Figura 6.**

Estructura de la cabecera TS.



### 2.3.1 Packet Elementary Stream

Un PES es un paquete de longitud variable con la información del elementary stream y un valor de máximo 65.536 bytes de los cuales 6 bytes son de cabecera, los primeros 3 bytes de la cabecera forman el *start code prefix*, formado por el código 00 00 01 que indica el inicio del paquete, después tenemos el byte del *stream ID*, el que indica el tipo de ES que se encuentra en el payload.

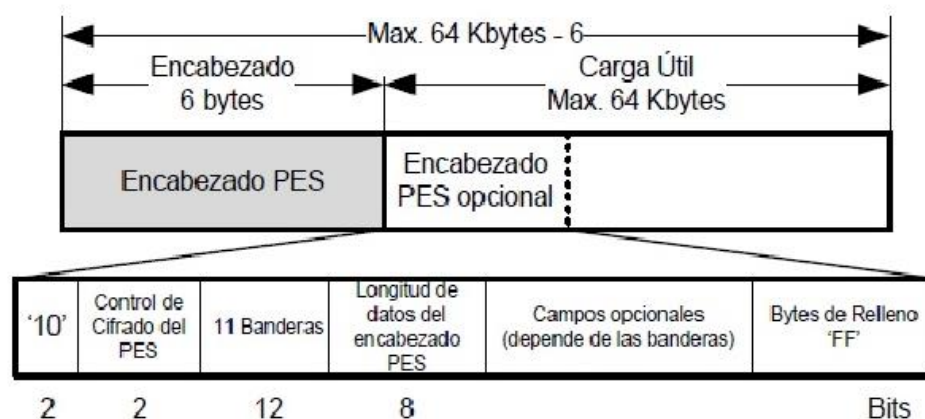
Después de la cabecera existe una “cabecera PES opcional” para adaptar los requisitos de transmisión del ES; tiene longitud variable e inicia con los bits “10” seguido de dos bits de control de cifrado, después se tiene 12 bits que contienen a las 11 banderas que indican los componentes que se encuentran presentes en los campos de la cabecera opcional PES, y contienen el PTS (Presentation Time Stamps) y el DTS (Decoding time stamps) que se encargan de la sincronización del audio y video.



Finalmente seguida de esta cabecera se encuentra la carga del ES o payload, esto se observa en la Figura 7.

**Figura 7.**

*Estructura de un Paquete PES.*



Equivalente a los paquetes PES están las Secciones que son paquetizadas y transportadas en los mismos TS, estas secciones comprenden las tablas PSI/SI y el carrusel de datos en caso de ser necesario.

### 2.3.2 Tablas PSI/SI

Las Tablas PSI/SI contienen la información necesaria para identificar los datos de entrada en el decodificador y de esta manera localizar y demultiplexar el video, audio e información adicional enviada.

Las tablas PSI (*Program Specific Information*) son parte del estándar MPEG-2, su función es especificar a qué programa pertenece la información y de esta manera demultiplexarla en el decodificador. Las tablas SI (*Service Information*) son en

complemento de las tablas PSI permitiendo el control de la recepción de datos de aplicaciones interactivas, EPG y otras, según la norma ABNT NBR 15603-1 existen 15 tablas de este tipo. En la Tabla 4 se encuentran explicadas las tablas PSI/SI más importantes.

#### Tabla 4.

##### *Tablas PSI/SI*

Tabla	Nombre	Función
PSI	PAT (Tabla de asociación de programas)	Indica los valores de PID de cada servicio de audio video o datos de los TS, crea un enlace entre los campos de "program number", "transport stream id" y "program map id".
	PMT (Tabla de mapeo de programas)	Identifica los PIDs de cada servicio que constituye el programa y la localización de la del PCR.
	CAT (Tabla de acceso condicional)	Proporciona información de los sistemas de acceso condicional y asocia las EMM transmitidas.
SI	NIT(Tabla de información de red)	Brinda la información y características de la organización física de los transport stream (TS) de una red, además de los datos necesarios para la sintonización de un servicio.
	SDT (Tabla de descripción de servicios)	Describe los servicios existentes en un transport stream (TS)
	AIT (Tabla de información de aplicaciones)	Transmite información de control de aplicaciones.
	EIT (Tabla de información de eventos)	Brinda información de los eventos en cada servicio, existentes en orden cronológico.

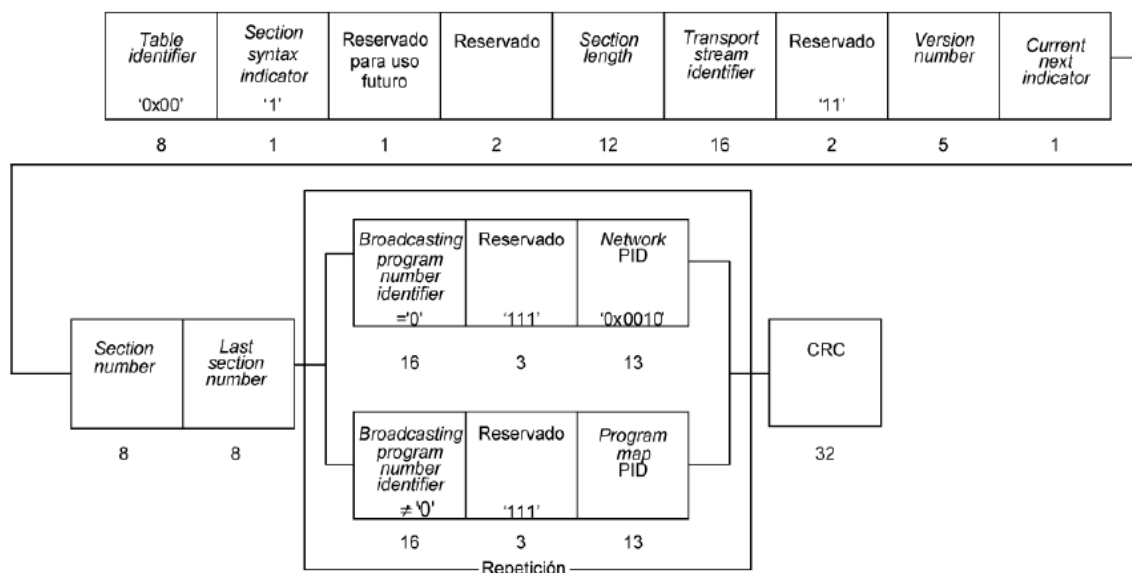
*Nota:* Información obtenida de ABNT NBR 15603-1

### 2.3.3 Tablas PAT

Las tablas PAT del inglés *Program Association Table* son indispensables en la multiplexación ya que indica el valor de los PIDs de los diferentes transport stream (TS) y deben tener un flujo constante de frecuencia de al menos 100ms. Su valor de PID es siempre 0x0000 y su "table\_ID" 0x00. Su estructura se muestra en la Figura 8.

**Figura 8.**

*Estructura de datos PAT*



*Nota:* Información obtenida en ABNT NBR 15603-1

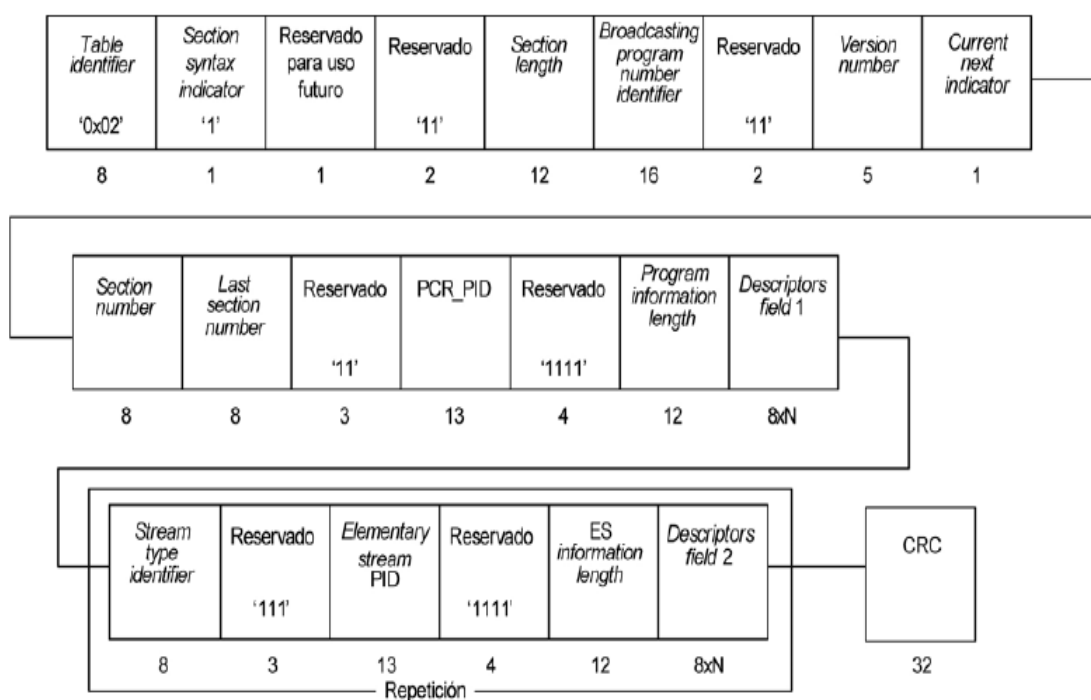
Adicional en su estructura se encuentra el valor de PID de la tabla PMT y de la tabla NIT; si el valor del *Broadcasting program number identifier* es igual a 0 entonces el PID pertenece a la tabla NIT, en cambio si el valor es cualquier número diferente de 0 el PID pertenece a la tabla PMT; pueden existir varias tablas PMT ya que pueden existir varios servicios en un mismo video.

### 2.3.4 Tablas PMT

Las tablas PMT vienen del inglés Program Map Table, estas tablas son obligatorias y su objetivo es localizar las transmisiones y el PCR (Program Clock Reference) que componen cada servicio.

**Figura 9.**

*Estructura de datos PMT.*



*Nota:* Información obtenida en ABNT NBR 15603-1

Como podemos observar en la Figura 9 en la estructura básica de un PMT tenemos que su "Table\_id" tiene el valor de 0x02 y se encuentra ubicado en la sexta posición del paquete, después de la cabecera y el contador. Para poder colocar un servicio adicional, como en este caso los subtítulos, se varía la parte final de la estructura de datos según la norma ABNT NBR 15603-2. Según el tipo de flujo, el

campo de “stream type” tiene un valor de 0x11 cuando se trata de un flujo de audio, un valor de 0x1B cuando se trata de video y un valor de 0x06 cuando se trata de un paquete PES de subtítulo o superposición, en la tabla 5 se detalla la estructura de datos para el servicio de subtítulos en una tabla PMT.

**Tabla 5.**

*Estructura de servicios en la PMT*

Campo	Valor	Bits
Stream type identifier	0x06	8
Reservado	‘111’	3
Elementary stream PID	(PID del paquete PES donde se encuentran los subtítulos)	13
Reservado	‘1111’	4
ES information length	(Longitud total de los descriptores )	12
Descriptors file 2		8xN

*Nota:* Información obtenida en ABNT NBR 15603-2

El descriptor file que usaremos en este documento será el descriptor de identificación cuyo valor de descriptor field es 0x52 y el valor del “component tag” o etiqueta de componente del título ES es 0x30 en caso de transmisión a través de capas distintas a la capa de recepción parcial y en caso de transmisión a través de la capa de recepción parcial es 0x87.

**Figura 10.***Descriptor de identificación.*

<i>Descriptor tag</i>	<i>Descriptor length</i>	<i>Component tag</i>
'0x52'		
8	8	8

*Nota:* Información obtenida de ABNT NBR 15603-1

Además se usara el descriptor de componente de datos que identifica el formato de la señal de datos y valor de su descriptor field es 0xFD y su estructura en la Figura 11. (ABNT, 2007)

**Figura 11.***Descriptor de componente de datos.*

<i>Descriptor tag</i>	<i>Descriptor length</i>	<i>Data coding method id</i>	<i>Additional identifier info</i>
'0xFD'			
8	8	16	8 x N

*Nota:* Información obtenida de ABNT NBR 15603-1

### 2.3.5 Closed Caption

Los datos de subtítulos o superposición se transmiten por un ES separado y por el mismo PMT simultáneamente con el programa principal sin distribuirse los datos antes de que inicie el programa. El texto de los subtítulos no se puede mostrar hasta que se reciban los datos de control de subtítulos, por lo tanto estos datos se envían con

una frecuencia máxima de 1 vez cada 0.3 segundos y una frecuencia mínima de una vez cada 5.0 segundos; esto puede ser interrumpido por comerciales, anuncios, etc.

Una transmisión puede ser asíncrona para EWBS o síncrona para transmisión de datos que deben estar sincronizados con las transmisiones de audio y video como los subtítulos. Los paquetes PES sincronizados deben cumplir con la sintaxis y formato especificados en la ISO/IEC 13818-1

Se aplica un método de transmisión PES síncrono con los siguientes parámetros.

- Se puede enviar solo un ES a la misma capa simultáneamente
- Cada ES acepta dos idiomas.
- Un PES debe pesar 32kB máximo.
- Un paquete PES se debe enviar en un intervalo mínimo de 100 ms.
- La tasa máxima de cada ES es de 256 kbit/s.
- El Búfer de recepción debe ser mayor o igual a 64 kB.

**Tabla 6.***Parámetros del paquete PES de caption.*

<b>Campo</b>	<b>Valor</b>	<b>Bits</b>
Prefijo de inicio	0x000001	24
Stream_ID	0xBD	8
PES_packet_length	Longitud del paquete PES, desde el siguiente byte	16
Data_identifier	0x80	8
PTS_DTS_flags	'10' PTS presente	2
ESCR_flag	'0'	1
ES_rate_flag	'0'	1
DSM_trick_mode_flag	'0'	1
Additional_copy_info_flag	'0'	1
PES_CRC_flag	'0'	1
PES_extension_flag	'1'	1
PES_header_data_length	Longitud de la cabecera de datos	8
Reserverd	'0010'	4
PTS[32..30]		3
Marker_bit	'1'	1
PTS[29..15]		15
Marker_bit	'1'	1
PTS[14..0]		15
Marker_bit	'1'	1
PES_private_data_flag	'1'	1
Pack_header_field_flag	'0'	1
Program_packet_sequence_counter_flag	'0'	1
P-SDT_buffer_flag	'0'	1
Reserved	'111'	3
PES_extension_flag_2	'0'	1
User area	Área no usada se coloca el bit '1'	8xN
Private_stream_id	0xFF	8
Reservado	0xF	4
PES data packet header length	0x0	4
PES_data_Private_data_byte	--	8xN

*Nota:* Información obtenida de ABNT NBR 15603-2



Como se observa en la Tabla 6 el valor de *stream\_id* en el caso de un flujo sincronizado de datos, se establece en '0xBD', el campo de *PTS\_DTS\_flags* enciende la bandera para la sincronización de datos, PTS (Presentation Time Stamp) es la información de gestión del tiempo utilizada para la visualización de datos de subtítulos. Los valores de *PES\_data\_Private\_data\_byte* se muestran en la Tabla 7.

**Tabla 7.**

*Parámetros de grupos de datos.*

<b>Campo</b>	<b>Valor</b>	<b>Bits</b>
Data_group_id	0x00	8
Data_group_version	No opera	8
Data_group_link_number	0x00	8
Last_data_group_link_number	0x00	8
Data_group_size	Longitud total a partir del siguiente byte hasta el crc16	8
Data_group_data_byte	Almacenamiento del grupo de datos.	8
CRC16	Código de redundancia cíclica.	16

*Nota:* Información obtenida de ABNT NBR 15603-2

Adicional los parámetros que pueden ser especificados en los datos de gestión de subtítulos, se muestran en la tabla 8. Los datos de gestión de subtítulos deben ser enviados máximo en un intervalo de 3 minutos, en caso de que el intervalo sea mayor la operación de inicialización de la unidad receptora se realiza al momento de seleccionar las estaciones, estos datos indican el lenguaje, formato, tipo de visualización entre otros.

**Tabla 8.***Parámetros de datos de gestión de subtítulos.*

Campo	Valor	Bits
TDM	'00' (libre) '01' (tiempo real)	4
Num_languages	1 o 2	4
Language_tag	0-1	4
DMF	'0010' (visualización automática cuando se recibe)  '1010' (elección de pantalla cuando se recibe)	4
ISO_639_language_code	Código del lenguaje a ser usado.	24
Formato	'1000'(escritura horizontal de 960x540) '1001'(escritura vertical de 960x540) '1010'(escritura horizontal de 720x480) '1011'(escritura vertical de 720x480)	4
TCS	'00' (código de caracteres)	8
Rollup_mode	'00' sin rollup '01' rollup '10' reservado '11' reservado	4
Data_unit_loop_length	Longitud del paquete desde el siguiente byte.	8
Data_unit	Almacenamiento de unidad de datos	8

*Nota:* Información obtenida de ABNT NBR 15603-2

### 2.3.6 Caracteres usados en paquetes PES para Closed Caption.

Los caracteres usados para Closed Caption y sobreimposición son códigos que constan de 8 bits, para países de habla hispana se usa el set de caracteres en latín que se observa en la tabla 9.

**Tabla 9.**

*Set de Caracteres en Latín.*

	0x	1x	2x	3x	4x	5x	6x	7x	8x	9x	Ax	Bx	Cx	Dx	Ex	Fx
0x	NU L		SP	0	@	P	`	p	BKF	COL	10/0	°	À	Ð	à	ö
1x			!	1	A	Q	a	q	RDF	FLC	ı	±	Á	Ñ	á	ñ
2x			"	2	B	R	b	r	GRF	CDC	é	²	Â	Ò	â	ò
3x			#	3	C	S	c	s	YLF	POL	£	³	Ã	Ó	ã	ó
4x			\$	4	D	T	d	t	BLF	WMM	€	Ž	Ä	Ô	ä	ô
5x			%	5	E	U	e	u	MGF	MACR O	¥	μ	Å	Ö	å	ö
6x		PAP F	&	6	F	V	f	v	CNF		Š	ŕ	Æ	Ö	æ	ö
7x	BEL		'	7	G	W	g	w	WH F	HLC	§	·	Ç	×	ç	÷
8x	APB	CAN	(	8	H	X	h	x	SSZ	RPC	š	ž	È	Ø	è	ø
9x	APF	SS2	)	9	I	Y	i	y	MSZ	SPL	©	¹	É	Ù	é	ù
Ax	APD		*	:	J	Z	j	z	NSZ	STL	ª	º	Ê	Ú	ê	ú
Bx	APU	ESC	+	;	K	[	k	{	SZX	CSI	«	»	Ë	Û	ë	û
Cx	CS	APS	,	<	L	\	l				¬	Œ	Ï	Ü	ï	ü
Dx	APR	SS3	-	=	M	]	m	}		TIME	ÿ	œ	Í	Ý	í	ý
Ex	LS1	RS	.	>	N	^	n	~			®	Ÿ	Î	Þ	î	þ
Fx	LS0	US	/	?	O		o	DE L			-	¿	İ	ß	ı	15/1 5

*Nota:* Información obtenida de ABNT NBR 15603-2

Con estos códigos se ingresa el mensaje a transmitir en los paquetes PES, además los caracteres usados para color, tamaño, fondo, separación, posición y demás funciones de control de subtítulos, se los encuentra en las columnas 0x, 1x, 8x y 9x.

**Tabla 10.**

*Códigos de control del set de caracteres, filas 0x y 1x.*

Código	Función
NUL	Espacio
BEL	BEL
APB	Posición activa hacia atrás
APF	Posición activa hacia adelante
APD	Avance de línea de operación
APU	Retiro de línea de operación
CS	Limpiar pantalla
APR	Retorno de posición activa
LS1	Bloqueo de turno 1
LS0	Bloqueo de turno 0
PAPF	Operación especificada avance de posición
CAN	Cancelar
SS2	Turno simple 2
ESC	Salir
APS	Especificación de posición activa
SS3	Turno simple 3
RS	Código de identificación del encabezado de datos
US	Código de identificación de la unidad de datos
SP	Espacio

*Nota:* Información obtenida de ABNT NBR 15603-2

En las tablas 10 y 11 se encuentra especificado la función de cada código para el control y configuración de los paquetes PES de clased caption.

**Tabla 11.**

*Códigos de control del set de caracteres, filas 8x y 9x.*

Código	Función
BKF	Color de primer plano negro
RDF	Color de primer plano rojo
GRF	Color de primer plano verde
YLF	Color de primer plano amarillo
BLF	Color de primer plano azul
MGF	Color de primer plano magenta
CNF	Color de primer plano cian
WHF	Color de primer plano blanco
SSZ	Talla pequeña
MSZ	Talla mediana
NSZ	Talla normal
SZX	Especificaciones de talla
COL	Especificaciones de color
FLC	Control intermitente
CDC	Ocultar control
POL	Patrón de polaridad
WMM	Modificación del modo de escritura
MACRO	Especificación macro
HLC	Control de cierre
RPC	Repetición de personaje
SPL	Dejar de subrayar y separar mosaicos
STL	Comience subrayado y separación de mosaico
CSI	Introducción de secuencia de control
TIME	Control de tiempo.

*Nota:* Información obtenida de ABNT NBR 15603-2

## 2.4 Código de Redundancia Cíclica

El código de redundancia cíclica es un método seguro y confiable para detectar errores en datos transmitidos o almacenados. El algoritmo de CRC trata los datos como un número binario el cual se divide por otro número binario llamado polinomio generador, el resto de la división es la suma de verificación CRC, que se agrega al final del mensaje transmitido. El receptor divide el mensaje (incluido el CRC), por el mismo polinomio generador que utilizó el transmisor, si el resultado de esta división es cero, entonces la transmisión fue exitosa caso contrario ocurrió un error durante la transmisión. Este tipo de código recibe como un número variable de datos y tiene una salida fija

El cálculo de CRC se describe como cálculos polinómicos ya que el valor del polinomio generador suele escribirse como un polinomio de cierto orden. Existen varios formatos para la implementación de CRC, como el CRC16, CRC-CCITT, CRC-DNP, CRC-32 u otros polinomios generadores.

### 2.4.1 CRC 16

Para hacer el cálculo del código de redundancia cíclica en el formato de CRC-CCITT en orden 16 para el cual el polinomio tiene el valor de:

$$x^{16} + x^{12} + x^5 + 1$$

Este polinomio también puede ser representado como 0x1021 al cual se lo traduce a binario para realizar la división aritmética de modulo-2 en la cual el cálculo se realiza mediante un XOR entre el polinomio generador y el mensaje o datos a transmitir. Para obtener nuestro CRC vamos a realizar los siguientes pasos:

- 1) Aumentar al mensaje tantos ceros como el orden del polinomio generador (en este caso 16).
- 2) Realizar un XOR entre el polinomio generador y los primeros 17 bits (el mismo número de bits del polinomio) del mensaje.
- 3) De la respuesta que se obtiene del XOR vamos a eliminar los ceros a la izquierda del primer 1 y colocamos a la derecha los siguientes bits del mensaje hasta completar 17.
- 4) Volvemos a realizar el paso 2 y 3 terminar con los bits del mensaje.
- 5) La respuesta del último XOR será nuestro CRC.

## **2.5 Subtitulado**

Los subtítulos son textos que aparecen superpuestos en la pantalla los cuales son la transcripción de todo o una parte del audio que se está transmitiendo, esto es necesario para la comprensión del programa por parte de las personas con discapacidad auditiva, adultos mayores, personas con poca audición o en casos de ambientes ruidosos que dificulten la audición.

Los subtítulos se encuentran normalmente en la parte inferior de la pantalla y son transmitidos en tiempo real o cuasi-real. En el país actualmente no se cuenta con este servicio en televisión nacional, los canales de televisión tanto público como privados solo cuenta con la subtitularían cuando se trata de algún aviso de emergencia o cadenas nacionales; haciendo que la programación normal sea inaccesible para ciertos sectores vulnerables de la población.

### 2.5.1 Tipos de Subtítulos

Los subtítulos se clasifican en diferentes tipos dependiendo de diferentes criterios. Se los clasificara según: su idioma, su tipo, su manera de transmisión. (Maseo, 2015)

Según su idioma:

- **Interlingüísticos**, son aquellos en los que el texto se encuentra en un idioma diferente del audio.
- **Intralingüísticos**, son aquellos en los que el texto y el audio se encuentran en el mismo idioma.

Según su forma de transmisión:

- **Diferido**, se insertan en el video previo a la transmisión, se usan en programas pregrabados
- **Directos**, se insertan en el trascurso del programa, son usados para programas en directo y se llevan a cabo mediante técnicas de estenotipia o reconocimiento de voz.

Según la forma de aparición en pantalla:

- **Scrolling**, o acumulativo estos subtítulos aparecen en la pantalla línea por línea y se desplazan hacia arriba y el nuevo texto va apareciendo en la línea inferior.
- **Bloque**, aparecen y desaparecen en la pantalla desplegados en una caja cuadrada o bloque y cada uno de ellos puede constar de dos o tres líneas.

Según su audiencia:



- **Abiertos**, están siempre en pantalla ya que son grabados junto con el video
- **Cerrados**, el usuario puede escoger si desea o no consumirlos ya que están grabados a parte del video.

## 2.5.2 Reconocimiento de Voz

El reconocimiento de voz es una herramienta computacional que permite procesar la voz humana y convertirla en información digital, la señal emitida por la voz comprende un conjunto de estudios de fonética, acústica, semántica, léxico; además de las ambigüedades que se debe tomar en cuenta para llegar a la correcta interpretación del mensaje.

Existen varias técnicas de aprendizaje para lograr un óptimo reconocimiento de la señal de voz; el Aprendizaje deductivo se basa en la transferencia de conocimientos sobre el lenguaje a un sistema computacional y el Aprendizaje inductivo se basa en la adquisición automática del conocimiento necesario sobre el lenguaje a transcribir, estos sistemas están basados en modelos de Markov o redes neuronales artificiales que trabajan a partir de muestras de datos.

Existen actualmente varias aplicaciones que consiguieron un nivel óptimo en el campo del reconocimiento de voz, además de contar con características adicionales que permiten cubrir cualquier necesidad. Entre las aplicaciones más reconocidas tenemos:

### 2.5.2.1 Google Cloud Speech-to-Text API

La API de Google para transcripción de voz usa modelos de redes neuronales y algoritmos de aprendizaje profundo para conseguir gran precisión en la conversión de

audio a texto. Esta API reconoce el lenguaje hablado de entre más de 120 idiomas, además de la comprensión de nombres propios y contexto.

Utiliza tecnología de aprendizaje automático para transcribir audio previamente gravados o en tiempo real. La precisión de Google Cloud mejora cada vez más y los modelos prediseñados se adaptan a las necesidades de cada usuario.

#### 2.5.2.2 IBM Watson Speech to Text API

Watson Speech fue creado por IBM para la transcripción de audio a a texto, utiliza inteligencia artificial con algoritmos de aprendizaje profundo, aplica conocimientos de estructura de lenguaje, forma de la señal, gramática entre otros, lo que hace de esta API una solución óptima para el reconocimiento de voz. Para un uso personalizado cuenta con las siguientes características:

- Transcripción en tiempo real o mediante carga de archivos.
- Diferentes modelos de idiomas.
- Opciones de frecuencia de audio.
- Modificación de características de salida de la transcripción.
- Organización de transcripciones mediante etiquetas, formato, palabra clave, tiempo y redacción.
- Control de seguridad de datos, IBM no recopila o utiliza los datos recopilados sin un acuerdo o aceptación.

Esta API se conecta a la nube de IBM mediante autenticación IAM (Identity and Access Management) basada en token.

### 2.5.2.3 Microsoft Azure Speech Services API

Azure es la API de Microsoft que parte del servicio de voz para convertirla en texto, cuenta con modelos personalizados con estilos de habla, terminología del idioma y ruido de fondo. Usa un aprendizaje con modelos de redes profundas para conseguir una paridad en el reconocimiento de voz humana.

Esta API permite la entrada de audio mediante micrófono en tiempo real o archivos almacenados a los que se puede acceder mediante Speech SDK compatible con WAV / PCM de 16 bits, 16 kHz / 8 kHz, audio de un solo canal y REST API que admite formatos de audio adicionales utilizando el punto final REST de voz a texto o el servicio de transcripción por lotes.

## 2.6 Transmisión

La transmisión imágenes y sonidos de forma óptima a través de televisión digital se realiza mediante la codificación de señales que se encuentran en forma binaria, de esta forma se puede lograr la compresión de aplicaciones interactivas, servicios adicionales e información de interés social que no se tiene en la televisión análoga. Para lograr esto se utilizan moduladores de señal y retransmisores, en este documento hablaremos de modulador DTU.215 de marca Dektec.

### 2.6.1 DTU-215

Dektec es una fábrica de dispositivos USB, tarjetas adicionales para PC, software profesional y convertidores IP para televisión digital. (Dektec, s.f.)

El DTU-215 es un modulador portátil alimentado desde USB-2 por lo que no requiere alimentación externa como se observa en la Figura 12, con atenuador

programable, conversión digital directa desde 47 a 1000 MHz para una excelente calidad de señal y un simulador completo de canales que incluyen AWGN y Multipath, en la Tabla 12 se encuentran todas las características con las que cuenta, además es amigable con el usuario puede ser usado en estándares terrestres y de cable.

**Tabla 12.**

*Características módulo DTU-215*

Características	
Frecuencia de RF	VHF, UHF 36 a 1002MHz
Ancho de Banda	Hasta 8MHz
Nivel de salida RF	-46 a -15dBm (QAM) -49 a -18dBm (OFDM)
Modulación	QAM-B, QAM-C, ATSC 8VSB, DVB-C, DVB- TATSC-M / H, CMMB, DAB, T-DMB, DTMB, DVB-C2, DVB-T2 multi PLP, ISDB-T.
Sistema Operativo	Windows 7, 8, 10 Linux $\geq$ 2.6.18, 3.x, 4.x
Procesador	Core i5 mínimo Core i7 recomendado CPU AMD o equivalentes

*Nota:* Información obtenida de ABNT NBR 15603-2

Para trabajar con este modulador existen varios softwares que pueden ser usados dependiendo de la necesidad. (Dektec, DTU-215, s.f.)

- StreamXpress: Reproduce archivos TS y parámetros de modulación de control.
- XpressSim: Simulador de canales (opción para StreamXpress)
- T2Xpress : Generador de señal DVB-T2 con control total de todos los parámetros de modulación y con simulador de canal integrado

- C2Xpress : Generador de señal DVB-C2
- TmmXpress: Generador de señal ISDB-Tmm
- SpRcApi: API de control remoto StreamXpress para reproducción y modulación automatizadas.
- DekTec SDK: Crea su propia aplicación totalmente personalizada con salida modulada a través del DTU-215.

**Figura 12.**

*Módulo DTU-215*



*Nota:* Información obtenida de Dektec

#### 2.6.1.1 Software de Transmisión StreamXpress

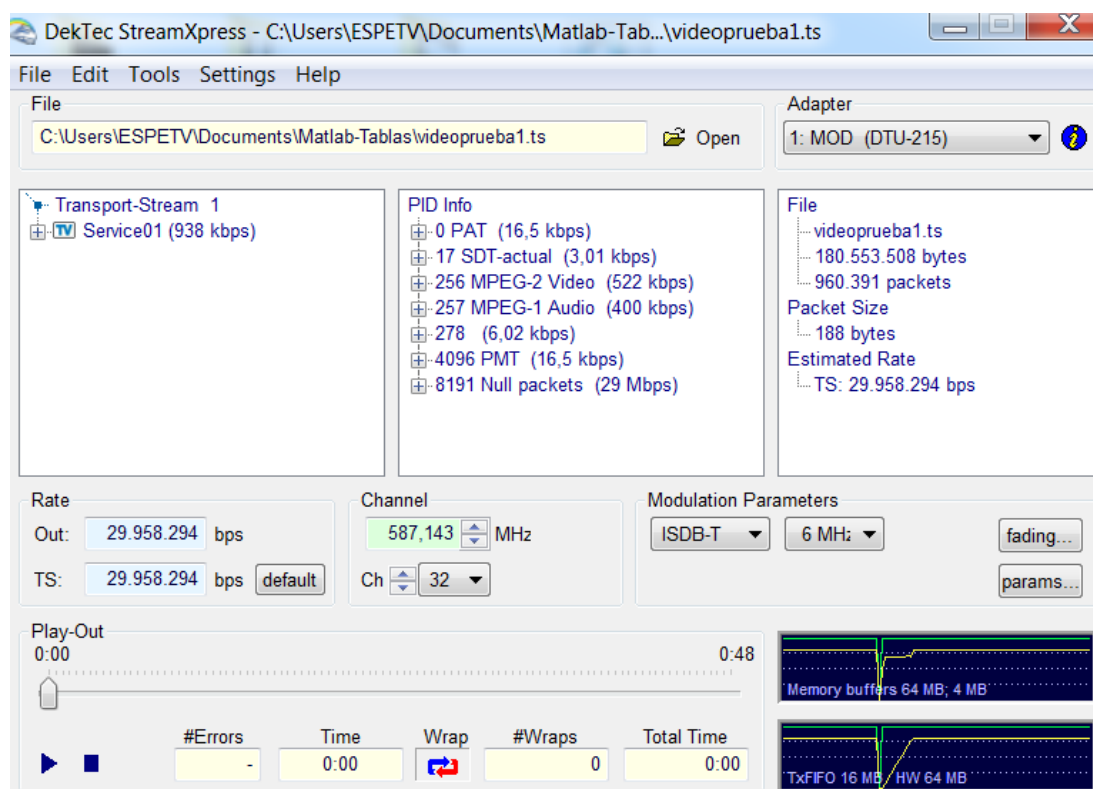
El reproductor de flujos de la marca Dektec es un software versátil de bajo costo y alto rendimiento que admite todos los adaptadores de salida DekTec como se observa en la Figura 13 además tiene corrección automática opcional de los campos de contador de continuidad, PCR / PTS / DTS y TDT / TOT y un fácil control de modificación de parámetros de modulación y simulación de canal.

- Varios tipos de flujos de transporte, SD / HD / 3G-SDI, muestras I / Q, archivo TMCC (para ISDB-T) y archivo T2-MI (para DVB-T2)
- Cálculo automático de velocidad

- Corrección de PCR
- Reproducción sin fin en bucle
- Generador de señal de prueba
- Inyección de error
- Control remoto
- Sintonización basada en canales

**Figura 13.**

*Interfaz de Software StreamXpress.*

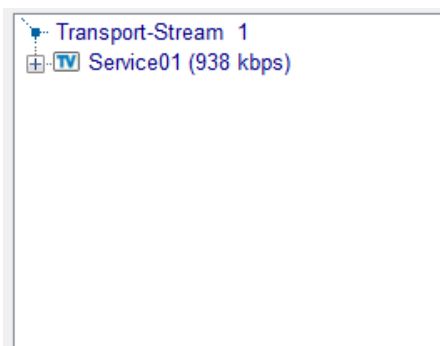


La interfaz cuenta con tres ventanas de información del archivo, en la primera ventana de la izquierda se representa la información del flujo de transporte que

mostrara la información de Transporte Stream ID, cada uno de los programas detectados con su nombre y tasa de bits totales, como se observa en la Figura 14.

#### **Figura 14.**

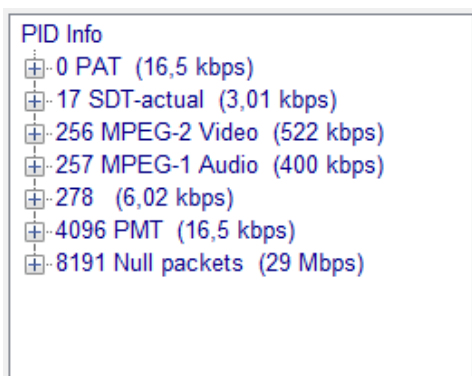
*Información de TS.*



En la ventana del medio se desplegara la Información PID que muestra cada PID y su asociación con un número de programa, bitrate del PID, presencia de PCR y Stream\_ID, como se observa en la Figura 15.

#### **Figura 15.**

*Información de PID.*

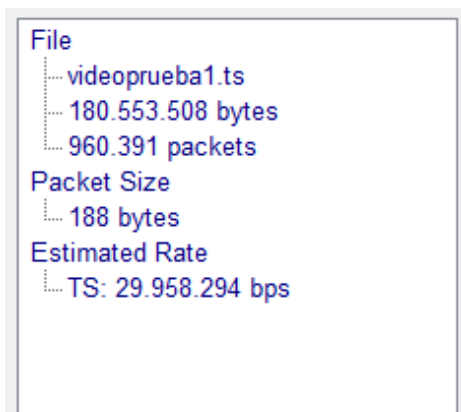


Y en la ventana del lado izquierdo se despliega la Información del archivo que muestra el nombre del archivo, tamaño de archivo en bytes, número de paquetes

completos en el archivo, tamaño del paquete y tasa estimada para el TS, como se observa en la Figura 16.

**Figura 16.**

*Información de archivo.*

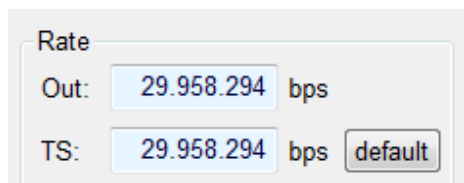


StreamXpress permite realizar algunos ajustes específicos para la reproducción de un flujo de transporte. En la sección de control de velocidad ofrece algunas capacidades poderosas con respecto a la configuración de la velocidad de salida de la reproducción. En la figura 17 podemos observar que se obtiene la tasa de TS del archivo derivado de PCR, la tasa de salida representa a la tasa de reproducción real y seguirá a la tasa de TS por defecto pero se puede modificar independientemente tomando en cuenta que pueden darse errores al momento de decodificar en el receptor. Además se tiene la opción de remultiplexar el TS abierto al agregar en tiempo real paquetes nulos (relleno) y recalcular PCR y PTS / DTS sobre la marcha.



**Figura 17.**

*Control de velocidad.*



Rate

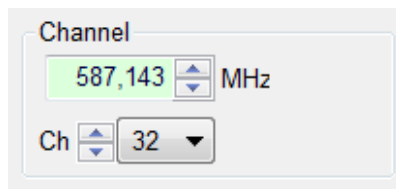
Out: 29.958.294 bps

TS: 29.958.294 bps default

En la sección de canal el usuario tiene que seleccionar que parte del espectro de RF enviará la señal modulada ingresando el canal y la frecuencia central de la señal digital modulada como se ve en la Figura 18. La mayoría de los dispositivos moduladores DekTec son completamente ágiles en la banda UHF / VHF.

**Figura 18.**

*Frecuencia de canal.*



Channel

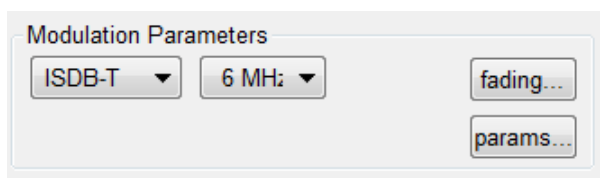
587,143 MHz

Ch 32

DekTec tiene licencia para admitir la modulación ISDB-T como se observa en la Figura 19. Existen muchos parámetros para ISDB-T, por lo que la mayoría de los parámetros se establecen en la opción "params". Permite modular en un ancho de banda de 6MHz u 8MHz pero se debe tomar en cuenta que el ancho de banda para ISDB-T es de 6MHz. Si el archivo abierto es un archivo TMCC y contiene parámetros de modulación ISDB-T, aparecerá una casilla de verificación.

**Figura 19.**

*Parámetros de Modulación.*



En la opción de “params” se modifican los parámetros de modulación ISDB-T que están organizados en 4 secciones como se observa en la Figura 20.

**Figura 20.**

*Parámetros ISDB-T.*

PID	Sys/Comp	Data	A	B	C
0x0...	PAT	16.544	A	-	-
	Service01	938.496			
0x0...	MPEG-2 Video	521.888	A	-	-
0x0...	MPEG-1 Audio	400.064	A	-	-
0x0...	SDT-actual	3.008	A	-	-
0x0...		6.016	A	-	-
0x1...	PMT	16.544	A	-	-

Map other PIDs to layer: --  Allow PID in multiple layers

ISDB-T Parameters  
 Broadcast Type: Television Mode: 3 Guard: 1/8 IIP PID: 0x1FF0  
 Partial Reception  Emergency Broadcasting

Layer Parameters

	#Segments	Modulation	Code Rate	Time Intlv	Rate (bps)	Selected (bps)
A	13	64QAM	3/4	I = 2	18.255.835	964.064
B	0	64QAM	3/4	I = 2	0	0
C	0	64QAM	3/4	I = 2	0	0
Total	13				18.255.835	964.064

Valid ISDB-T Settings

- 1- Componentes de MPEG-2 TS para mapeo de capas: Permite seleccionar qué componente del flujo de transporte se transportará en qué capa. Normalmente, un programa completo se llevará en una sola capa
- 2- Parámetros ISDB-T: Permite la selección de parámetros estándar ISDB-T.
  - Tipo de transmisión
  - Modo
  - Guardia
  - PID IIP
  - Recepción parcial
  - Transmisión de emergencia
- 3- Parámetros de capa: Esta sección permite la selección de modulación y asignación de segmentos para cada capa.
- 4- Centro de mensajes y configuraciones preestablecidas Parámetros 2-ISDB-T: Esta sección muestra cualquier error y permite al usuario seleccionar la configuración preestablecida.

Para la reproducción del video se tiene los botones de control para iniciar una reproducción y detener la reproducción o pausarla mientras se reproduce. Un contador de errores interno que se incrementa cuando el dispositivo descarta paquetes. La visualización del tiempo de reproducción desde total y parcial desde la última vez que se presionó el botón de parada.

**Figura 21.**

*Opciones de reproducción de video.*



Además muestra el tamaño del búfer de memoria del software utilizado para la reproducción y un gráfico que muestra la plenitud del búfer y el tamaño del búfer de memoria del dispositivo de hardware utilizado para la reproducción como se observa en la Figura 21. (Dektec, DTC-300-SP-StreamXpress USER MANUAL, 2017)

### 2.6.1.2 Software de Control de Transmisión ESPE Player

Este software fue desarrollado como proyecto de titulación en la Universidad de las Fuerzas Armadas por Roberto Mena, permite enviar un canal RF de videos TS y BTS con el estándar ISDB-Tb previamente generados además de la manipulación y control de transmisión a través de una tarjeta moduladora aprovechando todas sus funciones. El programa se presenta en una interfaz en java como en la Figura 22 y además cuenta con una librería dinámica en C++ (MENA BASTIDAS, 2015).

Esta plataforma permite controlar las funcionalidades de la transmisión para poder darle al usuario control y adaptación a sus necesidades. En la Figura 23 en la parte superior cuenta con el botón de abrir que permite escoger la dirección del video que se va a transmitir y una tabla PID que de ser posible indica los servicios del archivo TS cargado.

Figura 22.

Interfaz ESPE Player.

The screenshot shows the ESPE Player interface with the following sections:

- Top Section:** A text input field and an "Abrir" button.
- Table Section:** A table with columns "PID", "CAPA A", "CAPA B", and "CAPA C". Each row has a checkbox in each column.
- Parametros ISDB-T Section:**
  - Canal: 2
  - IIP PID: x1FF0
  - Utilizar Informacion TMCC:
  - Bytes/Paquete: -
  - Recepción Parcial:
  - Tipo de Transmision: Television
  - Modo: 1
  - Tiempo de Guarda: 1/32
  - Emergencia:
- Parametros de Capa Jerarquica Section:**

	# Segmentos	Modulación	Codificación	E. Tiempo	Capacidad Canal	Tasa Seleccionada
A	0	DQPSK	1/2	0/0/0		
B	0	DQPSK	1/2	0/0/0		
C	0	DQPSK	1/2	0/0/0		
TOTAL						
- Cuadro de Texto:** A large green text area.
- Bottom Section:**
  - Reproducir
  - Repetir:
  - Parar
  - Conectar
  - Cerrar

**Figura 23.***Lectura de video TS.*

PID	CAPA A	CAPA B	CAPA C
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Permite la configuración de los parámetros de ISDB-T como muestra la Figura 24, el tipo de transmisión, configuración de canal, modo que puede ser modo 1, modo 2, o modo3 con 108, 216 y 432 portadoras por segmento respectivamente, bytes por paquete, IIP PID con la información TMCC, tiempo de guarda, escoger si se va a realizar una recepción parcial del video en cuyo caso la transmisión one-seg debe tener un segmento en la capa A y además permite indicar si el video cuenta con transmisión de emergencia.

**Figura 24.***Parámetros ISDB-T.*

**Parámetros ISDB-T**

Utilizar Informacion TMCC

Canal: 2 IIP PID: x1FF0

Bytes/Paquete: -  Recepción Parcial

Tipo de Transmision: Television Modo: 1 Tiempo de Guarda: 1/32  Emergencia

En la Figura 25 se observa la sección de capa jerárquica que indica el número de segmentos de capa, modulación codificación, tasa de código interno, tiempo de entrelazado y tasa de transmisión.

**Figura 25.**

*Parámetros de capa jerárquica.*

Parámetros de Capa Jerárquica						
	# Segmentos	Modulación	Codificación	E. Tiempo	Capacidad Canal	Tasa Seleccionada
A	0	DQPSK	1/2	0/0/0		
B	0	DQPSK	1/2	0/0/0		
C	0	DQPSK	1/2	0/0/0		
TOTAL						

En la parte inferior se tiene los parámetros de reproducción del video como muestra la Figura 26, conexión con la tarjeta DTU-215, reproducción y pausa. Además un cuadro de mensaje que indica si existe algún error en la conexión con la tarjeta o transmisión.

**Figura 26.**

*Parámetros de reproducción.*

Cuadro de Texto				
Reproducir	<input type="checkbox"/> Repetir	Parar	Conectar	Cerrar

### 3 CAPITULO III: METODOLOGÍA

En este capítulo se explicará el desarrollo e implementación de audio por subtítulos como un servicio de televisión digital terrestre, mostrando todos los componentes necesarios de software y hardware para el reconocimiento de voz, creación de subtítulos, modificación de paquetes de texto, transmisión y recepción de la señal de TDT.

#### 3.1 Estructura del sistema

La estructura del sistema completo para la implementación de subtítulos en una señal de televisión digital se encuentra explicada en la Figura 27, en la cual se observa como primero ingresa el video en formato TS el cual es procesado por el servicio de reconocimiento de voz y se obtiene el texto del audio, este texto y el video original entran en el software multiplexor del que se obtiene el video subtulado en formato TS, este se envía al transmisor el cual a su vez lo transmite hacia el decodificador y finalmente al televisor digital.

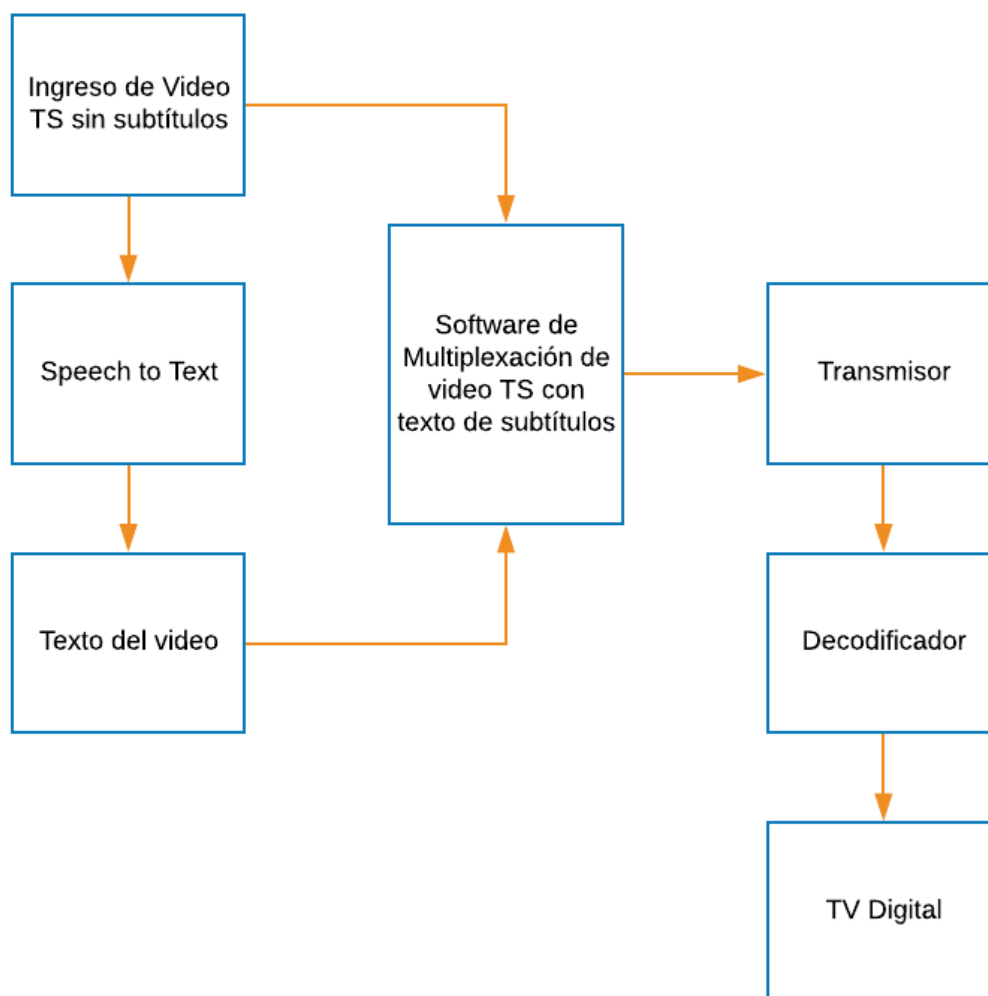
La implementación del proyecto se observa en la Figura 28 donde se muestran los equipos que se utilizan para la ejecución del proyecto. Los equipos son:

- PC Dell Inspiron 5458
- Transmisor Dektec DTU-215
- Decodificador Visiontec VT7200
- Televisor Panasonic



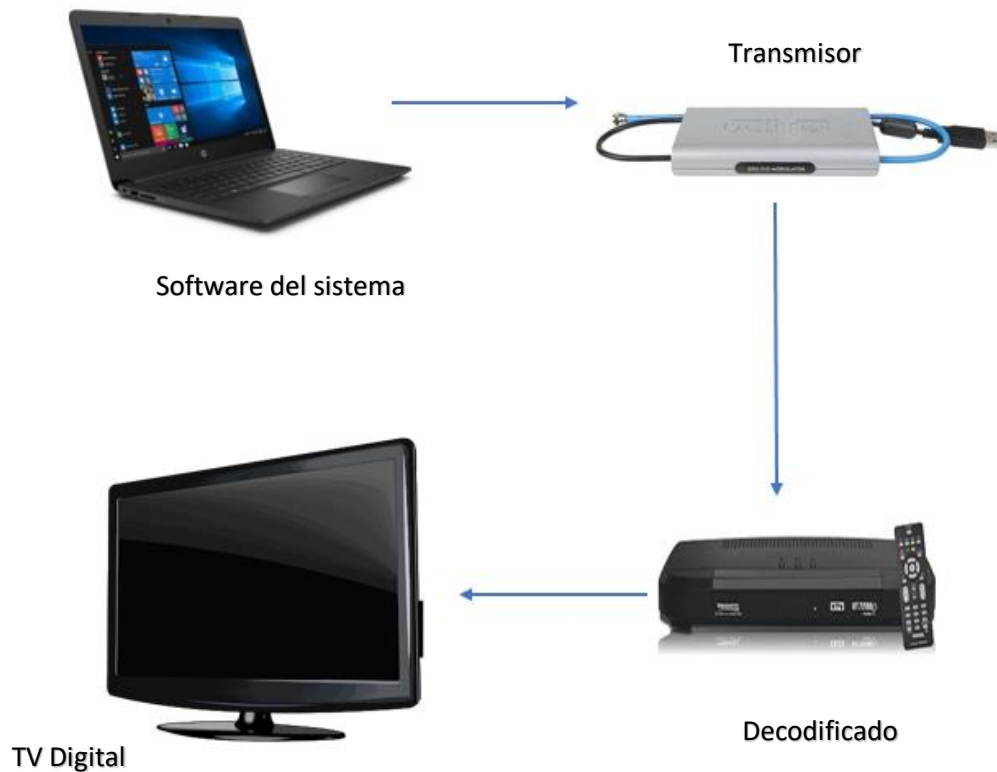
**Figura 27.**

Diagrama de bloques del sistema.



**Figura 28.**

*Esquema de implementación del sistema.*



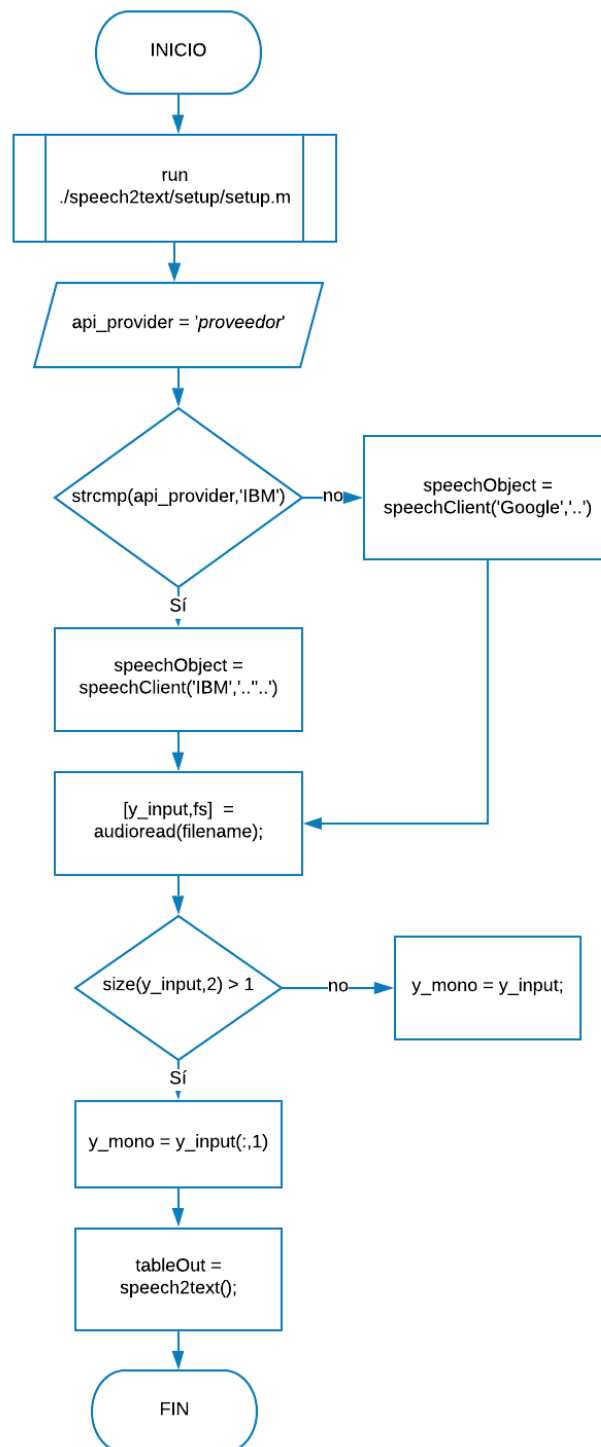
### 3.2 Reconocimiento de voz

Para el reconocimiento de voz se trabaja con la API de IBM, Watson Speech to text, y la API de Google, que son las aplicaciones que brindan todas las características que se busca para el diseño del proyecto como se indicó en la sección 2.5.2.

Se realizó un programa en Matlab que permite conectar el servicio de Google Cloud o de IBM Watson con el software de Matlab y de esta manera usar los recursos y características que ofrecen las APIs en el programa. En la Figura 29 se observa el diagrama de flujos del programa de reconocimiento de voz.

**Figura 29.**

*Diagrama de Flujos programa de reconocimiento de voz*



El programa ejecuta la función `speech2text` que es una función creada por Matlab específicamente para trabajar con las API de audio a texto de diferentes proveedores de este servicio en la nube, después de esto se seleccione el proveedor con el que se va a trabajar, en nuestro caso puede ser IBM o Google, dependiendo del proveedor se crea un objeto de voz con sus características. Se lee el archivo de entrada, en caso de ser estéreo se lo cambia a mono, y finalmente para obtener la salida se coloca en la función de `speech2text` el objeto de voz y el archivo de audio.

### 3.2.1 IBM Watson Speech to Text API

La API de IBM es accesible ya que cuenta con una versión gratuita que tiene todas las características necesarias para la conversión de audio a texto de los videos que se usaran en el proyecto.

Primero se debe crear una cuenta de IBM Cloud en la que se activa un servicio que contenga la API de voz a texto para agregarla a la cuenta, una vez hecho esto se ingresa al panel de servicios que contiene las credenciales, clave API y valores de URL como se muestra en la Figura 30.

#### Figura 30.


*Credenciales de API Watson Speech to text.*

#### Credenciales


Descargar ↓

Mostrar credenciales ©

Clave de API:

..... 

URL:

`https://api.us-south.speech-to-text.watson.cloud.ibm.com/instances/58259aeb-3746-4a7e-82c5-` 

Después se escribe las credenciales copiadas en un archivo "IBM\_Credentials\_Speech2text.json" con el atributo "apikey" para la clave y URL, como se muestra a continuación:

```
{  
  
    apikey: .....  
  
    URL: "https://api.us-south.speech-to-text.watson.cloud.ibm.com/insta...  
  
}
```

Finalmente el programa de Matlab mediante la función `speech2text` buscará "IBM\_Credentials\_Speech2text.json" para obtener la clave API y el URL necesarios para conectarse con la API. (MathWorks, 2020)

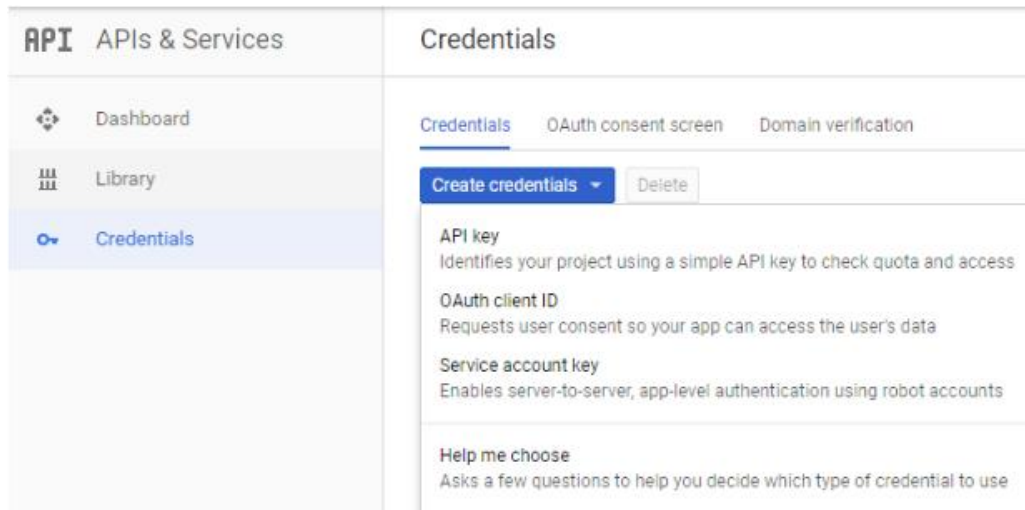
### 3.2.2 Google Cloud Speech-to-Text API

La API de Google Cloud, Speech-to-Text API también cuenta con todas las características necesarias en su versión gratuita, además consta de un panel de control con el cual se puede monitorizar la eficiencia de la aplicación mediante gráficas y estadísticas.

En la cuenta de Google Cloud se crea un proyecto y se incluye el servicio de la API Speech to text, después en el panel de servicios, en la opción credenciales se escoge la opción para crear credenciales como se observa en la Figura 30.

**Figura 31.**

*Credenciales de Google Cloud.*



Una vez creada la clave la copiamos en el archivo

"Google\_Credentials\_Speech2text.json" con el atributo "clave" como se indica a continuación:

```
{
  "key" : "AlzaSyCIUzlgafziYDwiJoKNI..."
}
```

Opcionalmente se puede colocar el atributo "recognizeUrl", pero si no se coloca el predeterminado es <https://speech.googleapis.com/v1/speech:recognize>.

Finalmente el programa de Matlab mediante la función speech2text, que buscará "Google\_Credentials\_Speech2text.json" para obtener la clave que la API Google Speech requiere para la autenticación. (MathWorks, 2020)

### 3.3 Creación de paquetes

En esta sección se realizara la modificación y creación de paquetes necesarios para la inclusión de subtítulos como un servicio más en el flujo de trasporte del video a transmitir, se modificaran las tablas PSI/SI para incluir el servicio de subtítulos y sus características y se crearan los paquetes PES de texto con las características del mensaje.

Se analiza primero el paquete PAT que contiene el valor de los PIDs de los diferentes archivos .TS incluidos en el video. Al analizar este paquete encontraremos en valor del PID de la tabla PMT.

#### Figura 32.

*Paquete de tabla PAT.*

```
47 40 00 10 00 00 B0 0D 00 01 C1 00 00 00 01 F0 00 2A B1 04 B2 FF FF FF
FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
```

Como se indicó en la sección 2.3.3 si el valor del *Broadcasting program number identifier* (señalado en color verde) es igual a 0 el PID corresponde a la tabla NIT y si el valor del *Broadcasting program number identifier* es diferente de 0 el valor corresponde a la tabla PMT, entonces según la Figura 32 el valor de PID de la tabla PMT es '1000' o 4096 en decimal.

### 3.3.1 Paquetes PMT

Como se observó en la sección anterior la tabla PMT tiene el valor de PID de 4096, ahora se procede a modificarla como se indica en la sección 2.3.4 colocando los servicios y descriptores necesarios.

Primero se identifican los tipos de flujos existentes en la tabla y los descriptores con los que cuenta cada uno con base a la Figura 9 de la sección 2 que indica la estructura del PMT. Se ubican los flujos de datos de audio y video que en este caso tienen el valor de '02' para video conforme la ITU y de '03' para audio conforme la ISO/IEC 11172-3.

**Tabla 13**

*Estructura de flujo de datos de paquetes PES*

Campo	Valor
Stream type identifier	0x06
Reservado	'111'
Elementary stream PID	0x116
Reservado	'1111'
ES information length	0x008
Descriptors file 2	Descriptores

Se añade el flujo de datos para paquetes PES en los que se colocaran los subtítulos, con los valores que se encuentran en la Tabla 13 cuya estructura se explica en la sección 2 y además la estructura de los descriptores para este servicio, que se encuentran detallados en la Tabla 14.





### 3.3.2 Paquetes de Control de Subtítulos

Son modificados los paquetes PES con valor de PID de 0x116 o 278 en decimal, en estos paquetes se coloca el flujo de datos de los subtítulos con la información de control antes de enviar el mensaje, tal como se indicó en la sección 2.3.5 de la sección anterior.

Se coloca la información inicial de los paquetes PES para subtítulos en TDT con los valores de la Tabla 15, cuya estructura se explica en la tabla 6 de la sección 2.3.5.

**Tabla 15**

*Valores iniciales de paquetes PES para subtítulos.*

<b>Campo</b>	<b>Valor</b>
Prefijo de inicio	0x000001
Stream_ID	0xBD
PES_packet_length	0X002E

Seguida de la información para datos sincronizados como lo son los subtítulos, con los valores de la Tabla 16. Estos valores se los ubican dentro del paquete PES como se observa en la Figura 34.

**Tabla 16**

*Valores de paquetes PES sincronizados.*

<b>Campo</b>	<b>Valor</b>
Data_identifier	0x80
Flags	0x81
PES_header_data_length	0x17
Reserverd	'0010'
PTS[32..30] y Marker_bit	0x03
PTS[29..15] y Marker_bit	0xDB6F

PTS[14..0] y Marker_bit	0XE7A9
Flags	0x8E
User area	0xFF

**Figura 34.**

*Parámetros de subtítulos en Paquete PES.*

```

47 41 16 35 83 00 FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
17 23 DB 6F E7 A9 8E FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
80 FF F0 00 00 00 00 0A 3F 01 1A 73 70 61 80 00 00 00 24 BB

```

Después colocamos la información de los parámetros de grupos de datos que se encuentran en la Tabla 7 con los valores indicados en la Tabla 17. Estos valores se los ubican dentro del paquete PES como se observa en la Figura 35.

**Tabla 17**

*Valor de parámetros de grupo de datos.*

Campo	Valor
Data_identifier	0x80
Private_stream_id	0xFF
Reservado	0xF
PES data packet header length	0x0
Data_group_id	0x00
Data_group_version	0x00
Data_group_link_number	0x00

Last_data_group_link_number	0x00
Data_group_size	0x0A
Data_group_data_byte	Almacenamiento del grupo de datos.
CRC16	Código de redundancia cíclica.

**Figura 35.**

*Parámetros de grupo de datos en paquete PES.*

```

47 41 16 35 83 00 FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
80 FF F0 00 00 00 00 0A 3F 01 1A 73 70 61 80 00 00 00 24 BB

```

Finalmente se colocan los parámetros de datos de gestión de subtítulos indicados en la Tabla 8 con los parámetros de la Tabla 18 y se los ubica en el paquete PES como se puede ver en la Figura 36.

**Tabla 18**

*Valor de parámetros de datos de gestión.*

Campo	Valor
TDM	'0000'
Num_languages	'0001'
Language_tag	'0001'
DMF	'1010'
ISO_639_language_code	0x737061
Formato	'1000'



sección 2.3.5.1, indicando el tamaño, color, posición, espaciado, modo de escritura, entre otros como se observa en la Tabla 19 y se coloca en el paquete PES en la Figura 37.

**Tabla 19**

*Caracteres de configuración para subtítulos en paquete PES*

Carácter	Valor	Función	Carácter	Valor	Función
0C	CS	Limpiar pantalla	36	6	
9B	CSI	Introducción a control de secuencia	3B	;	
37	7		33	3	
20	SP	Espacio	36	6	
53	S		20	SP	Espacio
9B	CSI	Introducción a control de secuencia	57	W	
31	1		9B	CSI	Introducción a control de secuencia
37	7		32	2	
36	6		20	SP	Espacio
3B	;		58	X	
31	1		9B	CSI	Introducción a control de secuencia
30	0		31	1	
31	1		36	6	
20	SP	Espacio	20	SP	Espacio
5F			59	Y	

9B	CSI	Introducción a control de secuencia	9B	CSI	Introducción a control de secuencia
36	6		38	8	
30	0		20	SP	
38	8		6E	N	
3B	;		88	SSZ	Tamaño pequeño
33	3		1C	APS	Especificación de posición
39	9		4D	M	
30	0		40	@	
20	SP	Espacio	87	WHF	Color de fondo blanco
56	V		90	COL	Especificación de color
9B	CSI	Introducción a control de secuencia	50	P	
33	3				

**Figura 37.**

*Características del mensaje en paquete PES.*

```

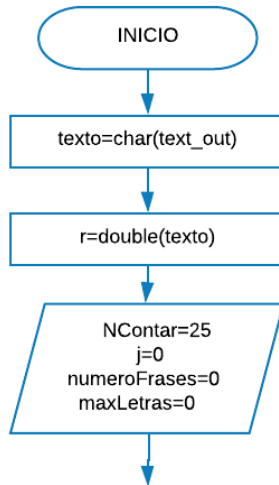
47 41 16 39 0C 00 FF FF FF FF FF FF FF FF FF FF FF FF 00 00 01 BD 00 A5 80
81 17 23 DB 79 2A 95 8E FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF FF
FF 80 FF F0 04 00 00 00 81 3F 00 00 7D 1F 20 00 00 78 0C 9B 37 20 53 9B
31 37 36 3B 31 30 31 20 5F 9B 36 30 38 3B 33 39 30 20 56 9B 33 36 3B 33
36 20 57 9B 32 20 58 9B 31 36 20 59 9B 38 20 6E 88 1C 4D 40 87 90 50 6C
61 73 20 72 65 66 6F 72 6D 61 73 20 61 20 6C 61 20 6C 65 79 20 6F 72 67
E1 6E 69 63 61 09 09 09 09 09 09 09 09 09 09 09 09 09 09 09 09 09 09 09

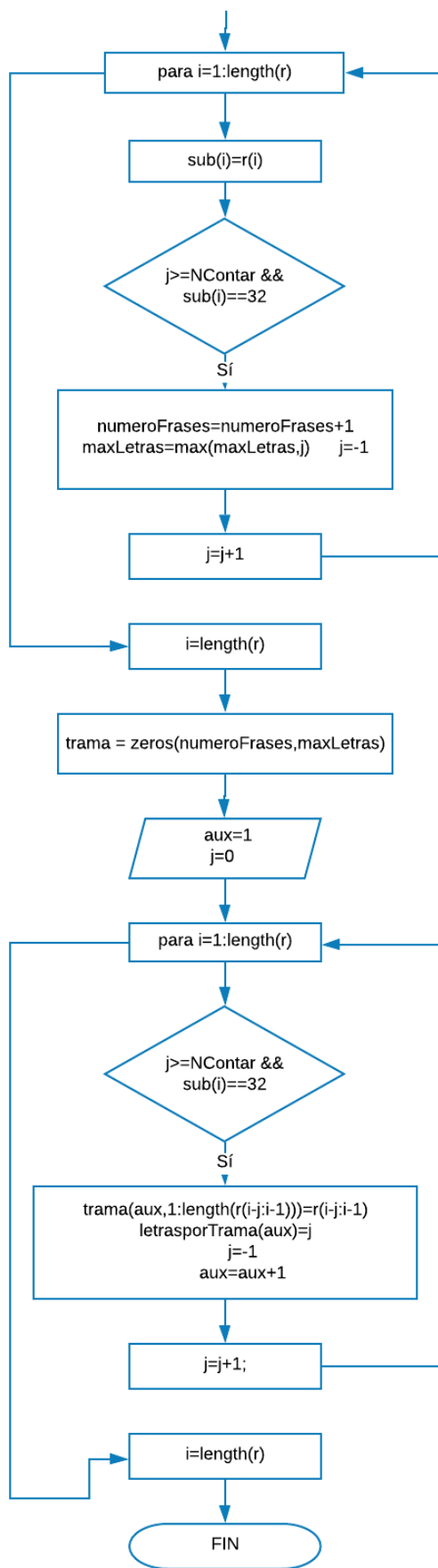
```











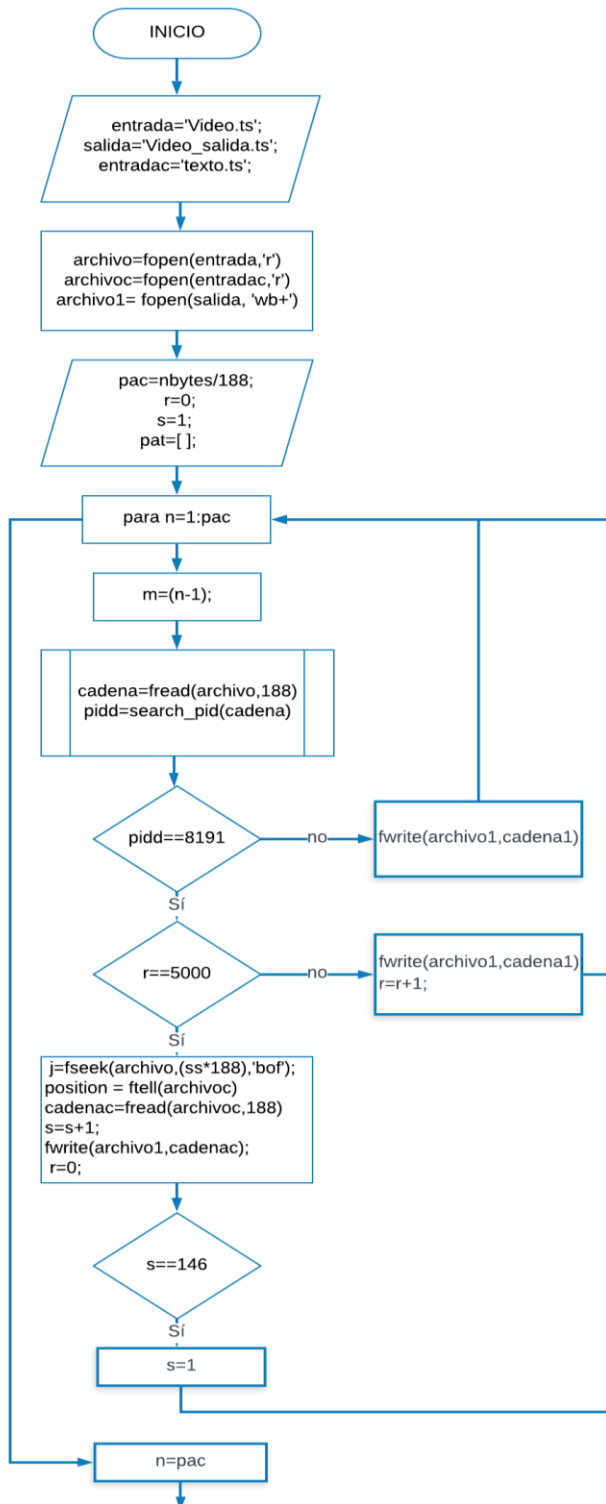
En el programa ingresa el texto convertido mediante el servicio de reconocimiento de voz, se lo divide en cadenas de 25 caracteres con las condiciones de que terminen en espacio para no dejar palabra inconclusas y finalmente se va colocando cada cadena en una matriz.

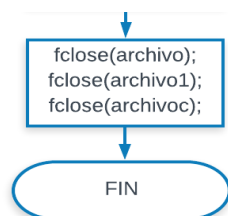
### **3.4 Multiplexación**

Con los paquetes de audio, video, datos con los subtítulos y la configuración deseada, se realiza la multiplexación que es la combinación de los paquetes que al momento los tenemos por separado, para esto se desarrolló un programa en Matlab “multiplexación” en el cual ingresa el video inicial sin subtítulos, y los paquetes PES con el texto de los subtítulos, los paquetes PES son colocados en el PID 8191 que inicialmente contenía paquetes vacíos y finalmente se guarda en un nuevo archivo .TS que contiene el video subtulado, tal como se observa en el diagrama de flujo de la Figura 40.

Figura 40.

Diagrama de flujo de programa de multiplexación.





### 3.5 Transmisión

Para la transmisión del video multiplexado se ocupa la tarjeta DTU-215 cuyas características están detalladas en la sección 2.6.1 y como software de transmisión se utiliza StreamXpress, que es un software de reproducción de Dektec y esta detallado en la sección 2.6.1.1. Además se transmite con un software adicional que se detalla en la sección 2.6.1.2, a este software se lo modificó para que permita transmitir de forma continua, en tiempo real.

#### 3.5.1 Software de Control de Transmisión ESPE Player

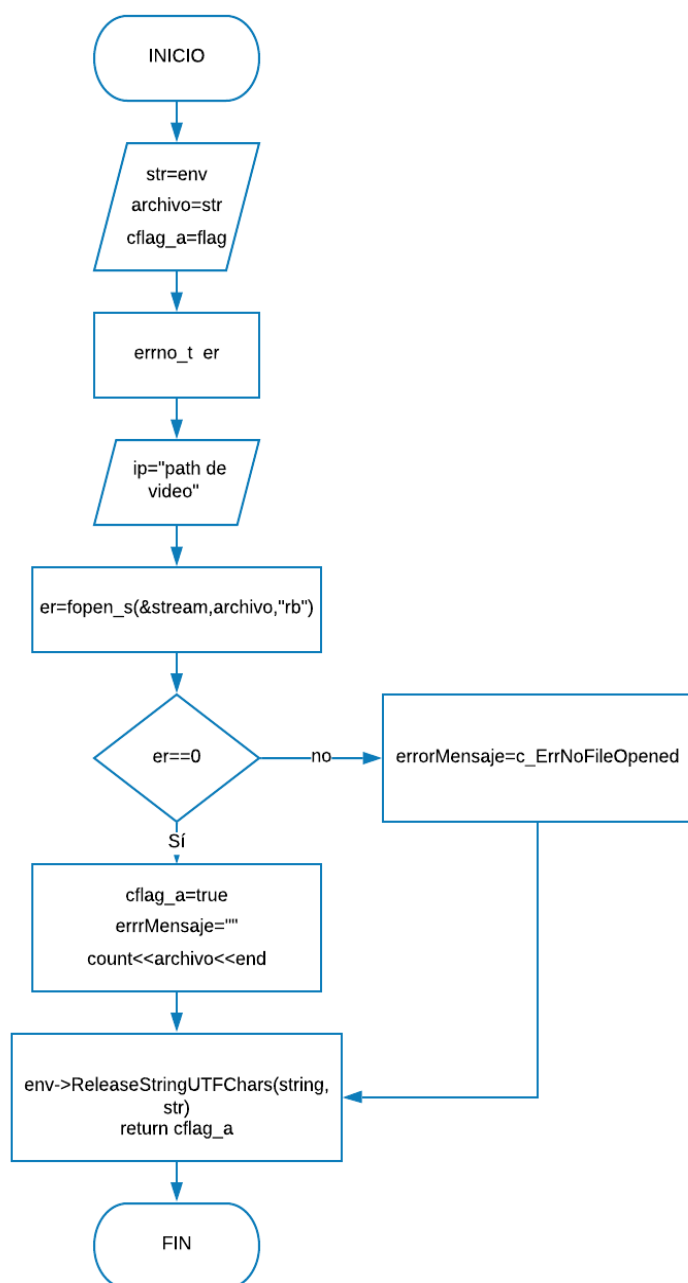
Este software es un reproductor de archivos .TS desarrollado como proyecto de titulación, para utilizarlo se verifica inicialmente que se tengan instalados los archivos JDK (Java Development Kit) y JRE (Java Runtime Environment) para 32 bits, se recomienda usar Java 1.8.0\_25. Este software cuenta con el proyecto de la librería dinámica en C++ y el proyecto de Java para realizar modificaciones a necesidad del usuario.

Para realizar la transmisión en vivo y que se puedan leer los datos mientras aún no se concluye la generación del video, se realizó la modificación al código fuente del software ESPE Player en la función que realiza la lectura del archivo de la librería

dinámica en C++. El diagrama de flujo del programa modificado se muestra en la Figura 41.

**Figura 41.**

*Diagrama de Flujo de la función modificada.*



Esta función permite la selección de la dirección en la que se encuentra el video y abre el archivo para su lectura, se añade al código la función `fsopen` que permite abrir una secuencia con uso compartido de archivos en la cual se coloca como argumentos que los archivos serán abiertos en forma binaria y con acceso a lectura y escritura, de esta manera el archivo puede ser abierto para su lectura al tiempo que está siendo escrito permitiendo una transmisión en tiempo real. Una vez realizada esta modificación se genera nuevamente el proyecto en Java y la aplicación que no tendrá ningún cambio en su diseño, como se observa en la Figura 42.

**Figura 42.**

*Aplicación ESPE Player*

The screenshot shows the ESPE Player application interface. At the top, there is a file path input field containing "C:\Users\ESPETV\Documents\Matlab-Tablas\videoprueba1.ts" and an "Abrir" button. Below this is a table for selecting video streams:

PID	CAPA A	CAPA B	CAPA C
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Below the table are the "Parametros ISDB-T" settings:

- Utilizar Información TMCC
- Canal: 7
- IIP PID: x1FF0
- Bytes/Paquete: 188
- Recepción Parcial
- Tipo de Transmision: Television
- Modo: 1
- Tiempo de Guarda: 1/32
- Emergencia

Next are the "Parametros de Capa Jerarquica" settings:

	# Segmentos	Modulación	Codificación	E. Tiempo	Capacidad Canal	Tasa Seleccionada
A	13	64QAM	1/2	0/0/0	13.276	
B	0	DQPSK	1/2	0/0/0	0	
C	0	DQPSK	1/2	0/0/0	0	
TOTAL					13.276	

At the bottom, there is a "Cuadro de Texto" (Text Box) displaying "La Tarjeta DTU 215 está conectada". Below the text box are control buttons: "Reproducir",  Repetir, "Parar", "Conectar", and "Cerrar".



## 4 CAPITULO IV: RESULTADOS

En esta sección se realizaron las pruebas de subtítulo en las que se ingresa un video sin subtítulos que ingresa los programas de reconocimiento de voz, creación de paquetes y multiplexación para ser transmitido mediante el software StreamXpress a través del modulador DTU-215 hacia el decodificador y finalmente ser visualizado en el televisor digital. Además se realizara una prueba de transmisión en tiempo real usando el software ESPE Player como software transmisor.

Para realizar las pruebas de subtítulo primero se ingresa el video que se desea subtítular al programa de reconocimiento de voz seleccionando el proveedor de servicio, la Figuras 43 contienen una parte del texto traducido del videos de prueba 1 usando como proveedor de servicio a Google y la Figura 44 contiene una parte del texto traducido del video de prueba 2 usando como proveedor de servicio a IBM.

### **Figura 43.**

*Conversión a texto de video de prueba 1 con Google.*

```
text_out =
    "las reformas a la Ley Orgánica de telecomunicaciones tienen como objetivo pasar de tener una ley
```

### **Figura 44.**

*Conversión a texto de video de prueba 2 con IBM.*

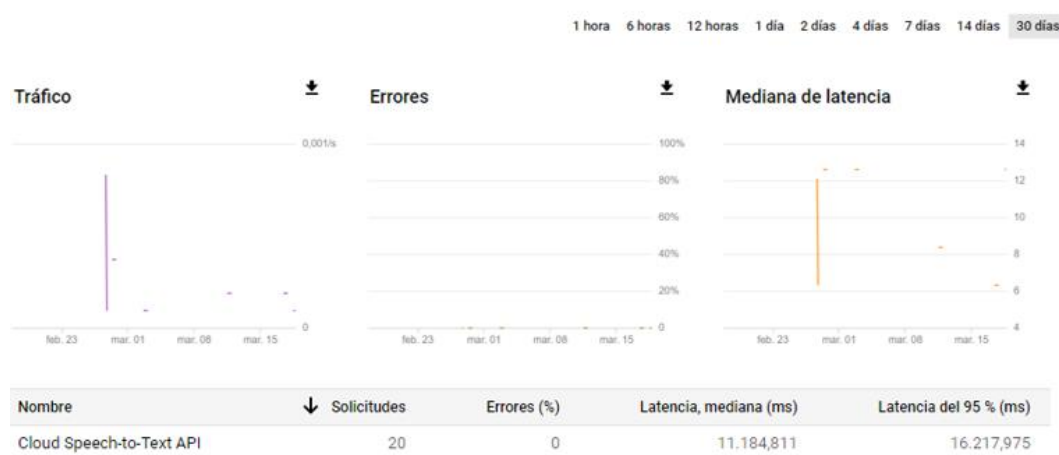
```
text_out =
    "Ecuador digital yo creo nos vemos como un instrumento muy apropiado para los tiempos de hoy fue la
```

Al realizar la traducción a texto con el programa de Google, este presenta un panel de control en el que se muestran los valores de tráfico, errores y latencia de todas

las solicitudes de traducción que se han realizado en los últimos 30 días, de esta manera se puede medir el nivel de confiabilidad de la API. Como se observa Figura 45 en las pruebas que se han realizado en los últimos 30 días se tiene un porcentaje de error de 0% por parte de la API.

**Figura 45.**

*Panel de control de API de Google.*



Con el texto traducido el software realiza la configuración de los subtítulos y la multiplexación junto con el video para finalmente tener un archivo .TS con todos los servicios deseados.

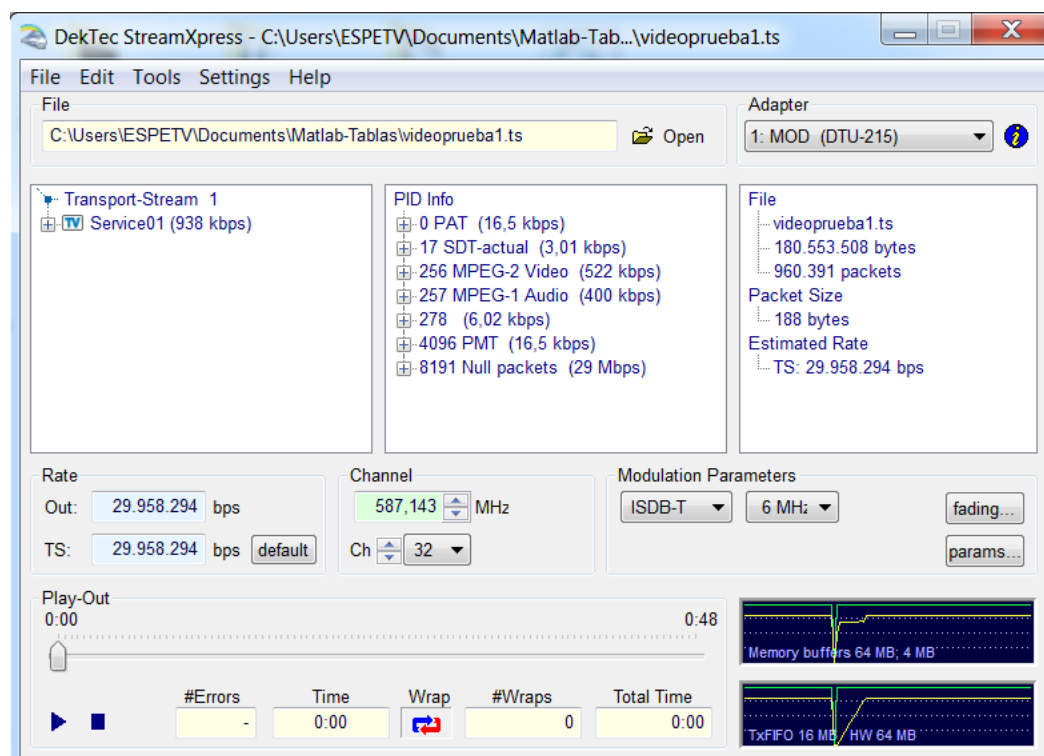
#### 4.1 Transmisión con StreamXpress

La transmisión del archivo .TS se lo realiza con el software StreamXpress, asociando en este caso la tarjeta DTU-215, usualmente esa opción suele seleccionarse automáticamente al momento de conectar la tarjeta, tal como se muestra en la Figura 46. En los recuadros se puede observar toda la información de video de prueba 1, los valores de PID de las tablas y servicios con los que cuenta, como se observa en la

Figura 46. El video cuenta con las tablas PAT, SDT, PMT, los paquetes de audio, video, datos y nulos. El tamaño del archivo de video 1 es de 180.553.508 bytes y cuenta con 960.391 paquetes de 188 bytes cada uno a una velocidad de transmisión de 29.958.294 bps. La velocidad de salida es configurada con la misma velocidad de TS para no tener problemas en la transmisión. Las pruebas se realizaron en el canal 32 que tiene una frecuencia de 587,143 MHz y en los parámetros de modulación con base al estándar ISDB-T. Una vez terminada la configuración, se inicia la transmisión. Los equipos necesarios para la transmisión se observan en la Figura 47.

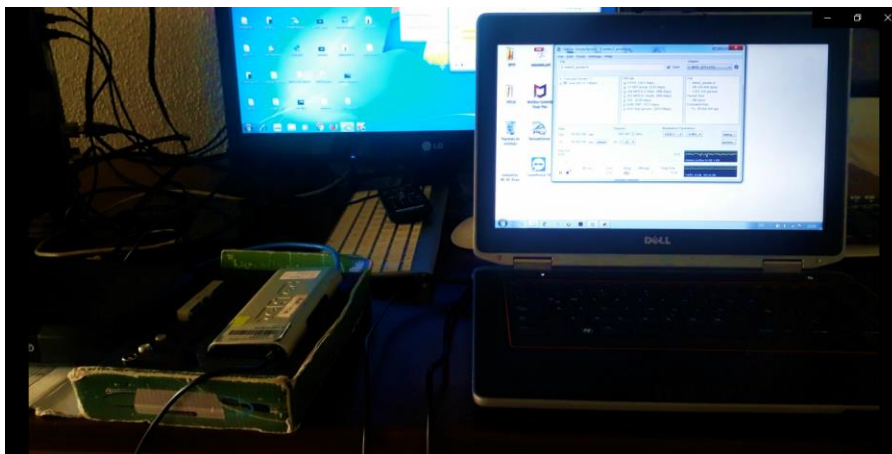
### Figura 46.

#### Configuración de software StreamXpress



**Figura 47.**

*Escenario de Transmisión de video.*



Para la recepción del video se utilizó el decodificador Visiontec VT7200 que recibe la señal del transmisor, la demultiplexa y la visualiza al televisor digital en la frecuencia del canal 32, esto se puede hacer de manera manual o automática. Una vez sintonizado el canal se muestra en la pantalla del televisor la reproducción con los subtítulos sincronizados con el audio y video como se observa en la Figura 48, para ver el video de prueba se tiene el código QR en la Figura 49.

**Figura 48.**

*Reproducción del video de prueba 1.*



**Figura 49.**

*Video de prueba 1.*



Podemos verificar la configuración de los paquetes del video TS con el analizador “TS & BTS ESPE-ANALYZER” en el cual se puede realizar una búsqueda por PID o por paquete para realizar el análisis necesario. En la Figura 50 se observa los bytes de uno de los paquetes de subtítulos del video de prueba 1.

**Figura 50.**

*Análisis de paquete de subtítulos de video de prueba 1.*

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
47	41	16	39	0C	00	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	00	00	01	BD	00	A5	80
81	17	23	DB	79	2A	95	8E	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF
FF	80	FF	F0	04	00	00	00	81	3F	00	00	7D	1F	20	00	00	78	0C	9B	37	20	53	9B
31	37	36	3B	31	30	31	20	5F	9B	36	30	38	3B	33	39	30	20	56	9B	33	36	3B	33
36	20	57	9B	32	20	58	9B	31	36	20	59	9B	38	20	6E	88	1C	4D	40	87	90	50	79
20	65	78	70	6C	6F	74	61	63	69	F3	6E	20	64	65	6C	20	65	73	70	65	63	74	72
6F	09	09	09	09	09	09	09	09	09	09	09	09	09	09	09	09	09	09	09	09	09	09	09
09	09	09	09	09	09	09	09	09	09	09	09	09	09	09	09	09	09	B7	6B				

PAT TABLE: 
 PMT TABLE:

NIT TABLE: 
 SDT TABLE:

Package Number TS:

En la Figura 51 se observa la reproducción del video de prueba 2 y para ver el video completo se tiene el código QR en el Figura 52, además se realizó el análisis de uno de los paquetes con el analizador “TS & BTS ESPE-ANALYZER” como se observa en la Figura 53.

**Figura 51.**

*Reproducción del video de prueba 2.*



**Figura 52.**

*Video de prueba2.*



**Figura 53.**

*Análisis de paquete de subtítulos de video de prueba 2.*

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
47	41	16	39	0C	00	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	00	00	01	BD	00	A5	80
81	17	23	DB	79	2A	95	8E	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF
FF	80	FF	F0	04	00	00	00	81	3F	00	00	7D	1F	20	00	00	78	0C	9B	37	20	53	9B
31	37	36	3B	31	30	31	20	5F	9B	36	30	38	3B	33	39	30	20	56	9B	33	36	3B	33
36	20	57	9B	32	20	58	9B	31	36	20	59	9B	38	20	6E	88	1C	4D	40	87	90	50	45
73	74	65	20	65	73	20	65	6C	20	67	72	61	6E	20	72	65	74	6F	20	71	75	65	20
74	65	6E	65	6D	6F	73	09	09	09	09	09	09	09	09	09	09	09	09	09	09	09	09	09
09	09	09	09	09	09	09	09	09	09	09	09	09	09	09	09	09	09	46	3A				

PAT TABLE:

PMT TABLE:

NIT TABLE:

SDT TABLE:

Package Number TS:

## 4.2 Software de Control de Transmisión ESPE Player

Una vez realizada la modificación de la función abrir, se generó nuevamente la aplicación ESPE Player, que ahora permite la lectura de un archivo que aún está siendo generado. Se selecciona el video .TS original que se desea transmitir, este puede ser un video grabado anteriormente o un video que este siendo generando en ese momento. Para la prueba se seleccionó el video denominado "videoprueba1.ts" que está siendo generado en el momento. La dirección donde se encuentra el video aparecerá en el recuadro superior, como se observa en la Figura 54. En el tipo de transmisión se coloca televisión, se elige el canal 7 y el modo 1 con un tiempo de guarda de 1/32. En los parámetros de capas se selecciona los 13 segmentos en la capa A con una modulación de 64QAM.

**Figura 54.**

*Configuración de software ESPE Player.*

C:\Users\ESPETV\Documents\Matlab-Tablas\videoprueba1.ts Abrir

PID	CAPA A	CAPA B	CAPA C
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Parametros ISDB-T**

Utilizar Información TMCC

Canal 7 IIP PID x1FF0

Bytes/Paquete 188  Recepción Parcial

Tipo de Transmision Television Modo 1 Tiempo de Guarda 1/32  Emergencia

**Parametros de Capa Jerarquica**

	# Segmentos	Modulación	Codificación	E. Tiempo	Capacidad Canal	Tasa Seleccionada
A	<span>13</span>	<span>64QAM</span>	<span>1/2</span>	<span>0/0/0</span>	<span>13.276</span>	
B	<span>0</span>	<span>DQPSK</span>	<span>1/2</span>	<span>0/0/0</span>	<span>0</span>	
C	<span>0</span>	<span>DQPSK</span>	<span>1/2</span>	<span>0/0/0</span>	<span>0</span>	
TOTAL					<span>13.276</span>	

**Cuadro de Texto** La Tarjeta DTU 215 está conectada

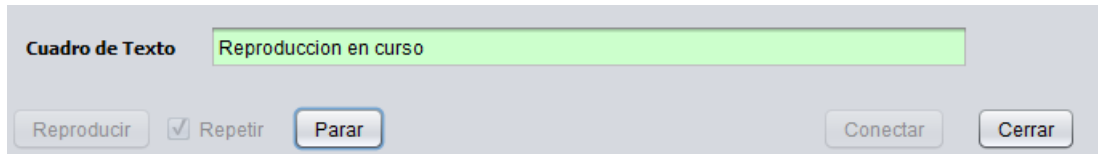
Reproducir  Repetir  Parar Conectar Cerrar

Una vez realizadas las modificaciones se selecciona el botón de conectar para verificar que exista conexión con la tarjeta DTU-215, el mensaje aparecerá en el recuadro inferior, si se visualiza el mensaje “La Tarjeta DTU 215 está conectada” entonces se puede iniciar la transmisión.



**Figura 55.**

*Reproducción de video en software ESPE Player*

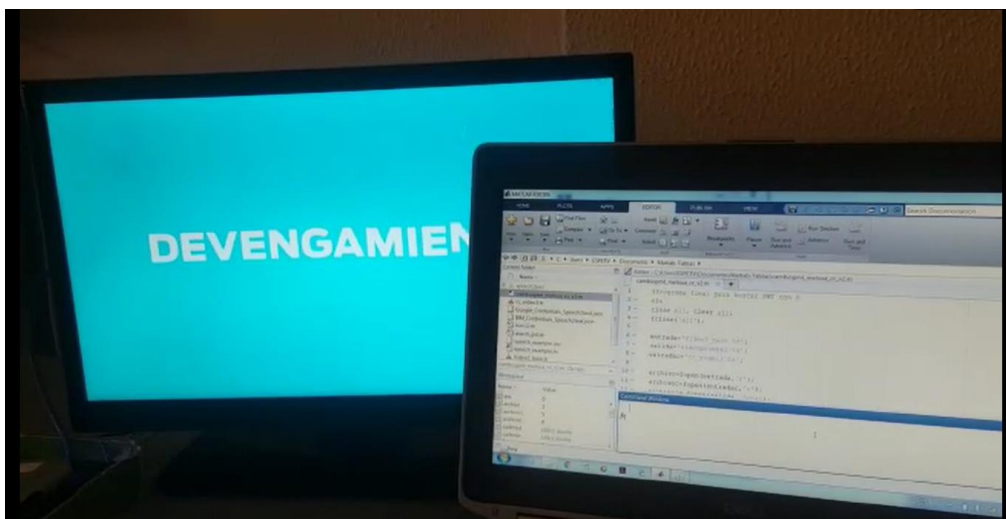


Se oprime el botón “*Reproducir*” y si el aparece el mensaje “*Reproducción en curso*” como se observa en la Figura 55, la transmisión se está realizando correctamente, si por el contrario indica algún error, se debe revisar la configuración y conexión de la tarjeta nuevamente.

Por último se observa en la Figura 56 la reproducción del video en pantalla, mientras se observa como el archivo en Matlab se está generando simultáneamente, para ver el video completo se tiene el código QR de la Figura 57.

**Figura 56.**

*Reproducción en tiempo real con ESPE Player.*



**Figura 57.**

*Video de reproducción en tiempo real.*



En el video de prueba 1 usando el servicio de Google se obtuvo los resultados esperados en la conversión a texto y se observó la eficacia del servicio en su panel de control, la creación y multiplexación de paquetes se realizó correctamente y se verificó con el analizador “TS & BTS ESPE-ANALYZER” y la transmisión y recepción del video se ejecutó sin problema, en el televisor se visualizó un texto sincronizado y claro.

En el video de prueba 2 se usó el servicio de IBM obteniendo también buenos resultados en cuanto a la conversión a texto, la creación y multiplexación de paquetes fueron verificados con el analizador “TS & BTS ESPE-ANALYZER” y la transmisión y recepción del video se ejecutó sin problema, dando como resultado la visualización del video subtulado con un texto claro y sincronizado.

Como última prueba se realizó una transmisión en tiempo real, en la cual mediante el software ESPE Player se logró transmitir un video que estaba siendo generado en el mismo momento mediante el programa de multiplexación en Matlab, siendo este el resultado esperado para esta prueba.

## 5 CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES

### 5.1 Conclusiones

Se realizó un sistema de transcripción de audio a texto mediante una aplicación en la nube que usa modelos de redes neuronales y algoritmos de aprendizaje profundo con gran eficiencia, el texto transcrito ingresa a un software que configura y sincroniza los subtítulos en paquetes PES que finalmente son multiplexados en un TS, con base en el estándar ISDB-T, para su posterior transmisión.

Se realizó un estudio del arte de las tecnologías actuales para la transmisión de televisión digital terrestre, investigando el estándar ISDB-T y su variación ISDBT-b junto con sus avances a nivel nacional e internacional.

Para la generación de las tablas PSI/SI y paquetes PES se utilizó las normas ABNT NBR15603-1 y ARIB STD-B24, utilizando la configuración para paquetes de datos con flujo sincronizado, en esta configuración se especifican los valores de los bytes que deben tener los paquetes de control y de texto para la transmisión de subtítulos.

Para la sincronización de subtítulos se utilizó la normativa ARIB STD-B37 que modifica la información de gestión del tiempo de los paquetes PES de control y texto para sincronizar el flujo datos de subtítulos con el de audio y video en un tiempo cuasi-real.

El texto a ser transmitido mediante los subtítulos se configuró con una longitud de 25 caracteres y características estándar en cuanto a tamaño, color, tipo de letra y espaciado, esto puede ser modificado en el software de acuerdo al set de caracteres de la norma ABNT NBR 15603-2.

Para verificar la correcta configuración de los paquetes del archivo TS usamos el analizador TS & BTS ESPE-ANALYZER que permite visualizar los bytes de cada paquete, indica las tablas PSI/SI y servicios existentes en el archivo y permite realizar una búsqueda por PID o número de paquete.

Para la transmisión se utilizó la tarjeta DTU-215 que es un modulador portátil fabricado por Dektec y mediante su software StreamXpress permite la reproducción de archivos TS y parámetros de modulación de control. Para la recepción se utiliza un decodificador marca Visiontec modelo VT7200 que va conectado al televisor digital.

Adicional se utilizó proyecto de tesis “DESARROLLO DE UNA PLATAFORMA DE CONTROL PARA LA TARJETA MODULADORA DE TELEVISIÓN DIGITAL DEKTEC DTU-215” y se realizó una modificación a su código para que permita la transmisión de un archivo .TS mientras este se está generando y permita una transmisión en vivo.

## **2.1 Recomendaciones**

Se recomienda que el dispositivo receptor final trabaje con el estándar ISDB-Tb para la correcta descomprensión de los paquetes enviados, además al momento de recibir el mensaje se debe activar la recepción de subtítulos en las configuraciones del decodificador o televisor digital.

Se recomienda al momento de realizar la traducción de audio a texto verificar el acceso a la API que se encuentre en uso y de ser posible ingresar al panel de control de la misma para comprobar que la traducción se haya realizado correctamente.

Si se requiere alguna modificación en la configuración de los subtítulos se recomienda verificar las normas y estándares que fueron base para la construcción de

los paquetes, así como la metodología usada para la generación de los mismos en el software de Matlab.

## BIBLIOGRAFÍA

- (Arib), a. O. (s.f.). *Data coding and transmission specifications for digital broadcasting, arib std-b24 version 5.0*.
- (nidcd), n. l. (2017). *Subtítulos para las personas sordas o con problemas de audición*.
- Abel suing, v. G. (2015). *Géneros y formatos para la televisión digital. Análisis en los países andinos*.
- Abnt. (2007). *Abnt nbr15603-1, televisión digital terrestre — multiplexación y servicios de información (si) parte 1: si del sistema de radiodifusión*.
- Arib. (2006). *Structure and operation of closed caption data conveyed by ancillary data packets. Arib std-b37 version 2.4*.
- Arib. (s.f.). *Reporte técnico de arib. Características del sistema isdb-t*.
- Dektec. (s.f.). Obtenido de <https://www.dektec.com/>
- Dektec. (2017). *Dtc-300-sp-streamxpress user manual*.
- Dektec. (s.f.). *Dtu-215*. Obtenido de <https://www.dektec.com/products/usb/dtu-215/>
- Ecuador, a. N. (2012). *Ley orgánica de discapacidades*.
- Figuroa, a. A. (2010). *Diseño de la red para interactividad en televisión digital terrestre e iptv en el campus espe sangolqui*.
- Google. (s.f.). *Google cloud*. Obtenido de <https://cloud.google.com/speech-to-text>
- Ibm. (s.f.). Obtenido de <https://www.ibm.com/watson/services/speech-to-text/>
- Ibm. (s.f.). *Watson speech to text*. Obtenido de <https://www.ibm.com/cloud/watson-speech-to-text>
- Iso/iec. (2000). *Iso/iec 13818-1. Information technology: generic coding of moving pictures and associated audio information: systems*.
- M, p., t, m., i, b., & d., m. (2015). *Desarrollo de un analizador de flujos únicos de paquetes de transporte ts conformados de acuerdo a la norma de tv digital isdb-tb*.
- Maseo, a. T. (2015). *Estudio descriptivo y experimental de la subtitulación en tv para niños sordos*.
- Mathworks. (2020). Obtenido de <https://www.mathworks.com/matlabcentral/fileexchange/65266-speech2text>
- Mena bastidas, r. J. (2015). *Desarrollo de una plataforma de control para la tarjeta moduladora de televisión digital dektec dtu-215*.

- Microchip. (2000). *Crc generating and checking an730*.
- Microsoft azure. (s.f.). Obtenido de <https://azure.microsoft.com/es-es/services/cognitive-services/speech-to-text/>
- Microsoft. (s.f.). *Microsoft azure*. Obtenido de <https://azure.microsoft.com/en-us/services/cognitive-services/speech-to-text/>
- Mintel. (2018). *Libro blanco de la sociedad de la información y del conocimiento*.
- Mintel. (2018). *Plan maestro de transición a la televisión diital terrestre*.
- Obregón, j. A. (s.f.). *Implementación de un laboratorio de pruebas de tv digital terrestre isdb-tb, para la emisión de aplicaciones interactivas para la escuela de ingeniería electrónica y telecomunicaciones de la unach*.
- Pisciotta, n. O. (2010). *Sistema isdb-t*.
- Rafael sotelo, d. D. (2011). *Sistema de transmisión isdb-t*.
- Ramírez, e. B. (2008). *La televisión digital terrestre en españa. Por un sistema televisivo de futuro acorde con una democracia de calidad*.
- Recalde vargas, c. E. (2016). *Análisis y generación del flujo de transporte con sobreimposición de texto para alerta temprana en servicios de televisión digital terrestre*.
- T, m., m, p., d, m., & i., b. (s.f.). *Generador de flujos únicos de paquetes de transporte ts en base a la norma isdb-tb*.
- Torrijos, a. L. (2012). *Sistema para el alineamiento de subtítulos y audio en escenarios de reablado en tv*.
- Valdivia, b. D. (2015). *Analisis del estandar isdb-t para ser utilizado en una posible implementacion de la television digital terrestre en nicaragua*.
- Visualfy. (s.f.). Obtenido de <https://www.visualfy.com/es/que-es-estenotipia-como-ayuda-personas-sordas/>