



REPÚBLICA DEL ECUADOR

Escuela Politécnica Nacional

" E SCIENTIA HOMINIS SALUS "

La versión digital de esta tesis está protegida por la Ley de Derechos de Autor del Ecuador.

Los derechos de autor han sido entregados a la "ESCUELA POLITÉCNICA NACIONAL" bajo el libre consentimiento del (los) autor(es).

Al consultar esta tesis deberá acatar con las disposiciones de la Ley y las siguientes condiciones de uso:

- Cualquier uso que haga de estos documentos o imágenes deben ser sólo para efectos de investigación o estudio académico, y usted no puede ponerlos a disposición de otra persona.
- Usted deberá reconocer el derecho del autor a ser identificado y citado como el autor de esta tesis.
- No se podrá obtener ningún beneficio comercial y las obras derivadas tienen que estar bajo los mismos términos de licencia que el trabajo original.

El Libre Acceso a la información, promueve el reconocimiento de la originalidad de las ideas de los demás, respetando las normas de presentación y de citación de autores con el fin de no incurrir en actos ilegítimos de copiar y hacer pasar como propias las creaciones de terceras personas.

Respeto hacia sí mismo y hacia los demás.

ESCUELA POLITÉCNICA NACIONAL

**FACULTAD DE INGENIERÍA ELÉCTRICA Y
ELECTRÓNICA**

**ESTUDIO DE LAS TÉCNICAS DE VISIÓN ESTEREOSCÓPICA
PARA EL ESTÁNDAR ISDB-T DE TELEVISIÓN DIGITAL
TERRESTRE**

**PROYECTO PREVIO A LA OBTENCIÓN DEL TÍTULO DE INGENIERO EN
ELECTRÓNICA Y TELECOMUNICACIONES**

AUGUSTO WLADIMIR DE LA CADENA RAMOS

wladimir_dlc@yahoo.com

DIRECTORA: ING. TANIA PÉREZ RAMOS

tania.perez@epn.edu.ec

Quito, Mayo 2012

DECLARACIÓN

Yo, Augusto Wladimir De La Cadena Ramos, declaro bajo juramento que el trabajo aquí descrito es de mi autoría; que no ha sido previamente presentada para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración cedo mis derechos de propiedad intelectual correspondientes a este trabajo, a la Escuela Politécnica Nacional, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.

Augusto Wladimir De La Cadena Ramos

CERTIFICACIÓN

Certifico que el presente trabajo fue desarrollado por Augusto Wladimir De La Cadena Ramos bajo mi supervisión

Ing. Tania Pérez

DIRECTORA DEL PROYECTO

AGRADECIMIENTO

Agradezco profundamente a mi Madre, a mi Padre y a mis hermanos por el apoyo constante a lo largo de toda la vida, ya que gracias a su amor, cariño y comprensión hoy logro culminar con una etapa trascendental para mí.

Agradezco a la Ing. Tania Pérez, no solo por su acertada guía y apoyo para realizar el presente proyecto, sino también por los conocimientos impartidos que me permitirán ser un gran profesional.

A la Escuela Politécnica Nacional, por brindarme una educación integral para la vida futura.

A mis profesores por compartir sus valiosos conocimientos y experiencia, para hacer de mí un profesional útil a la sociedad.

A toda mi familia por acompañarme y estar presente en los momentos importantes de mi vida.

Y a todos mis amigos por la compañía y los momentos alegres compartidos.

DEDICATORIA

A MIS PADRES MISBELL Y AUGUSTO

POR DARME LA VIDA Y APOYARME EN TODO MOMENTO;

A MIS HERMANOS CÉSAR Y WILLY,

POR SU CARIÑO Y POR SIEMPRE ESTAR A MI LADO.

A MIS ABUELOS,

POR SIEMPRE TENERME PRESENTE Y DEDICARME SU AMOR.

A TODOS ELLOS GRACIAS POR ACOMPAÑARME

Y SER PARTE DE ESTE LOGRO TAN IMPORTANTE.

CONTENIDO

CAPÍTULO 1.....	1
INTRODUCCIÓN Y DEFINICIONES	1
1.1 INTRODUCCIÓN.....	1
1.2 HISTORIA DE LA ESTEREOSCOPIA.....	3
1.3 FUNCIONAMIENTO DEL SISTEMA VISUAL HUMANO PARA CAPTAR IMÁGENES EN TRES DIMENSIONES	6
1.3.1 DEFINICIÓN DE ESTEREOSCOPIA.....	6
1.3.1.1 Estereoscopia.....	6
1.3.1.2 Hiperestereoscopia	6
1.3.1.3 Hipoestereoscopia.....	7
1.3.2 PERCEPCIÓN DE PROFUNDIDAD	7
1.3.2.1 Superposición.....	7
1.3.2.2 Perspectiva o puntos de fuga.....	7
1.3.2.3 Tamaño de objetos conocidos	8
1.3.2.4 Tamaño de la imagen sobre la retina.....	8
1.3.2.5 Estereopsis	8
1.3.3 DESCRIPCIÓN DEL OJO HUMANO ANTE LA PROFUNDIDAD	9
1.3.3.1 Descripción del ojo humano	9
1.3.3.2 Campo de Visión Humana	10
1.3.3.3 Convergencia Visual.....	11
1.3.3.4 Enfoque o acomodación Visual	12
1.3.3.5 Agudeza estereoscópica	12
1.3.3.6 Disparidad binocular.....	12
1.3.3.7 Definición de paralaje (Parallax)	12
1.3.3.8 El efecto Pulfrich.....	14
1.4 TÉCNICAS Y TECNOLOGÍAS PARA LA VISUALIZACIÓN DE IMÁGENES ESTEREOSCÓPICAS.....	15
1.4.1 GAFAS 3D (MINI PANTALLAS).....	16
1.4.2 PANTALLAS SEPARADAS Y GAFAS PRISMÁTICAS	16
1.4.3 TÉCNICA ANAGLÍFICA	17
1.4.3.1 Anáglifo cromático	17
1.4.3.2 Anáglifo por polarización o vectógrafo	18
1.4.3.2.1 Polarización de la luz.....	18
1.4.3.3 Anáglifo lenticular integrado (autoestereoscopia).....	19
1.4.3.4 Multiplexación Temporal.....	20
1.5 ESTÁNDAR DE TELEVISIÓN DIGITAL ISDB-T	20
1.5.1 CARACTERÍSTICAS DE LA TELEVISIÓN DIGITAL.....	21
1.5.1.1 Digitalización	22
1.5.2 ESPECIFICACIONES TÉCNICAS DE ISDB-T	23
1.5.2.1 Compresión de audio y video.....	24
1.5.2.2 Transmisión de la señal en ISDB-T.....	25
1.5.2.2.1 Modulación QPSK (Quadrature Phase Shift Keying).....	25
1.5.2.2.2 Modulación 16 QAM (Quadrature Amplitude Modulation), 64 QAM	25
1.5.2.2.3 Modulación DPSK (Differential Phase Shift Keying).....	26
1.6 ANÁLISIS DE LOS REQUERIMIENTOS DE UN SISTEMA DE TELEVISIÓN ESTEREOSCÓPICA.....	29
1.6.1 CONSIDERACIONES PARA LA EVALUACIÓN SUBJETIVA DE IMÁGENES.....	29

1.6.1.1	Factores de evaluación de la televisión monoscópica	29
1.6.1.2	Factores de evaluación de la televisión estereoscópica	30
1.6.1.3	Métodos de evaluación.....	31
1.6.1.3.1	Método de escala de calidad continua de doble estímulo DSCQS	31
1.6.1.3.2	Resultados experimentales del método DSCQS en imágenes estereoscópicas 32	
1.6.1.4	Condiciones de visualización	33
1.6.2	CONSIDERACIONES DE IMÁGENES DE TV ESTEREOCÓPICAS BASADAS EN SEÑALES L+R	34
1.6.3	REQUERIMIENTOS DE COMPATIBILIDAD Y VISUALIZACIÓN PARA UN SISTEMA DE TV ESTEREOCÓPICO	35
CAPÍTULO 2.....		38
ADQUISICIÓN DE IMÁGENES ESTEREOCÓPICAS Y TRANSMISIÓN DE LAS SEÑALES DE VIDEO 3D PARA EL ESTÁNDAR ISDB-T		38
2.1.	ADQUISICIÓN / GENERACIÓN DE CONTENIDOS 3D.....	38
2.1.1	CONVERSIÓN OPTO-ELÉCTRICA DE LA ESCENA	39
2.1.2	ADQUISICIÓN A TRAVÉS DE 2 LENTES (CÁMARA ESTEREOCÓPICA)	41
2.1.2.1	Aspectos geométricos y calibración de los lentes de las cámaras estereoscópicas.....	42
2.1.2.1.1	Configuración en paralelo.....	42
2.1.2.1.2	Configuración Toed-in (ejes cruzados)	44
2.1.2.1.3	Percepción cuantitativa de la profundidad de la cámara estereoscópica	47
2.1.3	ADQUISICIÓN A TRAVÉS DE MÚLTIPLES CÁMARAS (SISTEMA MULTIVISTA) ...	48
2.1.3.1	Interpolación Multivista	50
2.1.3.2	Reconstrucción 3D	51
2.1.4	ADQUISICIÓN A TRAVÉS CÁMARA 2D + Z-CAM (TOF)	51
2.1.5	ADQUISICIÓN A TRAVÉS CÁMARA 2D (Conversión 2D-3D)	53
2.1.5.1	Descripción del algoritmo de conversión.....	54
2.1.6	ADQUISICIÓN CON SISTEMA GENÉRICO PARA PRODUCCIÓN 3D	55
2.2	FORMATOS DE VIDEO PARA 3DTV.....	57
2.2.1	FORMATO ANÁGLIFO CROMÁTICO OPTIMIZADO (OCA)	59
2.2.2	PROCESAMIENTO INDEPENDIENTE DE LAS SEÑALES DE VIDEO (CSV).....	60
2.2.2.1	Concepto de Dual Stream	61
2.2.2.2	Transmisión Simultánea (Simulcast).....	62
2.2.3	FORMATO COMPATIBLE CON EL CUADRO HD (CFC)	63
2.2.3.1	Formato de Empaquetamiento de Cuadro	64
2.2.3.1.1	Empaquetamiento de cuadro para formato progresivo	64
2.2.3.1.2	Empaquetamiento de cuadro para formato entrelazado	65
2.2.3.2	Configuración Side by Side Half (SbS)	66
2.2.3.3	Configuración Top and Bottom (TaB).....	67
2.2.3.4	Configuración quincunx matrix (Tablero de ajedrez)	68
2.2.3.5	Configuración campos alternados	68
2.2.3.6	Configuración líneas alternadas	69
2.2.3.7	Configuración Side by Side (Full).....	69
2.2.3.8	Configuración "L + Profundidad" (Video L + Depth).....	71
2.2.3.9	Configuración L + D + Gfx + Gfx-L (Left + Depth + Graphics + Graphics – Left)	71
2.2.4	FORMATO VIDEO + PROFUNDIDAD (V+D)	72
2.2.4.1	El mapa de profundidad	73
2.2.4.2	Generación del mapa de profundidad a partir de 2D	74
2.2.4.3	Generación del mapa de profundidad a través de Cámara infrarroja.....	75

2.2.4.4	Generación del mapa de profundidad a través de Cámara Estéreo.....	79
2.2.4.4.1	Calibración de los planos de la Imagen.	80
2.2.4.4.2	Cálculo del mapa de profundidad.....	82
2.2.5	FORMATO MULTI-VIEW VIDEO CODING (MVC).....	83
2.2.6	FORMATO MVD (MVC + D)	85
2.2.7	FORMATO LAYERED DEPTH VIDEO (LDV).....	87
2.2.8	FORMATO DEPTH ENHANCED VIDEO (DES).....	90
2.3	PROCESAMIENTO DE LAS SEÑALES DE VIDEO ADQUIRIDAS.....	91
2.3.1	PROCESAMIENTO DE L+R HACIA ANÁGLIFO CROMÁTICO OPTIMIZADO	93
2.3.2	PROCESAMIENTO DE MEJORAMIENTO DE IMAGEN, TÉCNICAS CONVENCIONALES EN VIDEO MONOSCÓPICO.....	93
2.3.3	PROCESAMIENTO DE ESTIMACIÓN DE PROFUNDIDAD, INTERPOLACIÓN DE VISTAS, CREACIÓN DE MAPAS.	94
2.4	CODIFICACIÓN Y COMPRESIÓN DE LAS SEÑALES DE VIDEO ESTEREOSCÓPICO.....	94
2.4.1	MPEG-4 APLICABLE EN OCA, CSV Y CFC+	96
2.4.1.1	Características generales del estándar H.264, MPEG-4 Parte 10.	96
2.4.1.1.1	Elementos comunes en todos los perfiles.	98
2.4.1.1.2	Perfil de línea Base	99
2.4.1.1.3	Perfil Principal.....	99
2.4.1.1.4	Perfil Extendido	99
2.4.1.1.5	Perfiles Altos (High-Profiles).....	100
2.4.1.1.6	Estructura por capas	100
2.4.1.1.7	Algoritmo de codificación.....	100
2.4.1.1.8	Predicción Intra-Frame	101
2.4.1.1.9	Predicción Inter-Frame	102
2.4.1.1.10	Transformación y cuantificación	102
2.4.1.1.11	Proceso de codificación	103
2.4.1.1.12	Codificación de entropía.....	104
2.4.1.1.13	Tasa de transmisión.	105
2.4.1.2	Aplicación del H.264-AVC, MPEG-4 parte 10 a la TV3D.....	106
2.4.1.2.1	H.264 AVC en OCA (Optimized Chromatic Anaglyph)	106
2.4.1.2.2	H.264 AVC para CSV (Conventional Video Stereo).	107
2.4.1.2.3	H.264 AVC para CFC+ (Conventional HD Frame Compatible +)	108
2.4.2	CODIFICACIÓN PARA EL FORMATO V+D.....	109
2.4.2.1	MPEG-C parte 3	110
2.4.2.2	H.264 - Imágenes auxiliares.....	111
2.4.3	CODIFICACIÓN PARA EL FORMATO MVC.....	112
2.4.4	CODIFICACIÓN PARA EL FORMATO MVD	114
2.4.5	CODIFICACIÓN PARA EL FORMATO LDV	116
2.4.6	CODIFICACIÓN PARA EL FORMATO DES.....	117
2.4.7	TASAS DE BITS RESULTANTES DE LA CODIFICACIÓN.....	117
2.4.7.1	Tasa de bits para OCA.....	117
2.4.7.2	Tasa de bits para CSV	118
2.4.7.3	Tasa de bits para CFC+	118
2.4.7.4	Tasa de bits para V+D.....	119
2.4.7.5	Tasa de bits para MVC.....	122
2.4.7.6	Tasa de bits para MVD.....	124
2.4.7.7	Tasa de bits para LDV.....	125
2.4.7.8	Tasa de bits para DES	125
2.5	TRANSMISIÓN DE LAS SEÑALES DE TV-3D EN EL ESTÁNDAR ISDB-T.	126
2.5.1	GENERACIÓN DEL TRANSPORT STREAM Y MULTIPLEXACIÓN.....	126
2.5.1.1	Empaquetamiento y generación del TS con MPEG-2	126

2.5.1.2	Multiplexación MPEG-2 para señales TV3D que contienen múltiples informaciones.....	128
2.5.2	CODIFICACIÓN MODULACIÓN Y TRANSMISIÓN DE LA SEÑAL DE TV-3D.....	130
2.5.2.1	Características y capacidad de los 13 segmentos dentro de los 6 MHz.....	130
2.5.2.2	Señales de TV-3D en los segmentos OFDM.....	132
2.5.2.3	Señales de TV-3D en el segmento destinado One-seg.....	134
2.5.2.4	Transmisión de la señal de video.....	135
2.6	EFFECTO DEL SERV. 3DTV EN LAS CARACTERÍSTICAS TÉCNICAS DE ISDB-T.	135
CAPÍTULO 3.....		137
RECEPCIÓN DE LA SEÑAL DE VIDEO 3D PARA EL ESTÁNDAR ISDB-T.		137
4.1	DESCRIPCIÓN GENERAL DEL SISTEMA DE RECEPCIÓN PARA VIDEO ESTEREOSCÓPICO.....	137
6.1.1	DESCRIPCIÓN DEL IRD (Integrated Receiver-Decoder).....	139
6.2	PROCESO DE RECEPCIÓN DEL SERVICIO DE TELEVISIÓN ESTEREOSCÓPICA PARA ISDB-T.....	140
6.2.1	RECEPCIÓN DE LA SEÑAL ANÁGLIFA CROMÁTICA.....	140
6.2.2	RECEPCIÓN DE LA SEÑAL CSV.....	140
6.2.2.1	Recepción y desempaquetamiento del Dual-Stream.....	140
6.2.2.2	Recepción y desempaquetamiento del Simulcast.....	141
6.2.3	RECEPCIÓN DE LA SEÑAL PARA CFC+.....	141
6.2.4	RECEPCIÓN DE LA SEÑAL PARA V+D.....	141
6.2.4.1	Recepción para MPEG-C Parte 3.....	141
6.2.4.2	Recepción para Video auxiliar.....	142
6.2.5	RX. DE LA SEÑAL PARA FORMATOS AVANZADOS (MVC, MVD, LDV, DES).....	142
6.3	DESCRIPCIÓN DEL SET-TOP-BOX PARA 3DTV.....	142
6.3.1	COMUNICACIÓN E INTERFACES ENTRE EL SET-TOP-BOX Y EL TELEVISOR.....	145
6.4	REPRODUCCIÓN DEL VIDEO ESTEREOSCÓPICO.....	147
6.4.1	PANTALLAS ESTEREOSCÓPICAS.....	148
6.4.1.1	Pantallas Estereoscópicas Anaglíficas.....	148
6.4.1.2	Pantallas Estereoscópicas Polarizadas.....	149
6.4.1.3	Pantallas Estereoscópicas Basadas en la Multiplexación Temporal.....	151
6.4.1.3.1	Funcionamiento y estructura de los LCS (Liquid Crystal Shutter).....	152
6.4.2	PANTALLAS AUTO-ESTEREOSCÓPICAS.....	153
6.4.2.1	Pantallas Auto-Estereoscópicas de dos vistas.....	155
6.4.2.2	Pantallas Auto-Estereoscópicas con seguimiento de la cabeza.....	157
6.4.2.3	Pantallas Auto-Estereoscópicas multivista.....	159
6.4.2.4	Tecnologías para la elaboración de pantallas Auto-Estereoscópicas.....	159
3.2.4.2.1	Tecnología de multiplexación espacial.....	160
3.2.4.2.2	Tecnología Multiproyector.....	160
3.2.4.2.3	Tecnología secuencial en tiempo.....	161
6.4.3	PANTALLAS VOLUMÉTRICAS Y HOLOGRÁFICAS.....	162
6.4.3.1	Pantallas Volumétricas.....	162
6.4.3.2	Pantallas Holográficas.....	163
6.4.3.3	Infraestructura de las pantallas volumétricas.....	164
6.4.3.4	Algunos prototipos de pantallas volumétricas y holográficas existentes.....	165
CAPÍTULO 4.....		167
SITUACIÓN ACTUAL DE LA TV-3D Y TEMAS RELACIONADOS.....		167
4.1	INVESTIGACIÓN ACTUAL Y PROYECTOS VIGENTES.....	167

4.1.1	ORGANIZACIÓN “SD&A”	168
4.1.2	PROYECTO “3D VIVANT”	170
4.1.3	PROYECTO “3D4YOU”	172
4.1.4	PROYECTO “MOBILE 3DTV”	173
4.1.5	PROYECTO “TV3D 2.0”	173
4.1.6	AGENTES INVESTIGADORES Y DESARROLLADORES PARA LA TV3D	174
4.1.6.1	Organismos de estandarización.	174
4.1.6.2	Investigadores académicos y centros de investigación.	175
4.2	SITUACIÓN TÉCNICA ACTUAL DE LA TELEVISIÓN ESTEREOSCÓPICA.....	177
4.2.1	DESARROLLO EXPERIMENTAL Y PRUEBAS DE CAMPO REALIZADAS.	177
4.2.1.1	Pruebas y experimentos en la etapa de adquisición del video 3D.	177
4.2.1.2	Pruebas y experimentos en la etapa de procesamiento y codificación.	179
4.2.1.3	Pruebas y experimentos en la etapa de Transport Stream, Multiplexación y Transmisión y recepción de la señal.	180
4.2.1.4	Pruebas “Codificación y transmisión de 1080p 3D en 6 segmentos ISDB-T internacional.”	182
4.2.1.5	Pruebas y experimentos en la etapa de reproducción de la señal.	183
4.2.1.6	Pruebas de TV3D móvil.....	183
4.2.1.7	Pruebas de percepción de Imagen y experiencia del usuario.	185
4.2.2	EQUIPOS EXISTENTES PARA CONFORMAR EL SISTEMA DE TV3D.	186
4.2.2.1	Equipos para la adquisición.	186
4.2.2.1.1	Cámaras estereoscópicas	186
4.2.2.1.2	Cámaras de profundidad Z-Cam – TOF.....	187
4.2.2.1.3	Cámaras para sistema multivista.	187
4.2.2.1.4	Equipos para el procesamiento, codificación, transporte y transmisión de la Señal.	188
4.2.2.1.5	Equipos para la recepción de la señal.	188
4.2.3	PERSPECTIVAS DE NORMALIZACIÓN, EVOLUCIÓN DE LOS ELEMENTOS TÉCNICOS DEL SISTEMA 3DTV.....	190
4.3	SITUACIÓN DE MERCADO DEL SISTEMA Y SUS TENDENCIAS.....	193
4.4	RELACIÓN CON OTRAS TECNOLOGÍAS Y PERSPECTIVAS FUTURAS.	197
CAPÍTULO 5		199
CONCLUSIONES Y RECOMENDACIONES		199
5.1	CONCLUSIONES.....	199
5.2	RECOMENDACIONES	202
REFERENCIAS BIBLIOGRÁFICAS.....		204

ANEXOS

Anexo 1.-Extracto de la Especificación High-Definition Multimedia Interface Versión 1.4a (HDMI 1.4a) (porción de señalización para video 3D – Formato Digital)

Anexo 2.- Detalle de la codificación para los formatos LDV (Layered Depth Video).

Anexo 3.- Métodos de configuración y calibración de las cámaras para producción de Televisión estereoscópica (3DTV) (Formato Digital).

Anexo 4.- Extracto de algunos resultados de pruebas experimentales relacionadas a la 3DTV.

Anexo 5.- Extracto del estudio de la experiencia visual del usuario de señales de televisión estereoscópica (3DTV) al usuario (Estudio completo - Formato Digital)

Anexo 6.- Características de algunos equipos para televisión estereoscópica (3DTV) (Formato Digital)

ÍNDICE DE FIGURAS

CAPÍTULO 1

Figura 1. 1 Trayectoria Visual y descripción del ojo humano	10
Figura 1. 2 Campo visual humano para 2D y 3D.	11
Figura 1. 3 Convergencia visual de objetos A y B a diferente distancia.	11
Figura 1. 4 Ángulos de paralaje para objetos en diferentes posiciones.	13
Figura 1. 5 Geometría de las distancias en la convergencia.	14
Figura 1. 6 Naturaleza de la luz polarizada.....	19
Figura 1. 7 Sistema de Transmisión de Televisión Digital	21
Figura 1. 8 Espectro de la modulación OFDM	25
Figura 1. 9 Características de los esquemas de modulación posibles	26
Figura 1. 10 Características de las tasas de corrección de errores posibles.	26
Figura 1. 11 Distribución segmentada del canal.....	27
Figura 1. 12 Ejemplo de uso de los 13 segmentos del canal.	27
Figura 1. 13 Resultados del estudio ilustrativo utilizando el método DSCQS	33

CAPÍTULO 2

Figura 2. 1 Cadena de transmisión del video estéreo.	38
Figura 2. 2 Esquema del detector CCD	40
Figura 2. 3 Trayectoria de la luz en un lente de cámara.....	41
Figura 2. 4 Configuración en paralelo	43
Figura 2. 5 Aspectos geométricos de la configuración paralela	43
Figura 2. 6 Configuración Toed-In.....	45
Figura 2. 7 Aspectos geométricos de la configuración Toed-In.....	45
Figura 2. 8 Etapas multivista para un sistema 3DTV	49
Figura 2. 9 Vistas disponibles en la interpolación Multivista.....	50
Figura 2. 10 Cuantificación del mapa de profundidad extraído por la Z-Cam.	52
Figura 2. 11 Algoritmo de conversión 2D-3D.....	54
Figura 2. 12 Adquisición genérica de video 3D.....	55
Figura 2. 13 Cuadros resultantes de la configuración genérica.....	56
Figura 2. 14 Resumen del sistema anáglifo en ISDB-T.....	59
Figura 2. 15 Imagen Anáglifo de alta calidad.....	60
Figura 2. 16 Descripción del Dual-Streaming para TV3D.....	62
Figura 2. 17 Descripción del Simulcast para TV3D.	63
Figura 2. 18 Empaquetamiento de cuadro para formato progresivo.	65
Figura 2. 19 Empaquetamiento de cuadro para formato entrelazado.	66
Figura 2. 20 Empaquetamiento de cuadro para formato entrelazado.	67
Figura 2. 21 Configuración Top and Bottom (TaB).	67
Figura 2. 22 Configuración Tablero de ajedrez.....	68
Figura 2. 23 Configuración de Campos alternados.....	69
Figura 2. 24 Configuración Líneas alternadas	70
Figura 2. 25 Configuración SbS Full	70
Figura 2. 26 Configuración L + Profundidad	71
Figura 2. 27 Configuración L + D + Gfx + Gfx-L	72
Figura 2. 28 Descripción del mapa de profundidad	74
Figura 2. 29 Generación del mapa de profundidad a partir de 2D.....	75

Figura 2. 30 Estructura de una Cámara de Profundidad (Z-Cam).....	76
Figura 2. 31 Funcionamiento de las cámaras infrarrojas.....	76
Figura 2. 32 Adquisición de una imagen con la Z-Cam.....	77
Figura 2. 33 Generación del mapa de profundidad con cámara estéreo.....	80
Figura 2. 34 Generación del mapa de profundidad con cámara estéreo.....	81
Figura 2. 35 Puntos de los pixeles esquina.....	81
Figura 2. 36 Rectificación de rotación de un pixel.....	82
Figura 2. 37 Formato de 3D MVC (Perfil Multi-vista).....	84
Figura 2. 38 Formato de 3D MVD (Perfil Multi-vista + Profundidad).....	86
Figura 2. 39 Interpolación de vistas virtuales en MVD.....	87
Figura 2. 40 Información obtenida de LDV.....	88
Figura 2. 41 Vistas originales adquiridas en LDV.....	89
Figura 2. 42 Deformación de las vistas originales adquiridas en LDV.....	89
Figura 2. 43 Procesamiento de las vistas originales.....	89
Figura 2. 44 Información en DES para n=2 vistas.....	90
Figura 2. 45 Conversión L+R hacia anáglifo cromático.....	93
Figura 2. 46 Diagrama de bloques de la codificación H.264.....	101
Figura 2. 47 Procesos de CABAC.....	105
Figura 2. 48 Codificación para OCA.....	107
Figura 2. 49 Aplicación de H.264 en Dual-Streaming.....	107
Figura 2. 50 Aplicación de H.264 en Simulcast.....	107
Figura 2. 51 Aplicación de H.264 en CFC+.....	109
Figura 2. 52 Cadena de Transmisión para MPEG-C parte 3.....	110
Figura 2. 53 Aplicación del codificador H.264 a las señales de imágenes auxiliares.....	111
Figura 2. 54 Predicción Multivista para n=5 y GOP=8 (IPP).....	112
Figura 2. 55 Predicción Multivista para n=5 y GOP=8.....	113
Figura 2. 56 Codificador MVC de n vistas.....	114
Figura 2. 57 Cadena de transmisión en MVD.....	115
Figura 2. 58 Modo de predicción de imágenes para MVD.....	115
Figura 2. 59 Diagrama de bloques de un codificador MVD.....	116
Figura 2. 60 Tasa de bits en SDTV.....	118
Figura 2. 61 Tasa de bits en 720p.....	119
Figura 2. 62 Tasa de bits en 1080x1920.....	119
Figura 2. 63 Tasa de bits al codificar el mapa de profundidad.....	120
Figura 2. 64 Tasa de bits al codificar el mapa de profundidad.....	121
Figura 2. 65 Tasa de bits al codificar con MVC y n=2.....	122
Figura 2. 66 Tasa de bits al codificar con MVC y n=8.....	124
Figura 2. 67 Tasa de bits al codificar con MVD.....	124
Figura 2. 68 Tasa de bits al codificar con LDV.....	125
Figura 2. 69 Cabecera del Transport-Stream en MPEG-2.....	127
Figura 2. 70 Generación del Transport-Stream.....	127
Figura 2. 71 Multiplexación de vistas en dual Streaming.....	128
Figura 2. 72 Multiplexación de vistas izquierda + profundidad.....	129
Figura 2. 73 Multiplexación de canales 2D y 3D con un solo Multiplexores.....	129
Figura 2. 74 Multiplexación de canales 2D y 3D con varios Multiplexores.....	130

CAPÍTULO 3

Figura 3. 1 Infraestructura de recepción de la Televisión Digital Terrestre.....	138
Figura 3. 2 Estructura Interna de un Set-Top-Box.....	139
Figura 3. 3 Receptor IRD MPEG-C Parte 3.....	142

Figura 3. 4 Sintetización de vistas virtuales en un Set-Top-Box 3D.....	144
Figura 3. 5 Estructura de un Set-Top-Box 3D.....	145
Figura 3. 6 Interfaz para formatos de n=2 vistas.	146
Figura 3. 7 Interfaz para formatos de n>2 vistas.	146
Figura 3. 8 Efecto de la pantalla polarizadora.....	149
Figura 3. 9 Polarización en 2 dimensiones.	150
Figura 3. 10 Pantalla de ángulo de polarización variable.	150
Figura 3. 11 Pantalla de multiplexación temporal.	151
Figura 3. 12 Protocolo LCS – DLP Link.	153
Figura 3. 13 Discretización del espacio visual	154
Figura 3. 14 Métodos para direccionar las imágenes a cada ojo.	155
Figura 3. 15 Zonas de visión para n=2 vistas.	156
Figura 3. 16 Opciones de movimiento de la zona de visión.	157
Figura 3. 17 Métodos de movimiento de la zona de visión Xenotech y Sharp.	158
Figura 3. 18 Zonas de visión en pantallas autoestereoscópicas multivista.	159
Figura 3. 19 Funcionamiento de la tecnología multi-proyector.	161
Figura 3. 20 Funcionamiento de la tecnología secuencial en tiempo.	161

CAPÍTULO 4

Figura 4. 1 Prototipo de sistema Volumétrico.	169
Figura 4. 2 Arreglo de cámaras (100) en sistema Multivista.....	178
Figura 4. 3 Resultados experimentales de codificación MVC.....	179
Figura 4. 4 Condiciones de Prueba de la TV3D en algunos sistemas.....	180
Figura 4. 5 Laboratorio de Pruebas de la TV3D.	181
Figura 4. 6 Sistema de Pruebas de TV3D sobre ISDB-T	182
Figura 4. 7 Representación “Space-Ray” de un sistema de 100 vistas.....	183
Figura 4. 8 Descripción del sistema 3DTV móvil en DMB	184
Figura 4. 9 Parámetros de Prueba de 3DTV sobre DMB	184
Figura 4. 10 Generaciones del servicio 3DTV	190
Figura 4. 11 Fase de evolución de la 3DTV	191
Figura 4. 12 Comparación de los esquemas de codificación H.264.....	192
Figura 4. 13 Relación entre el número de vistas y el servicio brindado.....	193
Figura 4. 14 Unidades de Televisores 3D vendidas en 2010	196
Figura 4. 15 Tendencia de consumo de televisores 3D.....	197
Figura 4. 16 Servicios brindados por la TV a lo largo de la Historia	197

ÍNDICE DE TABLAS

CAPÍTULO 1

Tabla 1. 1 Resumen histórico de la estereoscopía.....	5
Tabla 1. 2 Valores y relaciones de luminancia de colores complementarios utilizados para la representación anaglífica	17
Tabla 1. 3 Características de los esquemas de muestreo.....	23
Tabla 1. 4 Características de transmisión de ISDB-T.....	28

CAPÍTULO 2

Tabla 2. 1 Representación de la profundidad por configuración	48
Tabla 2. 2 Formatos de 3DTV.....	58
Tabla 2. 3 Características técnicas de las cámaras infrarrojas.....	77
Tabla 2. 4 Pre procesamiento requerido por cada formato.....	92
Tabla 2. 5 Esquema de codificación según el formato.....	95
Tabla 2. 6 Evolución del estándar H.264	97
Tabla 2. 7 Tasas de transmisión en ISDB-T.....	131
Tabla 2. 8 Tasas de transmisión en ISDB-T para los formatos de TV3D.....	133
Tabla 2. 9 Características del estándar ISDB-T para el servicio 3DTV	136

CAPÍTULO 3

Tabla 3. 1 Prototipos de pantallas holográficas existentes.....	165
---	-----

CAPÍTULO 4

Tabla 4. 1 Miembros y funciones del 3D Vivant.....	171
Tabla 4. 2 Agentes Investigadores y aportaciones significativas.....	176
Tabla 4. 3 Modelos de cámaras estereoscópicas disponibles en el mercado.....	187
Tabla 4. 4 Modelos de cámaras de profundidad Z-CAM disponibles en el mercado.....	187
Tabla 4. 5 Equipos de transmisión disponibles en el mercado.....	188
Tabla 4. 6 Equipos de recepción 3DTV disponibles en el mercado.....	189
Tabla 4. 7 Oferta de Televisores 3D en el mercado	195

RESUMEN

El presente proyecto describe y detalla las técnicas y procedimientos para transmitir en el sistema de televisión, señales de naturaleza estereoscópica o tridimensional. Se abarcan todos los elementos que configuran el sistema de 3DTV, partiendo desde la adquisición y finalizando en la recepción y percepción del usuario.

En el primer capítulo, se exponen los conceptos teóricos necesarios para entender el fenómeno físico de la estereoscopía, así como los parámetros técnicos relevantes de la televisión digital, y en especial del estándar ISDB-T internacional.

En el segundo capítulo, se describen técnicamente los procesos en la cadena de transmisión del video estereoscópico. En cada etapa, se detallan los métodos existentes, las alternativas técnicas, los formatos de producción 3D y los parámetros técnicos de importancia. Se parte desde la adquisición, con el funcionamiento de los dispositivos transductores, hasta llegar finalmente la etapa de transmisión; haciendo hincapié en la aplicación del servicio sobre el estándar ISDB-T internacional.

El tercer capítulo trata sobre la etapa de recepción de las señales de 3DTV, describiendo los equipos para recibir y desempaquetar la información recibida. Posteriormente se trata el funcionamiento de las pantallas reproductoras, describiendo las tecnologías y alternativas existentes para todos los formatos posibles.

El cuarto capítulo comprende el análisis de los proyectos en marcha y estudios relacionados la 3DTV, se trata la proyección del servicio y la experiencia técnica del mismo en los lugares donde se ha implementado. Se analiza el impacto que puede tener el servicio, y como este puede evolucionar con el tiempo.

Por último se exponen las conclusiones y recomendaciones referentes al estudio realizado.

PRESENTACIÓN

La televisión constituye el medio de comunicación masivo más usado por el ser humano; a pesar de la evolución de servicios como Internet y telefonía móvil, las prestaciones y características de la televisión, provocan que este servicio, en lugar de ser reemplazado, sea susceptible de evolución en servicios y contenidos.

El presente proyecto se presenta como una guía para establecer los lineamientos y directrices, que desemboquen en la implementación del servicio de televisión estereoscópica. Dentro de todo lo tratado en el proyecto, se pretende hacer énfasis en los elementos técnicos que se deben desarrollar para la viabilidad del sistema. Se procura dar un conjunto de soluciones técnicamente posibles para la implantación del sistema, tomando en cuenta el proceso que todo servicio de telecomunicaciones atraviesa.

El presente proyecto, propone una nueva alternativa de interactividad y de evolución en la calidad del servicio. El estudio plasma la estructura de una aplicación novedosa y de última tecnología a la televisión digital.

El paso hacia la televisión digital supone una evolución en todos los sentidos, desde el aspecto de la infraestructura técnica, hasta la percepción del usuario; es por esto que, es necesario generar la comodidad y la innovación pertinente para que el espectador pueda disfrutar de una aplicación totalmente nueva y novedosa como la televisión estereoscópica.

Mediante el presente estudio, se busca dar una guía de la futura implementación del sistema, y de la aplicabilidad técnica que este servicio puede tener, en el corto, mediano y largo plazo. Tomando en consideración la situación técnica actual del estándar adoptado por el Ecuador (ISDB-T internacional)

CAPÍTULO 1

INTRODUCCIÓN Y DEFINICIONES

1.1 INTRODUCCIÓN

El ser humano por naturaleza ha sido y es capaz de distinguir su entorno en tres dimensiones, siendo la profundidad el aspecto que da realismo a una escena; la televisión desde sus inicios ha tenido un gran desarrollo y su objetivo principal ha sido llevar una escena de la realidad a un usuario final, que éste lo perciba fielmente y trate de interpretar lo que la escena muestra. Sin embargo, esta imagen ha sido siempre bidimensional, provocando que el sistema de TV no reproduzca la realidad en su totalidad, y que el ser humano no perciba realmente el sentido de la profundidad.

Durante los más de 70 años de la TV, desde sus primeras emisiones, el espectador siempre debió en su cerebro completar la información de profundidad de una escena de Televisión, es decir interpretar el elemento que a la escena le hacía falta para que adquiriera realismo. El ser humano por naturaleza adquiere realismo en todas las imágenes que observa de su entorno, aún así le es imposible que esto suceda en el sistema de imágenes más usado por la sociedad como lo es la televisión. Es decir la capacidad del ser humano para percibir realismo de una imagen, viene relacionada íntimamente con su capacidad de percibir la profundidad.

Desde hace muchos años se ha investigado y se ha tratado de encontrar la manera de reproducir ese realismo de una imagen de la vida diaria de alguna forma, recogerla y reproducirla. Esta necesidad ya fue interpretada por antiguos matemáticos y artistas como Euclides y Leonardo Da Vinci que ya estudiaron este fenómeno, con esto queda claro que este problema ha estado latente durante muchos años y de igual forma se ha desarrollado a pasos agigantados. Con estas inquietudes, nace lo que se conoce como estereoscopía¹ y con ello, el sinnúmero de pruebas técnicas y experimentos para lograr efectos 3D tan reales como las

¹ Término asociado a la visualización de una imagen en tres dimensiones.

escenas mismas del entorno cotidiano. En el momento en que la Televisión se vuelve masiva y tiene un grado alto de uso, surge también la necesidad por saber si en ese sistema se pueden ver imágenes de características estereoscópicas, y es así que en la industria del cine el 10 de junio de 1915 se emitieron las primeras proyecciones en 3D en un cine, sacando a la luz esta novedad de realismo en una pantalla perceptible para el ojo humano.

El desarrollo de la Televisión no ha parado, y hoy en vísperas de que nuestro país adopte un nuevo estándar de TV digital², es necesario conocer lo que eso conlleva. La digitalización es una etapa a la que todos los sistemas de Telecomunicaciones deben someterse, y la Televisión hoy se adentra a ese proceso. Una de las grandes ventajas de la digitalización es la facilidad del procesamiento de las señales digitales y la mejora de la calidad respecto a sistemas analógicos. La digitalización, también permite que se aplique la estereoscopia a las imágenes que se presentan en una pantalla de Televisión.

Para poder alcanzar este objetivo, se debe partir del estudio minucioso de las características del ojo que permiten la percepción de la profundidad, los principios físicos involucrados y las definiciones de los conceptos necesarios, para así posteriormente poder aplicar un procesamiento de las señales que representen una imagen de naturaleza tridimensional.

Después de determinar correctamente el proceso para convertir una imagen 3D a una señal eléctrica, se debe analizar el tratamiento de esta señal para transmitirla, visualizarla y en general hacerla útil para un sistema de TV.

Para la digitalización de la TV, existen múltiples estándares existentes que definen la manera de implementarlo. Debido a la importancia de la norma ISDB-T³ en nuestro país, el análisis por completo deberá relacionarse y ser referido a esta norma.

Todo lo que se debe analizar y estudiar contiene una base matemática importante que se debe entender de manera integral para cumplir con los objetivos

² Resolución según el informe de TDT de la Supertel del 26 de marzo del 2010

³ Integrated Services Digital Broadcasting (Terrestrial), estándar de TV digital aceptado en el Ecuador

planteados. El sistema de TV3D⁴ en la TV digital surge como algo novedoso e intrigante, plantea también la solución a la falta de realismo que actualmente se tiene en la TV y como perspectiva de un servicio que a futuro se convierta en algo de uso normal y cotidiano.

1.2 HISTORIA DE LA ESTEREOSCOPIA

La historia de este fenómeno comienza hace aproximadamente 16000 años, en dicha época se cree que algunos pintores dejaron en las cavernas de Lascaux (Francia) representaciones de algunas escenas de animales, que al ser observadas bajo la luz tenue de una antorcha, pueden tomar características tridimensionales. La siguiente noticia de la que se tiene referencia en el mundo de la estereoscopia, son los trabajos realizados por el matemático griego Euclides en el siglo III antes de Cristo (295 A.C). En su tratado Titulado "Óptica" se intenta encontrar una relación entre la visión estereoscópica y el hecho de que los seres humanos posean dos ojos.

En el siglo II de nuestra era, el médico griego Galeno describiría también lo que supuso una profundización de una teoría razonada de la visión de la perspectiva a través de los dos ojos. Supo, por observación, que si se cerraba un ojo, la imagen del entorno cambiaba ligeramente.

En 1833 el físico escocés Charles Wheatstone describió el fenómeno de la visión tridimensional, y construyó "el estereoscopio", el cual era un dispositivo muy sencillo, diseñado para apreciar el relieve de dibujos geométricos, es decir para observar la tridimensionalidad de los mismos.

En lo que respecta a la imagen tridimensional, Wheatstone informó que las pequeñas diferencias entre las imágenes proyectadas a los dos ojos ofrecen una sensación de profundidad del espacio en tres dimensiones. Dicho aparato constaba de dos espejos colocados en ángulo con respecto al observador y una imagen situada frente a cada espejo.

Gracias al invento del estereoscopio, la estereoscopia comenzó a tomar fuerza dentro del mundo de la ciencia, y para el año de 1849 David Brewster diseñó y

⁴ Nombre comercial asignado a la Televisión estereoscópica, también 3DTV.

construyó la primera cámara fotográfica estereoscópica con la que obtuvo las primeras fotografías tridimensionales, del mismo modo diseñó un visor para observarlas. Este desarrollo le permitió a Brewster iniciar la exploración de cámaras que permitan obtener imágenes tridimensionales en movimiento, pero no fue sino hasta la invención de los visores Viewmaster⁵ que la fotografía estereoscópica no se volvió popular.

En los años 50 se hicieron los primeros intentos de explotación comercial de películas en 3D, los cuales lograron tener muy poca incidencia en la industria cinematográfica. Esto, debido a que muchos de los técnicos de la época, desconocían por completo conceptos como “el paralaje”, que son fundamentales en la estereoscopía.

En los años noventa gracias a los adelantos de la electrónica e informática, fue posible llevar la técnica de la estereoscopía a los computadores, lo cual generó que áreas como la medicina, psicología, cartografía, sistemas de modelamiento CAD⁶ y educación entre otras, utilicen esta tecnología no solo para generar nuevas aplicaciones en el medio, sino para generar nuevos mecanismos de investigación en cada una de las áreas.

Hoy en día algunos receptores de Televisión y salas de cine, proyectan contenidos en 3D de alta calidad basándose en los principios básicos desarrollados hace varios años, pero con la diferencia de tener mejor manejo de las señales, lo que aumenta la fidelidad y la calidad.

Un resumen cronológico de los hechos relacionados a este fenómeno físico se presenta en la Tabla 1.1

Aún no se perfecciona de tal forma esta aplicación, como para que se use en la cotidianidad como si ha ocurrido con el audio (hoy en día el audio estéreo es muy usado), pero ya se tiene un alto grado de técnica que permitiría desarrollar una gran cantidad de contenidos para TV, que puedan ser apreciados de manera cómoda y realista.

⁵ Dispositivo visualizador de discos con 7 imágenes en 3-D inventado por William Gruber en 1935.

⁶ CAD (computer-aided design), es el uso de herramientas computacionales para diseño en distintas áreas.

Año	Evento
14000 A.C	En las cavernas de Lascaux, se observan dibujos que con una ligera variación de luz se ven en 3D
295 A.C	Euclides demuestra geoméricamente que cada ojo ve algo diferente, demostración de la esfera de Euclides
II siglo A.C	Galeno determina que cada ojo observa de un objetos partes en común y partes diferentes, lo que determina el principio de visión binocular
1584	Leonardo Da Vinci, usa la perspectiva para demostrar la visión Binocular.
1593	Baptista Porta sostiene que el ser humano observa una escena con un ojo y luego posteriormente con el otro en una sucesión rápida
1611	Kauffman afirma mediante la fisiología del ojo lo sostenido por Porta
1613	Aguillon, crea el concepto de Horóptero ⁷ , con el cual refuerza el concepto de visión Binocular
1833	Charles Wheatstone quien describió el fenómeno de la visión tridimensional, construye “ el estereoscopio “
1838	Charles Wheatstone afirma que la visión doble causada por la disparidad retinal causa la sensación de profundidad
1855	Helmholtz y Czemark contribuye con movimiento a las imágenes a través de discos Estereoboscópicos
1856	Brewster construye el estereoscopio lenticular
1862	Holmes construye algunos otros aparatos para percibir imágenes 3D en movimiento, Estereofantoscopio, Bioscopio, Estereosforoscopio
1900	Patente del sistema anaglífico polarizado.
1922	Primera proyección de cine en 3D película “ The power of Love”
1936	Se aplica en televisión la transmisión 3D en el sistema polarizado
1936	MGM, distribuye el primer sistema Anáglifo
1950	Se lanza el Viewmaster, con el fin de impulsar comercialmente el consumo de imágenes 3D
1980	Se lanza IMAX ⁸ con el cual fue posible crear películas tridimensionales de buena calidad,
1982	Se transmite el Alemania un contenido 3D, en anáglifo
1990	Con el avance de la computación las imágenes pueden ser manipuladas a 3D
2000	El cine 3D, empieza un fuerte auge, y la televisión digital se hace más popular en el mundo
2011	Europa con su estándar DVB-T, emite la especificación de TV3D en ese sistema.

Tabla 1. 1 Resumen histórico de la estereoscopía.

⁷Horóptero: lugar geométrico de los puntos del espacio cuyas imágenes se forman sobre puntos correspondientes de la retina.

⁸ Image Maximum, es un sistema de proyección de cine creado por IMAX Corporation que tiene la capacidad de proyectar representaciones de mayor tamaño y definición que los sistemas aleatorios de proyección.

1.3 FUNCIONAMIENTO DEL SISTEMA VISUAL HUMANO PARA CAPTAR IMÁGENES EN TRES DIMENSIONES

1.3.1 DEFINICIÓN DE ESTEREOSCOPIA

1.3.1.1 Estereoscopia

La estereoscopia puede ser definida, ya sea como una técnica de visualización de los objetos desde el punto de vista de la óptica, así como de una capacidad del ser humano para distinguir imágenes de la realidad. Para el primer caso se define como “cualquier técnica capaz de recoger información visual tridimensional o de crear la ilusión de profundidad en una imagen”⁹.

También se la puede definir como la capacidad del ser humano y de su sentido visual de percibir la profundidad, o de interpretarla de acuerdo a acciones y mecanismos que el sistema visual y el cerebro efectúan para obtener esta percepción. El hecho de poseer dos ojos lo hace posible, ya que se toman percepciones ligeramente diferentes de una imagen y se las fusiona después, esta ligera diferencia se da gracias a la llamada distancia interocular, la cual se define como la distancia entre las pupilas de los ojos humanos, y cuyo valor típico es de 65 mm.

La estereoscopia envuelve algunos otros conceptos que deben ser especificados para comprender claramente y desde varios puntos de vista (físico, psicológico, y fisiológico) los lineamientos que definen completamente a este concepto.

1.3.1.2 Hiperestereoscopia

Cuando la distancia interocular es mayor de 65 mm., el tamaño aparente de los objetos, así como la distancia a que parecen situarse, disminuye en la misma proporción. Es decir el efecto se mantiene pero las apreciaciones de distancia cambian, este proceso es muy usado en las cámaras de TV, debido a que mecánicamente es muy difícil situar dos lentes de alta resolución a 65 mm, por lo tanto se los sitúa más distantes y se procesa la información teniendo en cuenta este particular sin afectar la percepción de profundidad de los objetos.

⁹ Welling, William. Photography in America, Crowell Ed. pág. 23., USA, Nov. 1978

1.3.1.3 Hipoestereoscopia

Caso contrario al anterior, en el cual se disminuye la distancia entre lentes a menos de 65 mm., esta técnica también produce una variación proporcional en las apreciaciones de las distancias, muy usada en cámaras de microscopios 3D para resaltar detalles muy pequeños de las imágenes.

1.3.2 PERCEPCIÓN DE PROFUNDIDAD

A pesar de que los seres humanos poseemos 2 ojos, la imagen que observamos es única y la apariencia es de una sola escena, esto se debe al proceso que el cerebro realiza para “fusionar” las imágenes, sin embargo las razones fisiológicas de poseer dos ojos son:

- Aumentar el campo visual
- Obtener una imagen virtual “mejorada” utilizando las dos imágenes que vienen de cada ojo. Esto es lo que se denomina binocularidad.

El cerebro tiene muchas maneras de interpretar la profundidad de una escena, no necesariamente la estereopsis, la televisión monoscópica que hoy tenemos debe recurrir a estas técnicas para intentar mostrar la realidad de las escenas, a continuación se describen estas técnicas.

1.3.2.1 Superposición

Si un objeto se superpone a otros, el cerebro automáticamente atribuye que ese objeto que está “por delante” está más cerca que él, y de igual forma que el que se coloca detrás se encuentra más lejano.

El problema radica en no conocer exactamente la distancia entre los objetos en cuestión.

1.3.2.2 Perspectiva o puntos de fuga

Este mecanismo fue descrito por Leonardo Da Vinci y es simplemente nuestra capacidad automática de calcular la distancia a los objetos (o de los objetos entre sí) en base a la perspectiva de las cosas. Un ejemplo de ello es el observar las

líneas férreas como líneas no paralelas, lo cual da una idea de que los objetos situados donde las líneas están más cerca se encuentran más lejos.

1.3.2.3 Tamaño de objetos conocidos

Otro mecanismo importante es considerar la información que ya tenemos presente acerca de ciertos objetos. Sabemos el tamaño relativo de un auto y una mesa, entonces si vemos una imagen en la que los dos tienen el mismo tamaño vamos a entender que la mesa está más cerca (por comparación de tamaños). Este mecanismo es muy usado en televisión haciendo énfasis en la naturaleza física de los objetos.

1.3.2.4 Tamaño de la imagen sobre la retina

En base al tamaño de un objeto sobre nuestra retina el cerebro calcula la distancia a ese objeto. Por ejemplo si caminamos hacia una puerta, la imagen de la puerta se va agrandando con el acercamiento. De esta manera el cerebro calcula la distancia a la que estamos.

1.3.2.5 Estereopsis

Los dos ojos captan dos perspectivas ligeramente diferentes de una misma imagen, el cerebro fusiona esas dos imágenes para crear una sola. Este proceso le da volumen o solidez (estereopsis, del griego imagen sólida: estero=sólido) a esa imagen, pudiendo además, calcular distancias. Los ojos poseen la capacidad de converger en ciertos objetos cercanos para poder enfocar. La convergencia es un factor muy importante para el cálculo de distancias.

Estas dos visiones son ligeramente diferentes, porque se observan con una diferencia de alrededor de 65 mm., es decir, se perciben dos imágenes de una escena desplazadas la distancia interocular.

En la estereopsis intervienen diversos mecanismos, cuando observamos objetos muy lejanos, los ejes ópticos de nuestros ojos tienden a ser paralelos, cuando observamos un objeto cercano, nuestros ojos giran para que los ejes ópticos estén alineados sobre él, este proceso es conocido como convergencia. A su vez

se produce la acomodación para ver nítidamente el objeto. El proceso en conjunto se denomina fusión.

La capacidad para fusionar imágenes varía de persona a persona, de ahí se despliega el concepto de agudeza estereoscópica, la cual se define como la capacidad de discernir, mediante la estereopsis, detalles situados en planos diferentes y a una distancia mínima.

Hay una distancia límite a partir de la cual no existe la capacidad de apreciar la separación de planos, y que varía de unas personas a otras, esta distancia límite a la que se deja de percibir la sensación estereoscópica puede variar desde unos 60 metros hasta varios cientos de metros.

1.3.3 DESCRIPCIÓN DEL OJO HUMANO ANTE LA PROFUNDIDAD

1.3.3.1 Descripción del ojo humano

El funcionamiento del ojo humano para interpretar una escena es similar al de una lente de una cámara ya sea fotográfica o de video, es decir las variaciones de luz que una imagen presenta, estimula ciertos músculos oculares y logra distinguir los diferentes objetos que alrededor se encuentran, en cuanto a la percepción de la profundidad, la trayectoria de la imagen comienza en las pupilas y esta señal óptica llega a la retina (foto-receptor humano) la cual es la encargada de interpretar y retener la información de imagen, esto sucede en ambos ojos y para cada objeto situado a cierta distancia.

Los puntos de la retina en los cuales choca la señal óptica de cada objeto son diferentes, estas señales se fusionan en el quiasma óptico¹⁰ (ver Figura 1.1), y el cerebro interpreta esta diferencia de puntos de señal en la retina, como la distancia a la cual se encuentra un objeto, y por lo tanto se percibe la profundidad.

A continuación se detallan algunos aspectos importantes para comprender el comportamiento del ojo en el estudio en cuestión.

¹⁰ El quiasma óptico es el cruce de las fibras nasales de los nervios ópticos, de tal forma que, aproximadamente, la mitad de las fibras pasan del nervio óptico derecho a la cinta óptica izquierda, y viceversa.

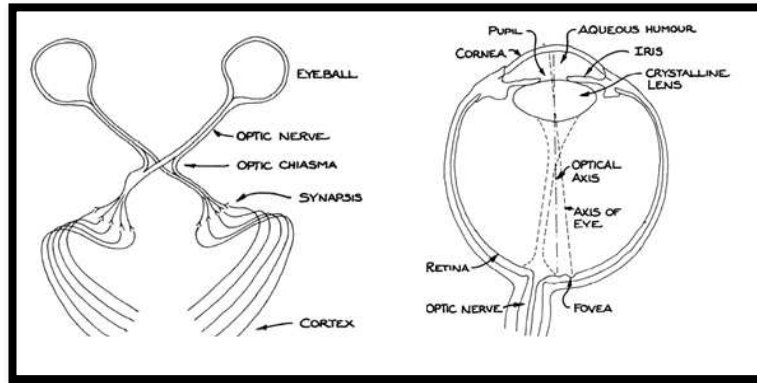


Figura 1. 1 Trayectoria Visual y descripción del ojo humano¹¹

Así pues, la percepción del relieve presenta un doble aspecto: en primer lugar monocular, un solo ojo si puede percibir profundidad a través de las técnicas que el cerebro usa para interpretar distancia de los objetos (técnicas desarrolladas en 1.3.2); y en segundo lugar el aspecto binocular, ya que al mirar simultáneamente con ambos ojos se obtiene la visión estereoscópica o estereopsis, la cual permite apreciar realmente el relieve, en forma independiente de la visión monocular. Este fenómeno, ya se abordó antes en 1.3.2.5 y constituye el principio fundamental que se lleva a las cámaras de TV para producir y captar una imagen con información de profundidad.

1.3.3.2 Campo de Visión Humana

Una de las características importantes de la visión humana es el campo de visión, el cual se define como el espacio o lugar geométrico en el que el ser humano puede percibir un objeto, este espacio se mide en grados, debido a que es un ángulo que se forma por los puntos extremos en los que el ojo distingue y el punto medio de las pupilas. Teóricamente el campo visual horizontal llega a 180° y en vertical llega a 130°, sin embargo estos valores varían de acuerdo a la persona. Este campo es variable para la visión monoscópica y para la estereoscópica, este fenómeno reduce de cierta forma los requerimientos de la 3DTV, ya que el campo de visión en 3D (ver Figura 1.2) es menor y puede significar alguna restricción como el no usar pantallas panorámicas o con relaciones de aspecto que prioricen el sentido horizontal en gran medida.

¹¹Fuente: Foundations of the Stereoscopic Cinema, Lenny Lipton, 1982, p.18, p.62.

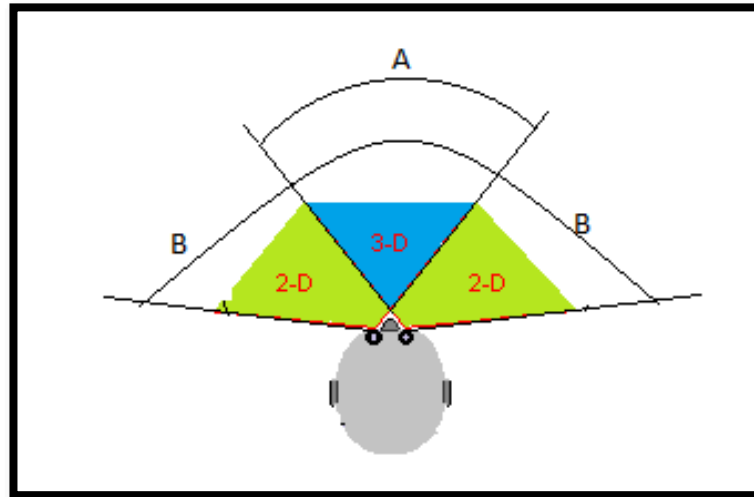


Figura 1. 2 Campo visual humano para 2D y 3D.

La Figura 1.2 expone la proporción del campo de visión tridimensional con el monoscópico, estos valores en promedio son de 170° en total para 2D+3D y para únicamente 3D, de 50° a 70° .

1.3.3.3 Convergencia Visual

Las pupilas de cada ojo proyectan una imagen separada de los objetos en cada retina. Para que esos sean vistos como una imagen única en el cerebro, la porción central de cada retina debe “ver” el mismo punto del objeto. Los músculos responsables de esta convergencia, de la rotación hacia dentro y hacia afuera de los ojos, proveen la información de la distancia como lo hace una cámara. Se denomina convergencia también al ángulo hecho por los dos ejes visuales del par (ver Figura 1.3).

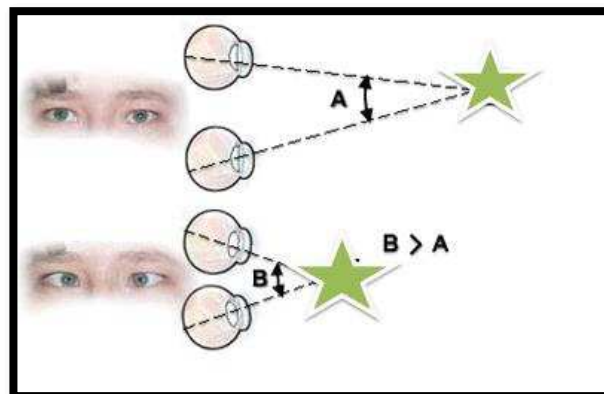


Figura 1. 3 Convergencia visual de objetos A y B a diferente distancia.

1.3.3.4 Enfoque o acomodación Visual

Este concepto se maneja de forma similar en una cámara y en nuestro sistema visual, en la cámara es el ajuste de la longitud focal de la lente, en el ser humano esto es realizado por los ojos a través del esfuerzo muscular del sistema visual. Esta tensión del músculo cambia la curvatura del cristalino, y por ende, la distancia focal del mismo, provocando así que la imagen sea vista con nitidez.

1.3.3.5 Agudeza estereoscópica

Los conceptos de convergencia y enfoque antes descritos, describen el proceso para captar un objeto por el ojo humano, en resumen un objeto ubicado a cierta distancia hace que converjan los ejes visuales y posteriormente se enfoca para verlo nítidamente, este conjunto de acciones se conoce como fusión, y cada ser humano tiene mejor o menor desarrollada esta capacidad, se define a la agudeza estereoscópica como la capacidad de distinguir, mediante la estereopsis, detalles situados en planos diferentes y a una distancia mínima. Es decir la habilidad del ojo humano para fusionar las imágenes e identificar la distancia entre nuestros ojos y los objetos explorados.

1.3.3.6 Disparidad binocular

Se define como la desigualdad o diferencia de varias percepciones de un objeto, que se proyecta en la retina del ojo. La disparidad binocular es considerada como la señal más importante en la percepción de profundidad sobre distancias medidas por la visión. El grado de disparidad entre las dos imágenes depende del ángulo paraláctico (o ángulo de convergencia). Este es el mecanismo que usa el cerebro para distinguir el relieve de un objeto, ya que estas pequeñas desigualdades, que son percepciones ligeramente diferentes de cada ojo, permiten la visualización binocular.

1.3.3.7 Definición de paralaje (Parallax)

Es el resultado del cambio de la posición de un objeto en el espacio debido tanto al movimiento del objeto como del punto de vista del observador. El paralaje visual del movimiento es una función en el cual la imagen de un objeto se mueve a través de la retina. Este fenómeno se describe como la desviación en ángulo

que un objeto presenta dependiendo del punto de vista de donde se lo mire. Dentro de la convergencia surge el concepto del ángulo paraláctico o de paralaje y constituye el principal elemento en la determinación de la distancia de un objeto. La disparidad binocular constituye el paralaje relativo de los objetos porque el ángulo de paralaje determina puntos diferentes de proyección de una imagen en la retina. Este concepto es el más usado en la producción de contenidos 3D, ya que significa una herramienta muy útil para crear efectos atractivos al ojo, por ejemplo paralaje variable para darle movimiento de profundidad a un objeto. A continuación se describe el desarrollo matemático del ángulo de paralaje, el cual será útil en el momento de estudiar las propiedades geométricas de los lentes de una cámara estereoscópica.

La Figura 1.4 muestra dos objetos a distancias y posiciones diferentes, cada punto se proyecta en la retina en un lugar diferente lo cual se define como paralaje relativo, $a'-a''$ y $b'-b''$ describen esta disparidad.

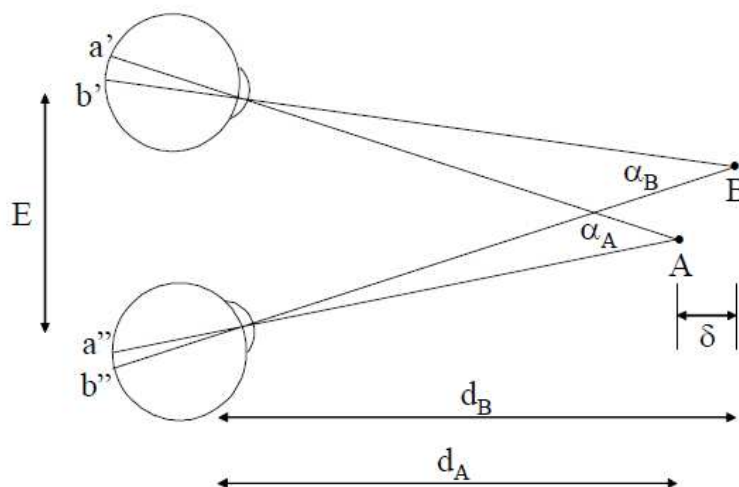


Figura 1. 4 Ángulos de paralaje para objetos en diferentes posiciones.

Los límites para visión estereoscópica se dan a una distancia de mínima de 150mm, dado que a menos de eso no se consigue fusión, y en el límite superior se tiene hasta 200 metros, esta restricción viene dada por la resolución angular del ojo.

Como se observa en la Figura 1.5 la trayectoria visual de los ojos sobre un objeto forma dos triángulos rectángulos con determinadas dimensiones:

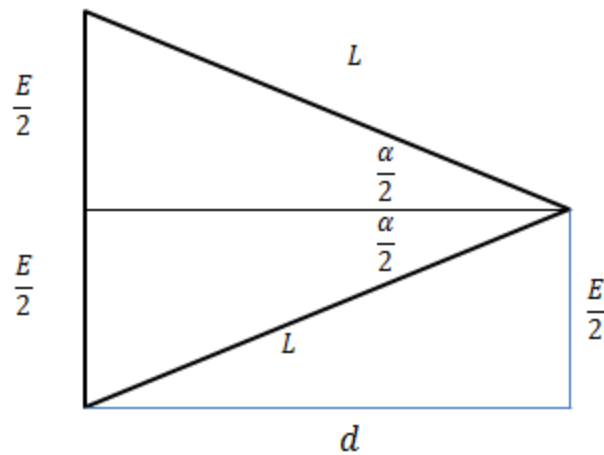


Figura 1. 5 Geometría de las distancias en la convergencia.

De donde se puede establecer que:

$$\sin \frac{\alpha}{2} = \frac{\frac{E}{2}}{L} \quad \mathbf{Ec\ 1.1}$$

Pero en radianes se puede decir que $\sin(\alpha) \rightarrow \alpha$ (si el ángulo es pequeño), entonces:

$$\begin{aligned} \frac{\alpha}{2} &= \frac{E}{2L} \\ \alpha &= \frac{E}{L} \end{aligned} \quad \mathbf{Ec\ 1.2}$$

Se tiene que E es 65 mm y $E^2/4$ es muy pequeño, por lo tanto la expresión del radical se puede simplificar y se obtiene:

$$\begin{aligned} L &= \sqrt{\frac{E^2}{4} + d^2} = d \\ \alpha &= \frac{E}{d} \end{aligned} \quad \mathbf{Ec\ 2.3}$$

1.3.3.8 El efecto Pulfrich

El llamado Efecto Pulfrich fue descubierto por el médico alemán Carl Pulfrich en 1922. El fenómeno describe la percepción de un efecto estereoscópico cuando se observa una imagen en movimiento horizontal sobre un plano y con un filtro

oscuro situado delante de uno de los ojos. Debido a la menor luminosidad que percibe el ojo con el filtro, la imagen llega al cerebro con un retardo de unas centésimas de segundo.

Por tanto, en la estereopsis el cerebro percibe la misma imagen pero con una pequeña diferencia de posición horizontal, lo que genera el efecto estereoscópico, con este fenómeno pueden obtenerse algunos efectos estereoscópicos filmando con una única cámara en movimiento.

El efecto supone que el ojo reacciona a velocidades diferentes dependiendo de si este es claro u oscuro, y esto hará que las imágenes lleguen en sucesión al cerebro produciendo un efecto tridimensional.

1.4 TÉCNICAS Y TECNOLOGÍAS PARA LA VISUALIZACIÓN DE IMÁGENES ESTEREOSCÓPICAS¹²

Definido los anteriores conceptos, es claro que se debe lograr presentar de alguna forma esas dos visiones ligeramente diferentes de una imagen a cada ojo, es decir separar las imágenes izquierda y derecha y hacer que solo el respectivo ojo las pueda ver, o intentar capturar el plano de una imagen y la información de la profundidad de forma separada, y posteriormente presentarla en una pantalla que no requeriría aditamentos extras, ya que presentaría la información tridimensional de una imagen.

Si analizamos el sistema estéreo del audio, la manera de hacerlo es tener dos fuentes sonoras, una para cada oído, en el caso de una imagen es complejo definir dos imágenes para cada ojo, por lo cual se han desarrollado diferentes técnicas para que una imagen contenga la información de los dos canales visuales y de igual forma algún sistema para separarlas a cada ojo y percibir la sensación de profundidad.

El objetivo es partir de la adquisición del par de canales visuales tal y como el sistema visual humano lo hace, es decir con dos “ojos” o transductores que están separados aproximadamente los 65mm. correspondientes a la distancia

¹²Clasificación y definición de técnicas según la Recomendación UIT-R BT.2088 del 2006

interpupilar; aún así se puede también, mediante procesamiento digital de imágenes, crear un pseudo 3D a partir de un cuadro de dos dimensiones, estas técnicas ya no requiere los procesos preliminares, lo que necesitan es un estudio y desarrollo minucioso de las técnicas DSP (Digital Signal Processing – Procesamiento Digital de Señales); al final, en cualquiera de los casos, la imagen que se debe presentar para el ojo humano debe tener información L+R¹³, y deben ser percibidas correctamente ante el ojo humano.

Las técnicas que permiten visualizar una imagen fija (cuadro) no han variado mucho desde el inicio de su desarrollo, los conceptos manejados son los mismos, sin embargo la precisión de los materiales y la tecnología de los mismos, han logrado que estas técnicas optimicen la calidad de visión.

A continuación se expone cada uno de estos sistemas, explicando de cada uno el principio físico que lo respalda.

1.4.1 GAFAS 3D (MINI PANTALLAS)

Este sistema supone una solución sencilla al problema de que cada ojo reciba la señal que le corresponde, en este sistema se utilizan 2 mini-pantallas acordes al tamaño del ojo en una estructura (montura) tal que permita sostener cada pantalla en cada ojo y mostrándole solo la información correspondiente; debido al tamaño de las pantallas la resolución es limitada y dado que las imágenes están en movimiento se crea una ligera molestia visual; además este es un sistema eminentemente personal y es imposible la visualización en grupo. En el inciso 1.2 se describe que este fue un primer elemento usado en fotografía pero hoy ha ido perdiendo mucha utilidad.

1.4.2 PANTALLAS SEPARADAS Y GAFAS PRISMÁTICAS

Esta técnica consiste en presentar pantallas separadas a una cierta distancia para que los rayos luminosos recibidos por el ojo sean desviados de forma correcta, lo cual se consigue mediante gafas con espejos o prismas. Esta técnica exige que el espectador mueva lo menos posible la cabeza, lo que puede causar gran fatiga.

¹³ Información estereoscópica de los canales visuales Izquierdo + Derecho, llamado I+D ó L+R.

1.4.3 TÉCNICA ANAGLÍFICA

Esta técnica propone la utilización de lentes anaglíficos, los cuales superponen dos imágenes que a continuación se separan para visualizarlas mediante un filtro óptico

1.4.3.1 Anáglifo cromático

Este sistema propone ubicar en un mismo cuadro las dos imágenes de dos capas de color, superpuestas pero desplazadas ligeramente una respecto a la otra para producir el efecto de profundidad. Cada imagen, para ser vista debe ser filtrada por lentes especiales en cada ojo que separa y hace que los ojos tengan esa diferencia de percepción de la imagen, estos filtros son óptimos al usar colores complementarios en la imagen y en las gafas, la imagen presentada por ejemplo en rojo no es vista por el ojo que tiene un filtro del mismo color, pero sí que ve la otra imagen en azul o verde. Es muy común utilizar el par rojo/cian ya que esta combinación tiene características simples de longitudes de onda paso bajo/paso alto. Debido a la alteración de los colores en los sistemas de televisión tradicionales, la relación de luminancia varía y por ello se trata de buscar los colores complementarios que den mejor resultado, la Tabla 1.2 muestra este efecto.

Color 1	Color 2	Relación de luminancia
Rojo	Cian	
$Y' = 0,299$	$Y' = 0,701$	2,34
Verde	Magenta	
$Y' = 0,587$	$Y' = 0,413$	1,42
Azul	Amarillo	
$Y' = 0,114$	$Y' = 0,886$	7,77

Tabla 1. 2 Valores y relaciones de luminancia de colores complementarios utilizados para la representación anaglífica¹⁴

Para esto manejan los filtros coloreados desde la captación, en el momento en que se recibe la imagen con dos cámaras, luego estas dos señales son

¹⁴Fuente: "Televisión Estereoscópica" - Informe UIT-R BT-2088, año 2006.

presentadas como una sola y los lentes se encargan de separar la imagen a cada ojo. Este sistema supone una técnica sencilla, sin embargo la pérdida de contraste y color hoy en día es inadmisibles ya que produce cansancio visual y alteración de las imágenes originales, algo que en un sistema de entretenimiento de TV no puede suceder, con todo esto se concluye que la técnica no se vuelve del todo conveniente en servicios de alta calidad.

1.4.3.2 Anáglifo por polarización o vectógrafo

Este sistema propone algo similar al cromático, la diferencia se basa en usar la naturaleza de la luz polarizada (ver Figura 1.6) para separar las imágenes izquierda y derecha, dicha separación de las imágenes se basa en la polarización de la luz a $+45^\circ$ para el ojo derecho y a -45° para el ojo izquierdo, se desea al igual que en el sistema anterior diferenciar cada imagen, pero si esta imagen se muestra como luz polarizada tiene componentes determinadas de campo eléctrico y magnético, las cuales pueden ser filtradas en recepción por gafas especiales. Este sistema posee varias ventajas como el ser pasivo, y mantener todo el espectro de colores de una imagen, haciendo que sea una técnica muy sencilla. A continuación se expone como se puede usar la polarización de la luz en una imagen.

1.4.3.2.1 Polarización de la luz

La luz puede tratarse como una onda electromagnética transversal. Una onda luminosa puede ser vista como superposición de varias ondas. Podemos considerarla como una superposición de dos ondas luminosas armónicas, perpendiculares, de igual vector de propagación y frecuencia (monocromáticas). En definitiva si una imagen, que es en esencia luz, se la propaga de forma polarizada, se obtienen dos componentes eléctrica y magnética determinadas y si desfasamos 90 grados dichas componentes para cada canal visual obtenemos una especie de “multiplexación visual”, que son separadas a través de gafas con materiales polarizadores capaces de distinguir estas componentes de campo eléctrico y magnético y separar las señales a cada ojo. Polarizar es modificar los rayos luminosos por medio de refracción o reflexión, de tal manera que queden incapaces de refractarse o reflejarse de nuevo en ciertas direcciones.

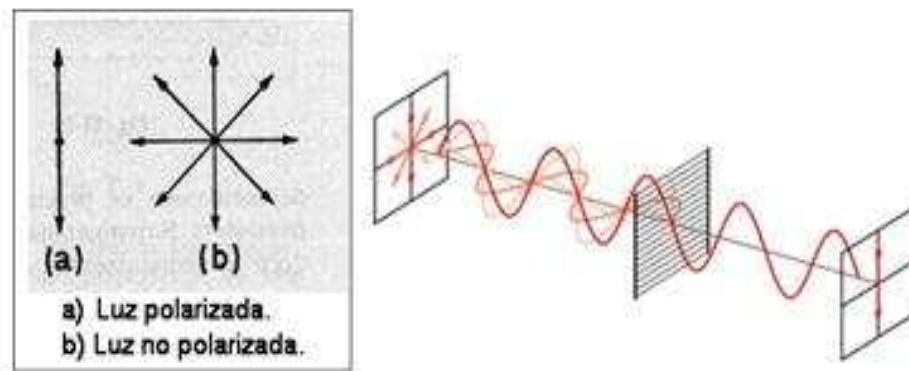


Figura 1. 6 Naturaleza de la luz polarizada

1.4.3.3 Anáglifo lenticular integrado (autoestereoscopia)

En este sistema, las dos imágenes (L+R) se presentan simultáneamente en tiras verticales intercaladas. Un sistema lenticular de lentes cilíndricas situado en la pantalla enfoca ambas imágenes en forma separada aproximadamente a una distancia interocular cercana a los 65mm.

Este sistema exige que el espectador no mueva su cabeza bruscamente, y restringe el perfil multivisión de una pantalla de televisión, es decir un cambio de posición del espectador afecta a la imagen que éste observa. Existe la sensación de que la escena gira con el movimiento del observador, este fenómeno se conoce como Free viewpoint (punto de vista libre), y hoy en día la tecnología ha desarrollado 8 de estos puntos. Cada Free Viewpoint requiere dos imágenes (una por cada ojo) lo que hace que la pantalla tenga una resolución mucho mayor que la HDTV (High definition Television). Con esto se desarrolla en algo la técnica que permita que existan múltiples espectadores, ya que no es necesario localizarlos en posiciones fijas. El principal cambio es la utilización de micro-lentes que permiten controlar la difracción de los haces de luz. Como no se necesitan gafas especiales, se lo conoce como sistema autoestereoscópico.

Hoy en día se han desarrollado algunas técnicas para hacer que este sistema pueda ser visto por varios espectadores y no produzca cansancio por el requisito de mantener estática la cabeza, en el capítulo 3 se estudian las pantallas y se conocerá a detalle su funcionamiento.

1.4.3.4 Multiplexación Temporal

Este procedimiento propone que las imágenes correspondientes a cada ojo se presenten en forma sucesiva una después de otra, en este sistema la información L+R no está contenida en la misma imagen, hay un par de cuadros independientes que se muestran en rápida sucesión. Es necesario separar estas imágenes ante el espectador, y esto se logra con obturadores de alta velocidad alternados y sincronizados, los cuales se montan en gafas especiales, estos obturadores son usualmente del tipo LCD (liquid crystal display). La velocidad a la cual se alternan las imágenes es conocida como velocidad de renovación, y por ello se pueden producir imágenes con parpadeo, sin embargo se ha mejorado y se han determinado ciertos valores donde se tiene buena calidad de visualización (12 imágenes/seg. para el cine tradicional y 12,5-15 imágenes/seg. para la televisión con exploración entrelazada). Este sistema funciona mucho mejor a velocidades de renovación más elevadas con exploración progresiva.

1.5 ESTÁNDAR DE TELEVISIÓN DIGITAL ISDB-T

De todos los servicios de telecomunicaciones existentes, la señal de video de la televisión ha sido la que más ha tardado en digitalizarse, los beneficios de una señal digital son indiscutibles; la facilidad de manejo, la robustez en la transmisión y la calidad al usuario final son entre otras, ventajas que hacen sólida y conveniente la digitalización. La mayor desventaja de una señal digital es sin duda el requerimiento de un mayor ancho de banda para su difusión con respecto a una señal analógica; esta dificultad ha significado que la televisión demore mucho más que los otros servicios en brindarse de forma digital.

La frecuencia de una señal de video es mucho más alta que la de otras señales analógicas, por lo tanto requiere mayor ancho de banda espectral (6, 7, 8 MHz)¹⁵ para su transmisión, lo cual ha determinado el número limitado de canales para una difusión de TV analógica radiada. La digitalización de esta señal se la hace al igual que cualquier otra mediante esquemas de muestreo definidos, esto provoca que el requerimiento para el canal de transmisión y el procesamiento se haga

¹⁵ Anchos de banda referenciales de la TV analógica para los sistemas tradicionales: NTSC, PAL, SECAM.

imposible de lograr en primera instancia, sin embargo la evolución de las técnicas de compresión de información digital son las que hoy permiten que el video en la TV se pueda digitalizar y difundir sin problema.

Alrededor del mundo se han desarrollado diversas técnicas respectivas para crear la Televisión digital, los estándares se han difundido en regiones del planeta determinadas y todas apuntan a un objetivo común; brindar un servicio con mayor calidad que la televisión convencional y dar servicios extras como interactividad, manejo de contenidos, y televisión 3D. El Ecuador en el año 2010 a través del organismo de regulación respectivo (Consejo Nacional de Telecomunicaciones CONATEL)¹⁶, se define que de todos los estándares desarrollados por diferentes países, se adopte el ISDB-T originario de Japón y con modificaciones Brasileñas, el estándar se conoce como ISDB-T internacional y hoy en día la mayoría de países de Sudamérica han decidido implementarlo.

1.5.1 CARACTERÍSTICAS DE LA TELEVISIÓN DIGITAL

La televisión digital se define como la transmisión de señales discretas de audio y video, es decir representadas por estados lógicos 0 ó 1. Para cualquier información de una imagen se asigna un valor binario determinado y estos valores en conjunto forman la señal digital de televisión, esta señal es susceptible de procesamiento y mejoras como la regeneración y protección ante perturbaciones externas. Las etapas que claramente se definen para un sistema de Televisión digital se exponen en la Figura 1.7.



Figura 1. 7 Sistema de Transmisión de Televisión Digital

¹⁶ Adopción del estándar según : “RESOLUCIÓN 084-05-CONATEL-2010”

1.5.1.1 Digitalización

Para que a la información de video analógico le corresponda un valor definido binario o cadena de bits, se debe muestrear la señal, lo cual significa analizar cada cierto tiempo a dicha señal y al valor que la señal muestra en ese instante asignarle un valor en bits. La señal de televisión está compuesta de dos señales analógicas: luminancia¹⁷ y crominancia¹⁸, y ambas deben ser muestreadas de manera coherente.

La televisión, muestra a la imagen como varios elementos de imagen (pixel¹⁹) que varían su iluminación de tal forma que ante la persistencia del ojo se nota una imagen continua. Para NTSC²⁰ la pantalla se divide en 525 líneas horizontales y la relación de aspecto horizontal / vertical es de 4:3. Con este punto de partida, el muestreo se realiza a 13.5 MHz en luminancia y 6.25 MHz en crominancia, para sistemas de 525 y 625 líneas y bajo el esquema de muestreo A:B:C, el cual determina que por A muestras de luminancia existen B muestras de crominancia azul y C muestras de crominancia rojo. Para la televisión los esquemas más utilizados son los de 4:2:2 en producción de alta calidad y de 4:1:1 y 4:2:0 para transmisión.

Después de definida la cantidad de muestras que se toman por señal, se debe establecer cuantos bits cuantizan a esta muestra, la UIT-R BT.601-6²¹ define que se pueden asignar 8 ó 10 bits por muestra con cuantificación uniforme. Los diferentes esquemas de muestreo existentes definen sus propios requerimientos del canal para determinada velocidad de transmisión, de igual forma las muestras de crominancia tanto verticales como horizontales varían de acuerdo al esquema; dando como resultado diferentes resoluciones horizontal/vertical. Considerando este conjunto, se obtienen los resultados de la Tabla 1.3.

¹⁷La luminancia "Y" es la componente del video que tiene la información de luminosidad (luz o brillo) de la imagen, matemáticamente en la TV a color se calcula mediante la combinación de los colores primarios en la relación: $Y = 0,299R + 0,587G + 0,114B$.

¹⁸La crominancia es el componente de la señal de vídeo que contiene las informaciones del color (R,G,B).

¹⁹ Pixel: La menor unidad homogénea en parte de una imagen digital.

²⁰ NTSC: National Television System Committee es un sistema de codificación y transmisión de Televisión en color analógico desarrollado en Estados Unidos.

²¹ Recomendación UIT-R BT.601-6, "Parámetros de codificación de televisión digital para estudios con formatos de imagen normal 4:3 y de pantalla ancha 16:9", 2007.

Muestreo	Tasa Binaria	Muestras horizontales de crominancia	Muestras verticales de crominancia	Resolución H/V
4:2:2	270 Mbps	360	240	1.5
4:1:1	202,5 Mbps	180	240	1.75
4:2:0	202,5 Mbps	360	120	3

Tabla 1. 3 Características de los esquemas de muestreo.

Con lo anteriormente descrito queda en evidencia la necesidad de técnicas de compresión que permitan que la señal de video disminuya su velocidad de transmisión a niveles aceptables para cualquier medio de transmisión sea este guiado o no guiado.

El anterior análisis está definido para SDTV (Standard Definition Television), debido a que también se puede tener HDTV (High Definition Television) con un mayor número de líneas, se ha establecido que un sistema HDTV tiene una frecuencia de muestreo de 74,25 MHz en luminancia; y se dan dos formatos de HDTV: 1080x1920i y 720x1920p, el primero con 1080 líneas activas en exploración entrelazada (exploración de líneas impares en primera instancia y posteriormente líneas pares) y el segundo con 720 líneas activas en exploración progresiva (se exploran las líneas en orden consecutivo), ambos sistemas mantienen 1920 pixeles por línea.

1.5.2 ESPECIFICACIONES TÉCNICAS DE ISDB-T

ISDB-T (International Service Digital Broadcasting – Terrestrial), se desarrolló en Japón en el año 1990 y hacia el año 2006, Brasil incluyó varias mejoras que lo convirtieron en un estándar internacional definido.

A continuación se describen las características del estándar, sin especificar a detalle las técnicas utilizadas en cada etapa, debido a que se analizarán posteriormente y estarán relacionadas a contenidos digitales 3D.

- La transmisión se realiza mediante OFDM (Orthogonal Frequency Division Multiplexing)²², que brinda eficiente protección al ruido y permite movilidad y portabilidad.
- Se pueden tener transmisiones fijas móviles y portátiles.
- Incorpora interactividad usando el Middleware Ginga²³, y soportando como canal de retorno al Internet.
- Permite la utilización de SFN²⁴ (Redes de frecuencia única).
- Utiliza un canal de 6 MHz dividido en 13 segmentos donde se pueden tener distintas configuraciones para canales HDTV y SDTV.
- Utiliza transmisión jerárquica priorizando con modulaciones robustas a los servicios vulnerables. Configura diferentes tipos de modulación y FEC (Forward Error Correction) para priorizar al servicio.
- El segmento 7 se define como “one-seg”, está destinado a la movilidad y soporta recepción en movimiento a altas velocidades.
- Utiliza el estándar MPEG-4 de compresión de imágenes permitiendo tasas menores a buenas calidades de imagen.
- Utiliza entrelazado en tiempo y frecuencia para facilitar la corrección de errores y hacer más robusto al sistema.

1.5.2.1 Compresión de audio y video

Las técnicas de compresión del estándar Internacional representan una de las mejoras sustanciales con respecto al sistema Brasileño dado que se permiten comprimir más canales para el mismo ancho de banda. El video se codifica con el estándar MPEG-4 AVC (H.264), y el audio en modo multicanal MPEG-4 AAC nivel 4 o MPEG-4 HE-AAC nivel 4 y al usar estéreo con las mismas modalidades pero en nivel 2.

²²OFDM: Técnica usada en varios sistemas de transmisión, es un sistema multiportadora, donde dichas portadoras se encuentran dispuestas ortogonalmente en el espectro, sistema con mayor protección a interferencias debido a la larga duración del símbolo OFDM.

²³GINGA: Conjunto de software ubicado entre el código de las aplicaciones y la infraestructura de ejecución (plataforma de hardware y sistema operativo), que sirve para desarrollar aplicaciones de interactividad para la TDT.

²⁴ SFN: En las redes de frecuencia única la difusión de una señal se hace en la misma frecuencia para todos los transmisores, requiere portadoras sincronizadas y en las zonas de solapamiento de cobertura no sufren interferencias ni pérdidas, al contrario reciben ganancia.

1.5.2.2 Transmisión de la señal en ISDB-T

ISDB-T procura tener gran fortaleza y robustez en la transmisión de la señal, debido a que las condiciones geográficas de los lugares donde se implementaría son variadas y se debe proveer de técnicas de transmisión apropiadas. El estándar utiliza OFDM para este propósito (ver Figura 1.8), este es un sistema que brinda gran inmunidad al ruido y a las pérdidas por multitrayectoria. El hecho de basar la transmisión en OFDM permite la utilización de SFN, y mejorar la calidad de recepción. Este tipo de transmisión es segmentada, dividiendo en 13 slots al canal de 6 MHz, cada segmento tiene un ancho de banda de 426 KHz, y especificando distintas funcionalidades para cada uno; OFDM es multiportadora y este número de portadoras puede variar de acuerdo a los requisitos, es así que se tienen 3 modos para ISDB-T que son 2k, 4k, 8k los cuales definen 1405, 2809, y 5617 portadoras respectivamente.



Figura 1. 8 Espectro de la modulación OFDM

1.5.2.2.1 Modulación QPSK (*Quadrature Phase Shift Keying*)

Es un tipo de modulación digital donde la fase de la onda portadora se modula para codificar bits de información digital en cada cambio de fase. QPSK posee cuatro estados para la modulación, asignando un estado a cada dos bits de la cadena de datos.

1.5.2.2.2 Modulación 16 QAM (*Quadrature Amplitude Modulation*), 64 QAM

QAM combina dos técnicas de modulación digital, variando la onda portadora en amplitud y en fase de acuerdo a la información digital, 16 QAM define 16 estados

de modulación asignando a 4 bits un valor en fase y amplitud. 64 QAM establece 64 estados para la modulación.

1.5.2.2.3 Modulación DPSK (Differential Phase Shift Keying)

En este tipo de modulación la información digital hace variar a la fase de la señal analógica a transmitirse, esta fase instantánea de la portadora sinusoidal conmuta entre los valores digitales de la señal de entrada (2 valores para señal de dos estados), en la modulación DPSK cada estado de modulación es codificado por un salto respecto a la fase que tenía la señal anterior.

Antes se mencionó a la transmisión jerárquica como característica importante, vale decir que entre las modulaciones que se usan para lograr esto se tienen a QPSK, 16QAM, 64QAM y DQPSK (ver Figura 1.9). Esta transmisión jerárquica se basa en combinar los esquemas de modulación y la tasa de corrección de errores en el código convolucional (ver Figura 1.10) de manera de priorizar y dar el tratamiento adecuado a cada necesidad de robustez en ambientes hostiles o capacidad para aplicaciones que demanden un mayor flujo de datos.

MODULACIÓN			
QPSK	16QAM	64QAM	DQPSK
Más robustez	←		Menos robustez
Menos capacidad	→		Más capacidad

Figura 1. 9 Características de los esquemas de modulación posibles

CORRECCIÓN DE ERRORES (FEC)				
1/2	2/3	3/4	5/6	7/8
Más robustez	←			Menos robustez
Menos capacidad	→			Más capacidad

Figura 1. 10 Características de las tasas de corrección de errores posibles.

Los 13 segmentos del canal conforman 2 secciones A y B que permite la versatilidad de tener posibilidades de ocupar los 6 MHz de distinta forma (ver Figura 1.11), combinando las opciones de HDTV, SDTV y móvil.

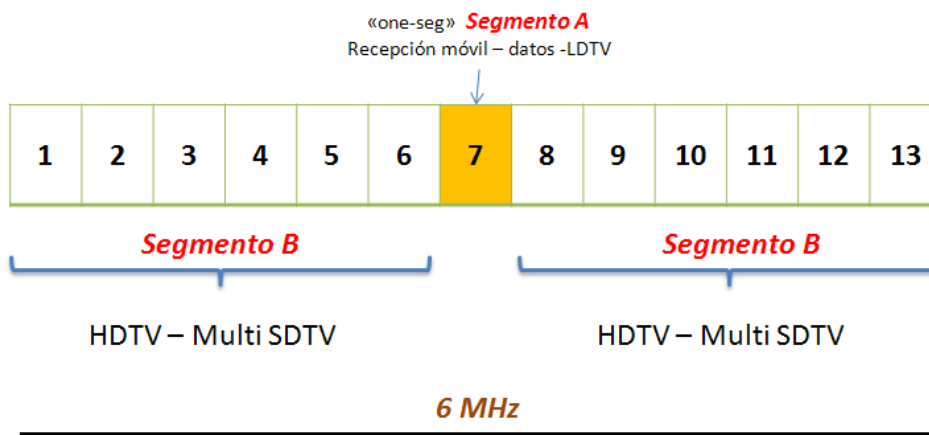


Figura 1. 11 Distribución segmentada del canal.

Un ejemplo de una de las posibilidades existentes para el sistema, donde se usa transmisión jerárquica se representa en la Figura 1.12.

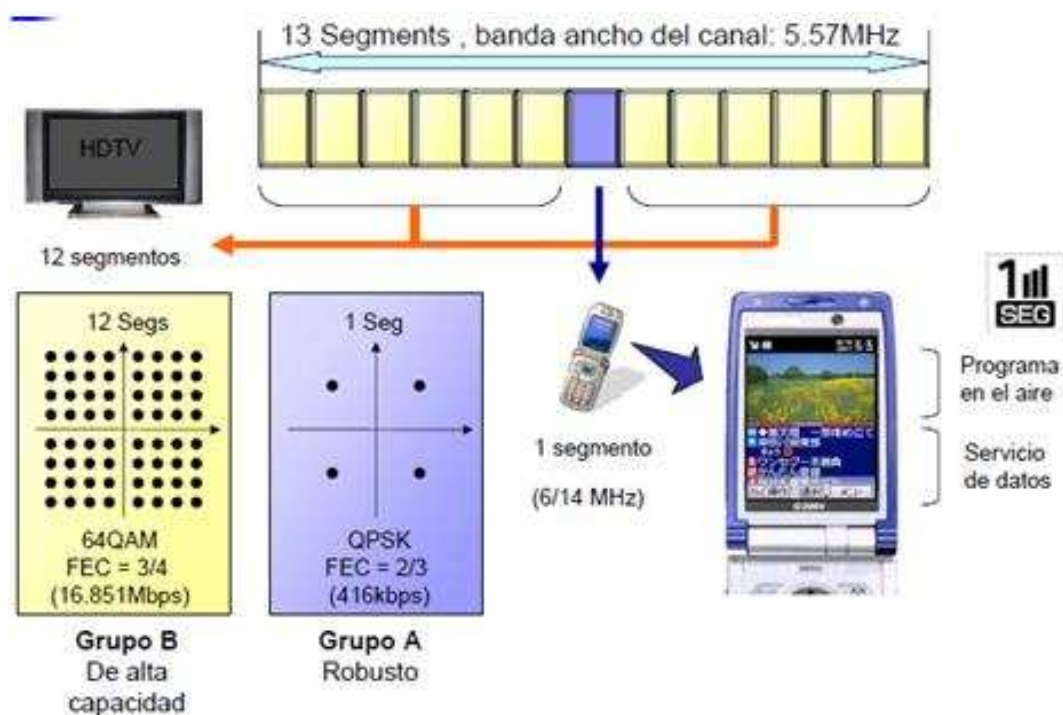


Figura 1. 12 Ejemplo de uso de los 13 segmentos del canal.

El ejemplo muestra claramente la combinación de la modulación y del FEC para diversas aplicaciones, se observa que el segmento 7 está usado para recepción móvil.

Con todo lo anteriormente expuesto, el estándar ISDB-T tiene algunas características que brindan muchas posibilidades para la transmisión del audio y del video.

Un resumen de dichas especificaciones técnicas se muestra en la Tabla 1.4.

	MODO 2K	MODO 4K	MODO 8K
# De segmentos OFDM	13		
Ancho de banda útil	5,575 MHZ	5,573 MHZ	5,572 MHZ
Espacio entre portadoras	3,968 KHZ	1,984 KHZ	0,992 KHZ
Total de portadoras	1405	2809	5617
Modulación	QPSK, 16QAM, 64QAM, DQPSK		
# De símbolos por cuadro	204		
Duración del símbolo activo	252 μ S	504 μ S	1,008 ms
Duración del intervalo de guarda	1/4, 1/8, 1/16, 1/32, de la duración del símbolo activo		
Codificación interna (canal)	Convolutacional ²⁵ 1/2, 2/3, 3/4, 5/6, 7/8		
Codificación externa (fuente)	Reed Solomon (204, 188) ²⁶		
Tiempo interleaving	Hasta 0,5 segundos		
Tasa de bits útil	3,651 Mbps - 23,23 Mbps		

Tabla 1. 4 Características de transmisión de ISDB-T.

²⁵ Para preparar la señal digital ante las degradaciones del canal se utiliza un código convolutacional que permita su detección, en esta codificación cada símbolo de m bits de información se transforma, al ser codificado, en un símbolo de n bits, donde m/n es la tasa del código ($n \geq m$), la transformación es función de los k símbolos anteriores, donde k es la longitud del código.

²⁶La codificación Reed Solomon (n,k) toma un bloque de bits y añade algunos bits redundantes, para ISDB-T $n=204$ y $k=188$, con esto se pueden corregir 8 bytes por cada paquete de 204 bytes que se reciben.

1.6 ANÁLISIS DE LOS REQUERIMIENTOS DE UN SISTEMA DE TELEVISIÓN ESTEREOSCÓPICA

Los principios básicos y las definiciones expuestas permiten establecer algunos requerimientos que un sistema de televisión estereoscópica debe cumplir. Es necesario definir: cómo utilizar las técnicas conocidas, elegir entre los mejores sistemas expuestos y aplicar los principios de visión estereoscópica a un sistema de televisión digital y más concretamente al estándar en estudio (ISDB-T), para ello a continuación se exponen algunos factores determinantes para la implementación de este servicio. La UIT-R ha definido algunas cuestiones y recomendaciones que establecen requisitos y objetivos que la TV estereoscópica debe cumplir, se analizarán dichas publicaciones y se concluirá qué exigencias y condiciones son indispensables para una favorable y eficiente aplicación del sistema.

1.6.1 CONSIDERACIONES PARA LA EVALUACIÓN SUBJETIVA DE IMÁGENES²⁷

La evaluación subjetiva de imágenes es un elemento vital en el diseño de los sistemas de televisión estereoscópica, por lo tanto es necesario establecer factores que permitan garantizar un sistema confortable para el usuario final, y condiciones de visualización básicas para el desempeño del sistema.

Por lo tanto la UIT-R ha establecido una serie de mecanismos, pruebas y procedimientos para recomendar los requisitos que un sistema de televisión estereoscópica debe cumplir.

1.6.1.1 Factores de evaluación de la televisión monoscópica

Los aspectos fundamentales que siempre han sido considerados en la televisión monoscópica tradicional, se deben mantener para nuevos sistemas. Estos factores, que han establecido una calidad de visualización adecuada en la TV2D, deben conservarse en su aplicación para la TV3D.

²⁷ Análisis según la recomendación UIT, recomendación UIT-R BT.1438 del año 2000.

- Resolución
- Representación del color
- Representación del movimiento
- Calidad general
- Nitidez de perfiles
- Profundidad

Estos factores han sido analizados y estudiados durante el desarrollo de la TV desde sus inicios y se ha logrado mantenerlos en valores que garantizan la correcta visualización de las imágenes, por lo tanto se hace indispensable que la TV3D también los cumpla.

1.6.1.2 Factores de evaluación de la televisión estereoscópica

Existen algunos factores que son exclusivos de la televisión estereoscópica, los cuales permiten establecer en principio las condiciones óptimas de visualización; estos factores no se han desarrollado en su totalidad y se piensa que con el avance de la tecnología se pueden ajustar a otros lineamientos. Estos factores son:

- *Resolución de profundidad*: La profundidad es la señal “extra” de la TV3D, se puede obtener a través de cualquier técnica, desde la adquisición con cámaras estereoscópicas o hiperestereoscópicas o utilizando cualquier método de visualización, debe ser óptima debido a que una reducida resolución en profundidad significaría una baja calidad de la imagen.
- *Movimiento en profundidad*: Este factor determina si en el movimiento de profundidad de los objetos en una imagen se producen discontinuidades.
- *Efecto teatro de marionetas*: Describe una distorsión en las imágenes tridimensionales en la que los objetos se perciben como muy grandes o pequeños en relación a la realidad.
- *Efecto papel de cartón*: Al igual que al anterior es un tipo de distorsión que ocurre en imágenes tridimensionales. Las posiciones de los objetos que se describen en (x,y,z) , es decir tridimensionalmente se perciben

estereoscópicamente sin problema, pero los objetos parecen anormalmente delgados.

En principio estos son los factores que suponen algún tipo de condición adversa en la visualización, por lo tanto se deben medir y evaluar para establecer valores límites que garanticen un servicio apropiado.

1.6.1.3 Métodos de evaluación

Los métodos de evaluación conocidos son los que se han utilizado a la televisión en general. Existen dos clases de evaluaciones subjetivas. En primer lugar, hay evaluaciones que determinan la calidad de funcionamiento de sistemas bajo condiciones óptimas, las que se denominan evaluaciones de calidad. En segundo lugar, hay evaluaciones que determinan la capacidad de los sistemas de mantener la calidad en condiciones no óptimas que se relacionan con la transmisión o emisión. Éstas se denominan evaluaciones de degradación. Se debe elegir cuál de estas evaluaciones sugieren una correcta aplicación a la TV3D, así por ejemplo si se dispone de una imagen de referencia, puede utilizarse el método de escala de calidad continua de doble estímulo o el método de escala de degradación de doble estímulo. Si no se dispone de ninguna referencia, puede utilizarse el método de juicio categórico para identificar los méritos de los sistemas estereoscópicos. Los métodos de evaluación de factores específicos de los sistemas de televisión estereoscópica requieren estudios adicionales.

1.6.1.3.1 Método de escala de calidad continua de doble estímulo DSCQS

El método DSCQS se ha utilizado con éxito para evaluar la calidad de imagen subjetiva en imágenes monoscópicas. La experiencia ha demostrado que este método es adecuado también para la evaluación de imágenes estereoscópicas y puede adaptarse fácilmente para medir otros atributos de la imagen además de la calidad, tales como la nitidez y profundidad percibida. Este método se ha adaptado fácilmente a la medida de otros atributos además de la calidad subjetiva de la imagen. Ello se ha conseguido realizando cambios específicos en las instrucciones dadas a los observadores. Por ejemplo, el método se ha adaptado a medir la nitidez percibida y la impresión general de la profundidad de las

secuencias de imágenes estereoscópicas. En cada sesión sólo se midió un atributo (la calidad subjetiva de imagen percibida, la nitidez percibida o la profundidad percibida).

1.6.1.3.2 Resultados experimentales del método DSCQS en imágenes estereoscópicas

Se han recolectado algunas experiencias de visualización en televisión estereoscópica que han partido de experimentos con condiciones específicas, en particular para este caso se tienen las siguientes condiciones iniciales:

- La visión del ojo derecho de las secuencias de imágenes estereoscópicas se someten a un filtrado paso bajo a tres niveles: no filtrado, resolución mitad y resolución de un cuarto.
- Las parejas de imágenes izquierda y derecha de una secuencia estereoscópica se visualizan a 120 Hz utilizando un método secuencial en el tiempo, sobre una pantalla de visión directa de 29 pulgadas (74 cm) o sobre un retroproyector de 65 pulgadas (165 cm).
- Las imágenes de los ojos izquierdo y derecho fueron entrelazadas temporalmente y visualizadas en sincronía con la apertura y cierre de un par de cristales obturadores de cristal.
- Los obturadores de cristal líquido tenían una transmitancia de aproximadamente el 30% y tiempos de respuesta de 0,2 ms y 2,8 ms de cierre y apertura respectivamente.
- La distancia de observación es 4 H (H es la altura de la pantalla).

Con estas consideraciones se realizó el ensayo en el cual se muestra una imagen de referencia y a continuación la de prueba para medir los diferentes puntos en cuestión.

La Figura 1.13 muestra estos resultados para diferentes tipos de resolución en vertical y en horizontal; en el gráfico el eje Y indica la puntuación negativa de referencia de la prueba. Una puntuación cero indica que la secuencia de prueba fue puntuada igual que la secuencia de referencia estéreo no procesada. Una puntuación negativa significa que la secuencia de prueba fue puntuada más bajo que la secuencia de referencia.

Resultados del estudio ilustrativo utilizando el método DSCQS

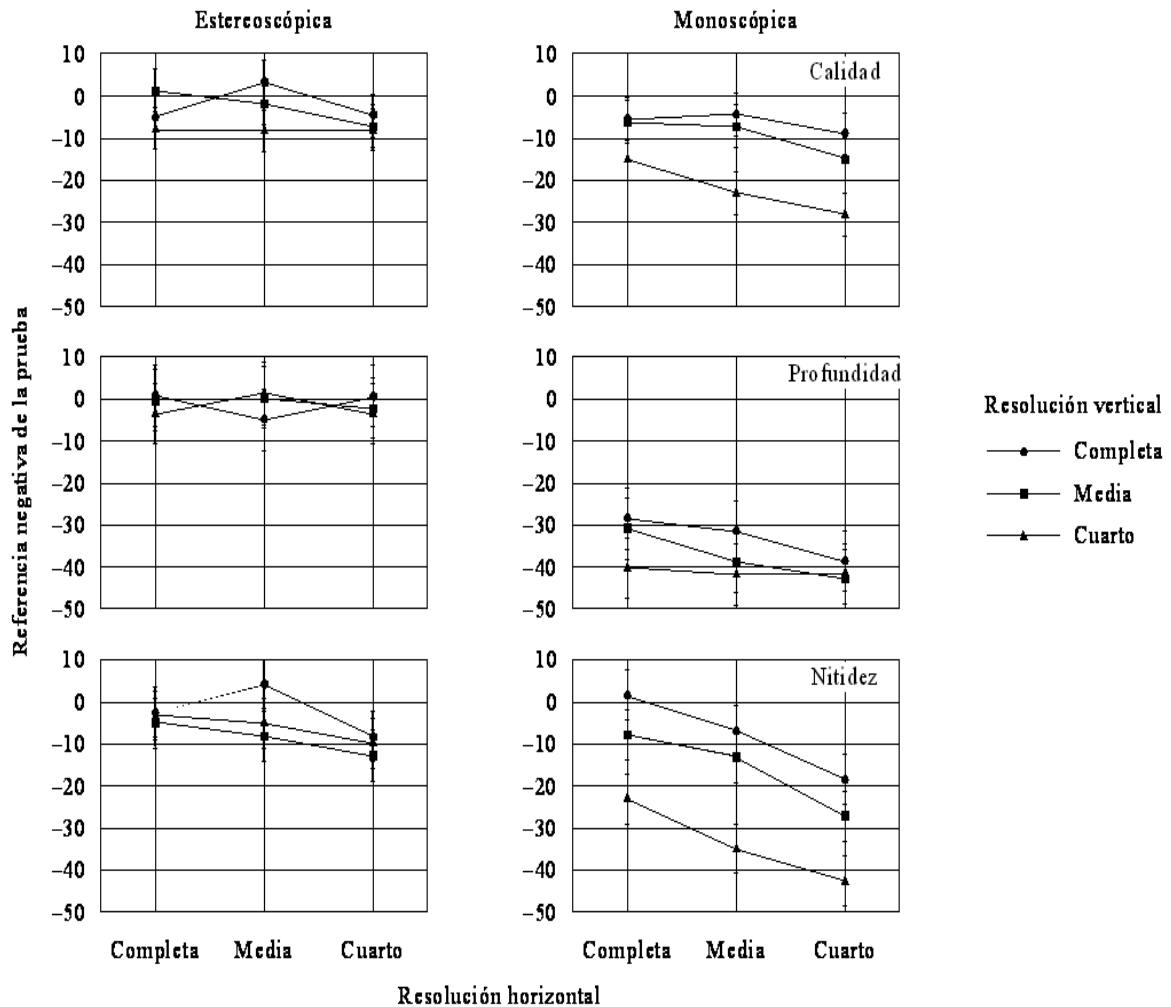


Figura 1. 13 Resultados del estudio ilustrativo utilizando el método DSCQS

1.6.1.4 Condiciones de visualización

Se pueden tener muchos efectos anormales en la presentación de imágenes estereoscópicas, el tamaño de los objetos es comúnmente mal interpretado por el ojo humano, en especial cuando las imágenes se acercan o se alejan, una medida para evitar ello es utilizar pantallas grandes debido a que el observador con esto será menos consciente de estas diferencias de profundidad. El ojo humano enfoca un objeto tomando como referencia su distancia. La convergencia es controlada automáticamente por el ojo. Por lo tanto, naturalmente no existe inconsistencia entre acomodación y convergencia. Sin embargo, cuando se visualizan imágenes estereoscópicas, el enfoque se fija

siempre en la pantalla, con independencia del punto de convergencia que se obtiene de la disparidad de las señales. Es decir el enfoque escapa del control del observador y dado que la convergencia está en la imagen, puede surgir una visualización de difícil acomodación.

El valor mínimo de profundidad de campo del ojo humano se ha establecido en $\pm 0,3 D$ (dioptrías; valor inverso de la distancia (m))²⁸. El objeto será enfocado correctamente cuando este encuentra a $\pm 0,3 D$. Dado que muchos programas de televisión incluyen imágenes a una distancia infinita (es decir, $D=0$), se considera que la profundidad que debe visualizarse con sistemas estereoscópicos se encuentra en el rango de 0 a 0,6 D. Por lo tanto, la distancia de visualización óptima es 0,3 D, es decir, 3,3 m.

Para los aspectos geométricos de la cámara (separación de la cámara, ángulo de convergencia de la cámara, longitud focal de las lentes), la resolución del sistema es importante para determinar las condiciones de visualización (tamaño de la pantalla). En el caso de HDTV, cuando se mira a la distancia de visualización normalizada de 3 H, la distancia de 3,3 m. corresponde a una pantalla de 90 pulgadas (229 cm). En el caso de la televisión de definición estándar (SDTV), cuando la distancia de visualización es la normalizada de 6 H, dicha distancia se corresponde con una pantalla de 36 pulgadas (91 cm). Utilizando un sistema de HDTV estereoscópico se realizó una evaluación subjetiva de la relación entre el tamaño de la pantalla y la percepción de profundidad, resultando que la percepción de profundidad más natural se obtuvo con una pantalla de 120 pulgadas (305 cm), lo cual da como resultado una distancia de visualización de 2,2 H (Sistemas poco prácticos para usuarios en el hogar).

1.6.2 CONSIDERACIONES DE IMÁGENES DE TV ESTEREOSCÓPICAS BASADAS EN SEÑALES L+R²⁹

De todo lo anteriormente analizado está claro que el servicio de TV estereoscópica tiene varias técnicas posibles para su implementación. Pero de lo estudiado sobresale que para lograr un sistema estéreo de video es

²⁸ Hiruma y Fukuda, 1990

²⁹ Análisis según la Recomendación UIT-R BT.1198, año 2000.

imprescindible la utilización de dos señales visuales, una para cada ojo, es así que se concluye, que la forma de llevarlo a cabo es acorde a esa técnica. Además el sistema debe garantizar que no exista fatiga como en la televisión monoscópica y que se mantengan los parámetros de la calidad de la imagen.

Otro aspecto de gran importancia es que, al considerarse dos señales la etapa de codificación debe garantizar el poder procesar este par de imágenes. Y que para efectos de transmisión el ritmo binario se pueda reducir para no crear un sistema con requerimientos difíciles de cumplir. También cabe mencionar que al usar señales L+R, se brinda versatilidad en cuanto al formato para aplicar, se puede hacer en sistemas polarizados, cromáticos o secuenciales; para cualquiera de las soluciones se debe procurar evitar sistemas incómodos al usuario y que demanden una capacidad excesiva en el canal de comunicaciones.

Los avances determinan que no solamente el sistema que separa las señales para cada ojo es posible de realizarse, se pueden contar con aquellos que convierten una señal de 2D a 3D y los que combinan información de la imagen plana con la de profundidad, que se obtiene mediante varias técnicas, en estos otros sistemas siempre se debe buscar mejorar en calidad y comodidad además de evitar requerimientos difíciles de cumplir.

1.6.3 REQUERIMIENTOS DE COMPATIBILIDAD Y VISUALIZACIÓN PARA UN SISTEMA DE TV ESTEREOSCÓPICO

Es importante que la implementación de esta funcionalidad para la televisión no cause retroceso en los avances ya logrados y que se convierta atractivo para el público, esto se logrará solo si no hay transiciones bruscas de tecnología y se guarda compatibilidad con sistemas anteriores. También es importante empezar a buscar una homologación en los equipos, en la producción y en elegir las técnicas apropiadas para evitar que el sistema sea costoso y poco conveniente. Al contar con un marco referencial claro y definido como el ISDB-T, la línea de estudio debe procurar mantener los objetivos de este estándar, es decir aplicabilidad, compatibilidad y el uso de las tecnologías que se describen sin tener que hacer considerables modificaciones a lo ya establecido; de ahí que es indispensable el análisis y estudio minuciosos de las posibilidades para optimizar la solución y

cumplir las metas de la mejor forma. Entre los requisitos que se deben cumplir para estos sistemas se tiene:

- Que la señal estereoscópica debe contener también la información monoscópica para guardar compatibilidad con receptores anteriores.
- Es obligatorio que la nueva señal mantenga y mejore los parámetros de calidad que se tienen actualmente en la televisión.
- La calidad y nitidez de la imagen debe ser tal que se elimine cualquier tipo de fatiga o molestia visual para el espectador.
- Con el avance de la tecnología las pantallas planas son las que permiten tener la impresión de profundidad, y debido a que la calidad debe superarse, las nuevas tecnologías de pantallas deben ser las más utilizadas.
- El color de una imagen no debe alterarse, por lo tanto los sistemas anáglifos cromáticos que degraden la tonalidad de la escena quedan descartados al interferir en la naturaleza de las imágenes.
- Con el desarrollo de la computación y los juegos en 3D, las tecnologías LCD se vuelven cada vez más accesibles, lo que sugiere que los sistemas 3DTV usen este principio o el polarizado que es un poco menos costoso.
- El procesamiento de la señal debe ser obligatoriamente digital: codificación, transmisión y todo proceso relacionado; solo esto garantiza la verdadera compatibilidad y flexibilidad del sistema.
- La producción de video estereoscópico no debe suponer procesos mucho más complejos que la del video monoscópico, porque en ese caso sería un sistema poco conveniente.
- El sistema en su conjunto debe presentar una alternativa atractiva al usuario, es por ello que el enfoque para generar contenidos debe cambiar y mejorar, apuntando a explotar al máximo el nuevo servicio.
- Se debe procurar en lo posible establecer formatos comunes en pantallas, prácticas de producción determinadas y normas de calidad suficientes para sistemas exitosos.

- Los nuevos sistemas de transporte de esta señal deben mejorar la inmunidad al ruido en condiciones topográficas adversas, de no lograrlo se deben mantener los parámetros que el estándar ISDB-T define.
- Las técnicas que sirven para desarrollar la televisión estereoscópica deben ser compatibles y ajustarse a las normas existentes, teniendo en cuenta que los estándares establecen la posibilidad de incorporar nuevos servicios y dan la flexibilidad pertinente.

CAPÍTULO 2

ADQUISICIÓN DE IMÁGENES ESTEREOSCÓPICAS Y TRANSMISIÓN DE LAS SEÑALES DE VIDEO 3D PARA EL ESTÁNDAR ISDB-T

2.1. ADQUISICIÓN / GENERACIÓN DE CONTENIDOS 3D

Existen diversas técnicas para obtener un cuadro estereoscópico válido para la transmisión en ISDB-T, el punto de partida para ello es identificar los métodos existentes para lograrlo, y en cada método, los procesos correspondientes. Se pueden clasificar en dos clases a esta generación de contenidos, la primera es la creación de video animado, a través de la producción por computadora; aquí se incluyen todos los métodos que requieren sistemas informáticos en los que no se involucra una cámara de video, y a todas luces es una técnica que solo depende de las herramientas informáticas disponibles; es por ello que en el cine la producción de películas animadas 3D es relativamente sencilla y abundante.

La adquisición es la primera etapa en el proceso de crear y transmitir señales de 3DTV; la cadena a la que el video estereoscópico debe someterse se describe en la Figura 2.1

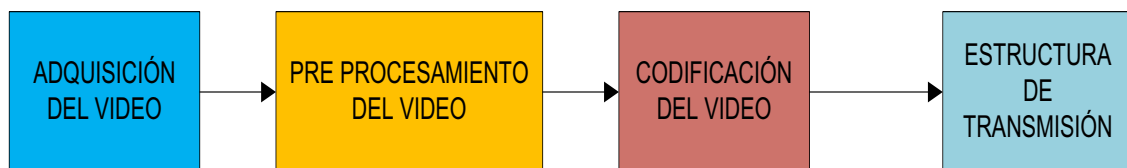


Figura 2. 1 Cadena de transmisión del video estéreo.

Dado que en televisión los contenidos en su mayoría representan a la realidad, y por tanto son captados a través de una cámara de video; se vuelve importante analizar cómo adquirir estas imágenes que por naturaleza son tridimensionales. En el capítulo 1 se establecieron ciertos lineamientos y requerimientos que son importantes considerar, así como aspectos fisiológicos del ojo que el transductor

estereoscópico debe emular para una correcta producción de imágenes estereoscópicas para televisión.

Se han propuesto y desarrollado algunas técnicas para adquirir video 3D y hoy en día las cámaras de TV realizan esta adquisición con formato propietario; existen algunos estándares y normas existentes que ayudan a definir algunos procedimientos para la adquisición pero nada estrictamente definido.

Los diferentes formatos para la codificación, transmisión y visualización hacen que existan muchas formas de adquirir video 3D y de procesarlo. Del capítulo anterior se desprende la conclusión de que para crear una imagen estereoscópica es necesario al menos tener información de imágenes L+R, lo que supone que se necesitan dos lentes para captar estas informaciones; sin embargo los nuevos métodos proponen algunas otras soluciones.

Es necesario especificar que la digitalización de la imagen como tal no varía a lo expuesto en el capítulo 1, es decir se muestrea de igual forma señales de luminancia y crominancia. Lo nuevo en el caso de la TV3D es el generar en algunos casos nuevas señales y en otros casos aumentar las veces que se realiza el muestreo, pero como tal los principios de adquisición de una cámara digital se mantienen. A continuación se exponen, las técnicas que utilizan y cómo el transductor debe actuar y digitalizar esta información de video.

2.1.1 CONVERSIÓN OPTO-ELÉCTRICA DE LA ESCENA

Existen varios modelos que permiten obtener la información 3D, y todos ellos en principio captan la escena, ya sea de una o varias vistas. Todos los cuadros captados requieren el mismo proceso optoelectrónico para la señal visual. Los dispositivos que permiten esta adquisición son detectores del tipo CCD, los cuales tienen una composición interna que ha evolucionado con la tecnología; la Figura 2.2 muestra lo que típicamente compone internamente a un detector del tipo CCD (charged-coupled device) desde el punto donde incide la luz hasta el punto en el cual se obtiene una diferencia de potencial correspondiente a las variaciones de luz expuestas.

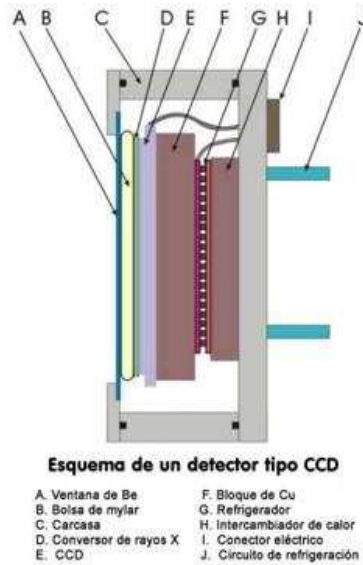


Figura 2. 2 Esquema del detector CCD

La imagen atraviesa los componentes químicos y mecánicos del detector para que la variación de las moléculas involucradas permita el paso de la luz y se obtenga una señal eléctrica que represente a la imagen, en la adquisición 3D, los detectores CCD existen en mayor número (al menos el doble en el caso estereoscópico) y con una escala de integración más alta.

Después de esta conversión la imagen debe digitalizarse, este proceso dependerá del formato y del tipo de compresión aplicado; estos factores serán detallados posteriormente.

Los dispositivos encargados de captar constan de componentes que configuran bloques funcionales para los siguientes procesos: adquisición, digitalización, codificación y transmisión.

La variedad de formatos hace que los dispositivos que adquieran el video varíen en sus funcionalidades, y los procesos de digitalización, codificación y transmisión, se hagan de forma distinta.

Más adelante se exponen las técnicas existentes para cada proceso, teniendo en cuenta que los dispositivos usan una o varias de las técnicas expuestas, y que contarán con los componentes para realizar cada uno de estos procesos.

2.1.2 ADQUISICIÓN A TRAVÉS DE 2 LENTES (CÁMARA ESTEREOSCÓPICA)

Este método se presenta como la solución elemental al problema de la adquisición, debido a que es la manera más próxima de emular el sistema de visión humano. Inicialmente se puede pensar en usar cámaras CCD comunes montadas en una estructura que permita la adquisición de imágenes; con ello es importante tomar en cuenta los aspectos geométricos y de calibración para captar de manera óptima y precisa las imágenes.

El desarrollo de algunos algoritmos permite que mediante 2 lentes se pueda obtener la siguiente información:

- Cuadros de Imágenes izquierda y derecha independientes.
- Información 2D del video más información de profundidad.
- Reconstrucción de múltiples vistas a partir de la información L+R.

Con esto se han determinado algunos formatos para el tratamiento de la 3DTV, y se han generado muchas opciones para formar el cuadro estereoscópico.

La tecnología de cada lente para captar la imagen es igual a la tradicional de los sistemas de video monoscópico, es decir cámaras CCD. Cada lente trabaja bajo sus principios para capturar la imagen independientemente, el proceso que el lente efectúa para que la imagen sea captada se puede apreciar en la Figura 2.3, el proceso descrito se conoce como un sistema óptico modelado por la ecuación de Gauss.

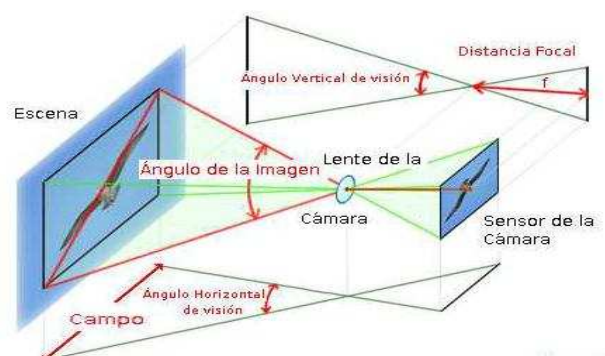


Figura 2. 3 Trayectoria de la luz en un lente de cámara

En la actualidad las cámaras de estudio basan su funcionamiento en modelos más complejos, sin embargo la idea principal se mantiene y permite comprender el funcionamiento de los transductores involucrados en la adquisición del video estereoscópico.

2.1.2.1 Aspectos geométricos y calibración de los lentes de las cámaras estereoscópicas.

La idea principal es asemejar los lentes a la visión humana, es por ello que la separación interlenticular debe ser de 65 mm. y tener un ángulo de convergencia igual al humano. Un requerimiento elemental para el diseño de un sistema de adquisición estereoscópico es la comprensión total del espacio real (espacio donde se encuentra el objeto) para reproducir el espacio del video estereoscópico.

De varios estudios se ha determinado que lo más importante para los lentes, es la disparidad binocular, el cual se torna determinante en la percepción de profundidad. Esta disparidad dependerá del ángulo de convergencia de los dos CCD's y de la distancia de estos al elemento visualizado³⁰. En principio todo par de lentes debería evitar la paralaje vertical y además estar libre de distorsión. La disposición de los lentes puede tener dos configuraciones:³¹

- Configuración en paralelo (con los ejes cámaras paralelos)
- Configuración Toe-in (con los ejes de las cámaras cruzados)

2.1.2.1.1 Configuración en paralelo

Para esta configuración los ejes de los lentes se alinean paralelamente, en este caso la convergencia se logra con el movimiento horizontal de los lentes; en esta configuración los elementos en el infinito son mostrados en el infinito debido a que se mantiene una separación horizontal constante H_c entre las vistas izquierda y derecha cuando estas son proyectadas (ver Figura 2.4).

³⁰ Ferre, M. and R. Aracil, "Interfaces for telerobotics studies on stereoscopic vision", in Proceedings of the 15th IFAC Triennial World Congress, Barcelona (España), pp. 1315-1320.

³¹ Configuraciones recomendadas en la ITU-R BT.2160-1 de Octubre de 2010, pp. 26-27

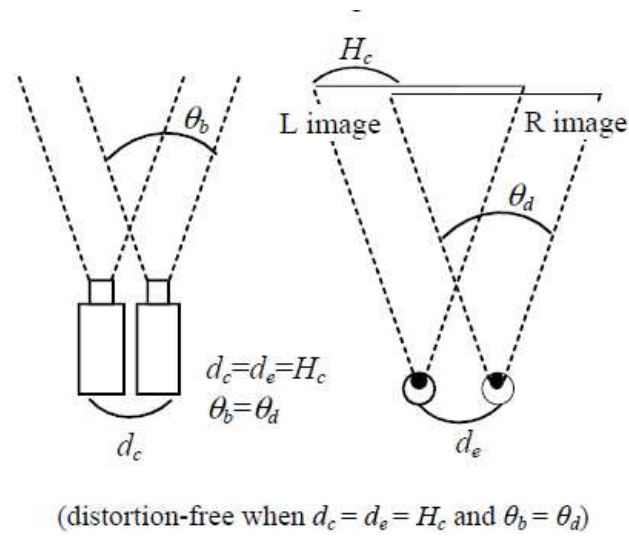


Figura 2. 4 Configuración en paralelo

Con la configuración en paralelo, para enfocar una imagen en un punto del espacio $W(x,y,z)$, los lentes se estructurarían como lo muestra la Figura 2.5.

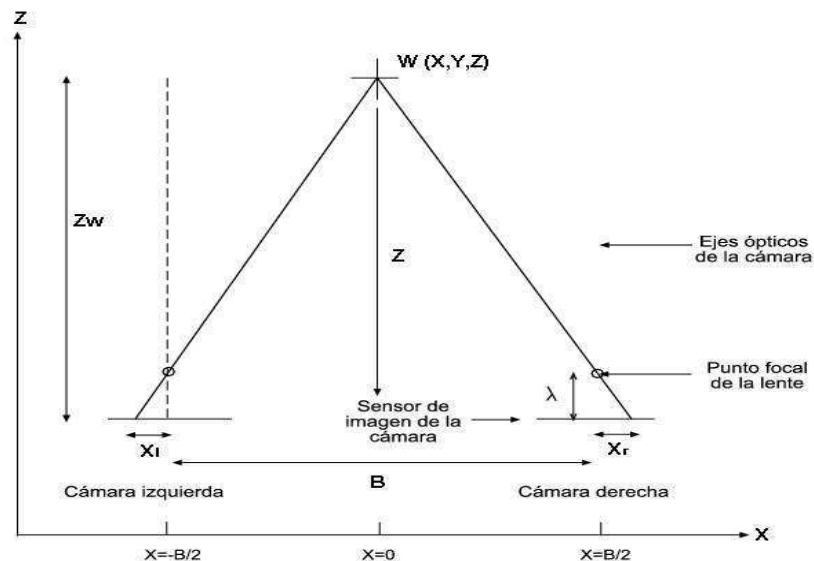


Figura 2. 5 Aspectos geométricos de la configuración paralela

El punto situado en W se proyecta a los lentes izquierdo y derecho, y la disparidad lenticular será función de B (distancia interlenticular), la distancia focal³² λ , y la distancia desde el eje de la cámara hasta W . Para capturar al objeto se debe efectuar un proceso de traslación de los lentes en el eje X, las coordenadas de

³² La distancia focal de una lente es la distancia entre el centro óptico de la lente y el foco cuando se enfoca al infinito.

proyección de la cámara, tanto izquierda como derecha respecto a W se exponen en las Ecuaciones 2.1 a 2.4.

$$x_{izq}(X, Z) = \lambda \frac{X + \frac{B}{2}}{\lambda - Z} \quad \text{Ec 2.1}$$

$$y_{izq}(Y, Z) = \lambda \frac{Y}{\lambda - Z} \quad \text{Ec 2.2}$$

$$x_{der}(X, Z) = \lambda \frac{X - \frac{B}{2}}{\lambda - Z} \quad \text{Ec 2.3}$$

$$y_{der}(Y, Z) = \lambda \frac{Y}{\lambda - Z} \quad \text{Ec 2.4}$$

Se observa que las ecuaciones 2.2 y 2.4 son iguales, lo cual demuestra que W se proyecta en el mismo punto “ y ” para las dos cámaras, lo cual es correcto porque la disparidad vertical debe ser nula; por otro lado la disparidad en X se expresa de acuerdo a las ecuaciones 2.5 y 2.6

$$disparidad_{horizontal} = x_{der} - x_{izq} \quad \text{Ec 2.5}$$

$$disparidad_{horizontal} = \lambda \frac{X - \frac{B}{2}}{\lambda - Z} - \lambda \frac{X + \frac{B}{2}}{\lambda - Z} = \lambda \frac{-B}{\lambda - Z} \quad \text{Ec 2.6}$$

2.1.2.1.2 Configuración Toed-in (ejes cruzados)

En esta configuración los ejes de las dos cámaras se intersecan en un punto (ver Figura 2.6). Para esta configuración si un objeto se sitúa en este punto de intersección la posición de su profundidad coincide con su posición en la pantalla, es por ello que en esta configuración se hace fácil identificar elementos detrás y delante de esta línea paralela a la línea base y que pasa por el punto de convergencia. Este es el método más usado.

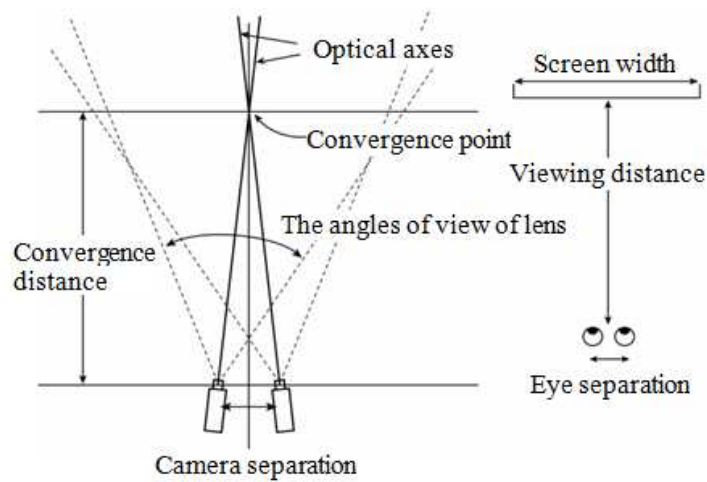


Figura 2. 6 Configuración Toed-In

Geoméricamente, en la configuración Toed-In (ver Figura 2.7) se vuelven muy importantes, la distancia inter-lenticular B y el ángulo de convergencia β , los cuales cumplen la relación matemática descrita en la ecuación 2.7

$$Z \cong \frac{B}{2 \tan \beta} \quad \text{Ec 2.7}$$

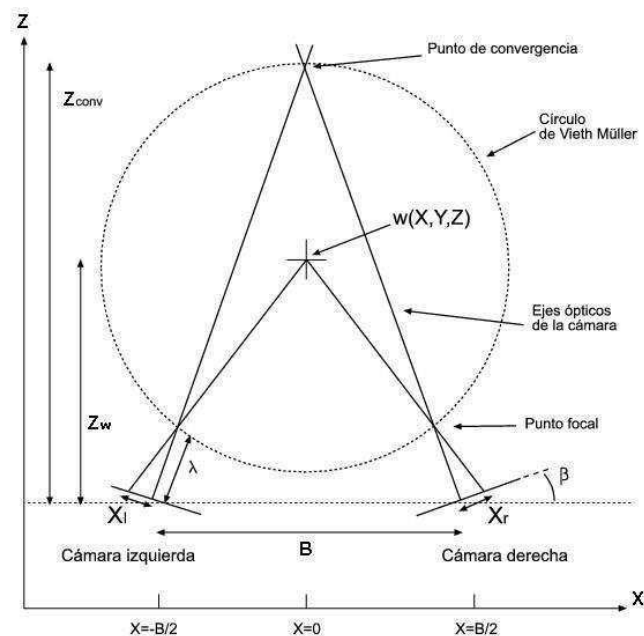


Figura 2. 7 Aspectos geoméricos de la configuración Toed-In.

En este caso el desarrollo matemático de la proyección del punto W en el plano de la cámara es más complejo que el caso anterior. En Toed-In los ejes de la cámara no son paralelos al eje Z , lo que obliga a realizar una traslación horizontal (eje X) y una rotación para el eje Y , estas variaciones se explican a través de matrices que representan las posiciones en los ejes, de la siguiente forma:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \xrightarrow{\text{traslación}} \begin{bmatrix} X + \frac{B}{2} \\ Y \\ Z \end{bmatrix} \quad \mathbf{Ec\ 2.8}$$

$$\begin{bmatrix} X + \frac{B}{2} \\ Y \\ Z \end{bmatrix} \xrightarrow{\text{rotación}} \begin{bmatrix} \cos \beta \left(X + \frac{B}{2} \right) - \sin(\beta)Z \\ Y \\ \sin \beta \left(X + \frac{B}{2} \right) + \cos(\beta)Z \end{bmatrix} \quad \mathbf{Ec\ 2.9}$$

Con esto las expresiones que determinan la proyección de perspectiva para la cámara izquierda y derecha son:

$$x_{izq}(X, Z) = \lambda \frac{\cos \beta \left(X + \frac{B}{2} \right) - \sin(\beta)Z}{\lambda - \sin \beta \left(X + \frac{B}{2} \right) - \cos(\beta)Z} \quad \mathbf{Ec\ 2.10}$$

$$y_{izq}(X, Y, Z) = \lambda \frac{Y}{\lambda - \sin \beta \left(X + \frac{B}{2} \right) - \cos(\beta)Z} \quad \mathbf{Ec\ 2.11}$$

$$x_{der}(X, Z) = \lambda \frac{\cos \beta \left(X + \frac{B}{2} \right) + \sin(\beta)Z}{\lambda + \sin \beta \left(X + \frac{B}{2} \right) - \cos(\beta)Z} \quad \mathbf{Ec\ 2.12}$$

$$y_{der}(X, Y, Z) = \lambda \frac{Y}{\lambda + \sin \beta \left(X + \frac{B}{2} \right) - \cos(\beta)Z} \quad \mathbf{Ec\ 2.13}$$

Con estas consideraciones, la calibración de las dos cámaras debe dar como resultado una adquisición de imágenes libres de distorsión.

El desarrollo matemático presentado describe el fundamento de la geometría de los lentes, actualmente se han desarrollado cámaras con mayores prestaciones que también describen un desarrollo matemático más complejo para su geometría. Ejemplo de esto es el modelo de cámara con detector de plano focal; este modelo propone una solución más eficiente a las distorsiones y fallas en la adquisición, el modelo se basa en el método de adquisición de imágenes estereoscópicas a través sistemas de imagen de onda milimétrica³³.

2.1.2.1.3 Percepción cuantitativa de la profundidad de la cámara estereoscópica

Para cualquiera de las configuraciones antes descritas, es muy importante determinar la distancia que representa profundidad de un objeto en una escena, es así que se definen dos parámetros importantes L_b y L_d , los cuales representan la posición de la profundidad en el espacio real y la posición de un objeto estereoscópico respectivamente; ambos parámetros se relacionan de acuerdo a la Ec. 2.14.

$$L_d = \frac{1}{\frac{1}{L_s} - \frac{a_1 \cdot a_2}{L_c} + \frac{a_1 \cdot a_2}{L_b} - \frac{H_c}{L_s \cdot d_e}} \quad \text{Ec 2.14}$$

Donde:

L_d : Posición de un objeto estereoscópico.

L_b : Distancia de disparo de la cámara.

L_c : Distancia de convergencia

L_s : Distancia de vista

d_e : Distancia interocular

d_c : Distancia entre los lentes.

H_c : Disparidad horizontal.

$a_1 = d_c / d_e$ y $a_2 = w / w'$.

w : Ancho de la pantalla

w' : Ancho de la pantalla virtual a la distancia del espectador.

³³ Paper IEEE "A Stereoscopic Model of Focal Plane Detector Array", Jung-Young Son, Seokwon Yeom, Dong-Su Lee, Kwang-Hoon Lee, and Min-Chul Park, JOURNAL OF DISPLAY TECHNOLOGY, VOL. 7, NO. 5, MAY 2011 p281

Se puede distinguir claramente que para el caso de la configuración paralela la distancia de convergencia es $L_c \rightarrow \infty$ y que la disparidad horizontal corresponde a la distancia de . Por otro lado para la configuración de ejes cruzados se observa que $H_c=0$; con estas consideraciones la Ec. 2.14 tiene distintas expresiones dependiendo el caso, tal como se describe en la Tabla 2.1.

<i>Configuración ejes paralelos</i>	<i>Configuración ejes cruzados</i>
$Ld = \frac{Lb}{a1 \cdot a2}$	$Ld = \frac{1}{\frac{1}{Ls} - \frac{a1 \cdot a2}{Lc} + \frac{a1 \cdot a2}{Lb}}$

Tabla 2. 1 Representación de la profundidad por configuración

Con estas especificaciones los lentes de las cámaras existentes tienen el lineamiento base para que en la adquisición no existan distorsiones y se procure aumentar la calidad.

2.1.3 ADQUISICIÓN A TRAVÉS DE MÚLTIPLES CÁMARAS (SISTEMA MULTIVISTA)

Sin duda la calidad de las imágenes 3D y la exactitud de los detalles de la misma son los objetivos principales en la adquisición, la solución natural para poder captar todo el volumen de una escena es ubicar n lentes que cubran toda la superficie, los cuales permitan obtener tantas vistas como sean necesarias para reconstruir una imagen volumétrica de gran calidad. Esta metodología se puede entender como la generalización de la adquisición de imágenes estereoscópicas que se realizan con más de un lente; a pesar de que se han desarrollado algunos algoritmos que generan múltiples vistas a partir de las captadas por un par de lentes, los múltiples lentes generan mayor calidad para la representación de la escena. A todas luces la implementación de este tipo de adquisición es más costosa y requiere una infraestructura más compleja; es claro también que no se requieren los aspectos geométricos antes descritos, al contrario se requiere precisión para la adquisición y un proceso de post-producción muy minucioso para que se obtenga una escena de alta calidad. Para la implementación de estos sistemas se utilizan lentes monoculares tradicionales dispuestos de tal forma que permitan obtener los detalles que proporcionen información de profundidad para

una escena. Estos sistemas, al ser complicados, quedan reservados para aplicaciones muy específicas y determinadas, tales como espectáculos deportivos y culturales, donde al espectador se le puede incluso brindar mayores posibilidades de interactividad. En estos sistemas la producción de imágenes es más compleja y robusta que en otros casos, ya que con muchas informaciones visuales los sistemas deben ser capaces de procesar gran flujo de datos, tener altas tasas de compresión y obtener variables útiles para el desarrollo.

Para que un sistema de TV, pueda soportar este tipo de aplicación las imágenes atraviesan el siguiente descrito en la Figura 2.8.

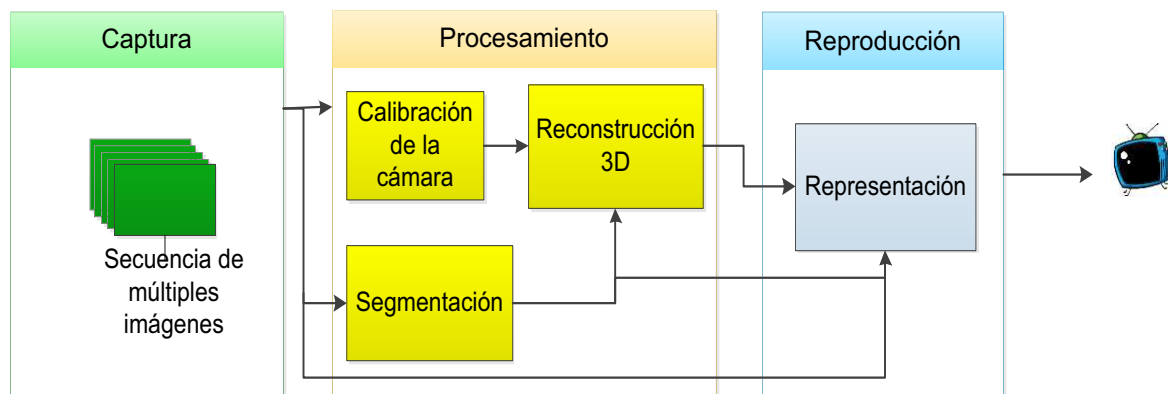


Figura 2. 8 Etapas multivista para un sistema 3DTV

La etapa de captura se desarrolla mediante adquisición sincronizada en el tiempo la cual se logra para todas las cámaras dispuestas, esta sincronización utiliza Genlock³⁴. Estas imágenes captadas por la cámara son almacenadas directamente sin compresión en el disco, para un procesamiento offline (fuera del proceso).

Estas cámaras deben permitir una calibración automática de los lentes de manera simultánea para evitar latencias por la calibración sucesiva de las n cámaras. Dado que la aplicación en espectáculos deportivos es en tiempo real, esta sincronización es exacta y debe responder en tiempo real.

³⁴ Genlock: Técnica común en video donde una referencia de video específica sirve para sincronizar todas las fuentes de vídeo que se dispongan.

La calibración de la cámara depende de los parámetros intrínsecos y extrínsecos de cada una de ellas incluyendo la distorsión de los lentes; las múltiples vistas son sometidas a una etapa de segmentación donde es preciso un procesamiento exacto de cada vista para que al aplicar los algoritmos de procesamiento generen resultados aceptables. Por otro lado la reconstrucción 3D se realiza después de la calibración y la segmentación.

El hecho de obtener mucha más información de imagen permite tener algunos resultados como video estéreo y video con libre punto de vista, el cual se da cuando una escena puede verse desde cualquier ángulo por la persona y esta percibirá una visión volumétrica de la respectiva vista elegida.

Se han investigado 2 metodologías para la representación de imágenes multivista los cuales son Interpolación y Reconstrucción.

2.1.3.1 Interpolación Multivista

La interpolación de una vista estima directamente la apariencia de la escena desde los puntos de vista originales, sin una reconstrucción explícita de la estructura de una escena 3D como un paso intermedio. Esto evita el tener la reconstrucción 3D como tal, pero se limita la representación solo a las vistas que estén entre las cámaras (ver Figura 2.9).

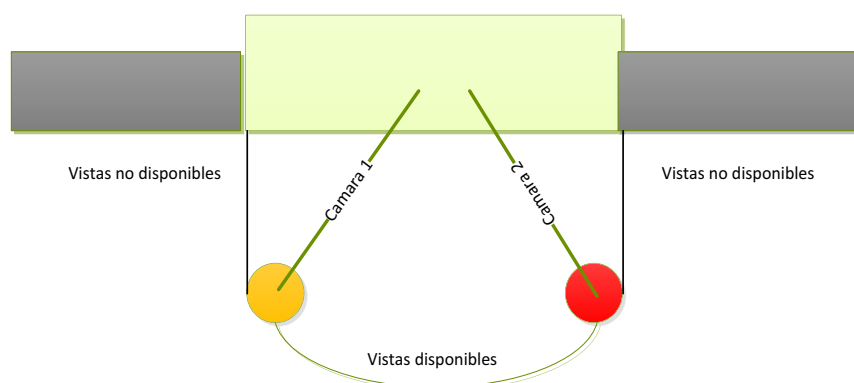


Figura 2. 9 Vistas disponibles en la interpolación Multivista.

En la interpolación se tiene la gran ventaja de reducir errores en la calibración, por lo que las fallas en la reconstrucción disminuyen también.

La calidad de la representación de las vistas depende del alineamiento que se aplica a los lentes para captar la escena.

2.1.3.2 Reconstrucción 3D

Este es el método más usualmente empleado para la representación de las vistas originales, teniendo en cuenta que las escenas captadas son dinámicas. Existen algunos métodos para la reconstrucción, como casco-visual, casco de la fotografía, estéreo, y optimización global de las figuras. El método de casco visual interseca los conos de la silueta de las múltiples vistas para reconstruir el volumen máximo ocupado por los objetos en la escena, esto requiere priorizar la segmentación del primer plano. El casco de la fotografía reconstruye a partir del volumen máximo de una fotografía que existe entre las múltiples vistas, en este caso existe la ventaja de no priorizar la segmentación del primer plano. La reconstrucción estéreo es usada para reconstruir escenas dinámicas, en este sistema se asegura la reconstrucción de la superficie que alinea las imágenes de las múltiples vistas reduciendo distorsiones en la representación de la escena. Para el método de reconstrucción global se combinan las técnicas volumétricas y las técnicas estéreo generando una reconstrucción más robusta.

La producción de un sistema multivista será exclusivo de determinadas aplicaciones, por lo cual no sería del todo aplicable a todas las estaciones de televisión; sin embargo se han publicado lineamientos importantes para producir estos sistemas, y para que se genere interés en algún sector y se empiece a generar producción de espectáculos multivista.

2.1.4 ADQUISICIÓN A TRAVÉS CÁMARA 2D + Z-CAM (TOF)

Este método es muy útil en algunos casos, dado que la información de profundidad permitiría generar el video estereoscópico. Se propone un sistema en el cual se capte el plano 2D con un lente normal y por otro lado se capte la información de profundidad de cada pixel de este plano, eso se hace a través de las conocidas Z-Cam o también llamadas TOF (Time of Fligth), que permiten generar información para dar profundidad a cada pixel.

Este método no es el único para obtener esta información, debido a que hoy en día también se puede extraer esta información de profundidad a través de la adquisición de imágenes L+R, por ello el usar Z-Cam para televisión es una opción más aplicable a la televisión digital. La calibración y el uso de estas cámaras es más complejo que el usar una cámara estereoscópica, la Z-Cam genera pulsos de luz infrarroja con diodos láser que son reflejados por los objetos que la cámara capta, al capturar esta reflexión el dispositivo calcula para cada pixel la distancia exacta a la que se encuentra y genera un mapa de profundidad, el cual no es más que el video original en escala de grises, en donde el blanco representa los pixeles más cercanos y el negro los pixeles con mayor componente de profundidad Z.

El mayor problema en estos sistemas, es que las 2 cámaras usadas deben tener resoluciones similares y ser compatibles para realizar asignaciones exactas de cada pixel con su información de distancia Z. De acuerdo a la precisión de la cámara, la escala entre blanco y negro tendrá mayor número de grises y se podrán reproducir más valores de distancia Z para representar la profundidad de cada pixel que tanto las cámaras 2D y la Z-Cam capten. Este problema de perspectiva que puede presentarse, se soluciona extrayendo la componente Z de estos pixeles. Con la información en escala de grises, se efectúa la codificación de cada pixel, y eso se logra asignando valores discretos de 1 ó 0 a las divisiones pixelares más pequeñas posibles y generando una secuencia de bits que refleje la distancia de profundidad de un objeto (ver Figura 2.10).

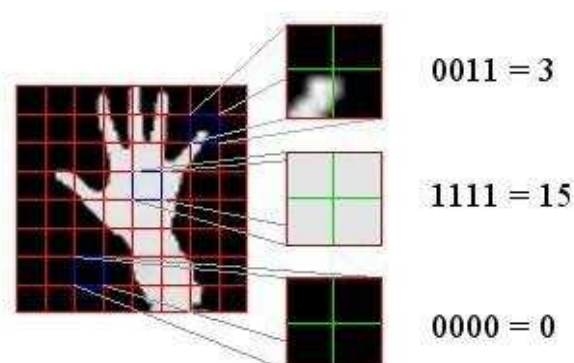


Figura 2. 10 Cuantificación del mapa de profundidad extraído por la Z-Cam.

Para construir un modelo consistente en 3D de la escena, la componente de la profundidad de la representación del objeto debe ser independiente de si el objeto está delante de la cámara o hacia el borde de la imagen.

Es necesario que para estos sistemas, en los cuales se inserta información de profundidad, se determine un estándar para que esta información pueda ser procesada por los equipos y pueda ser transmitida junto con la información del plano x, y.

Hoy en día los sistemas de TV tienden a usar esta información de profundidad (depth) para transmitir contenidos 3D, ya que requiere menos recursos que utilizar otra imagen que exija mayor capacidad de transmisión, sin embargo la forma de obtener esta información es variada y no necesariamente se puede hacer con una Z-Cam, se puede hacer también a través de dos lentes estéreo o incluso con procesamiento y estimación de profundidad de una señal de un solo lente (2D).

2.1.5 ADQUISICIÓN A TRAVÉS CÁMARA 2D (Conversión 2D-3D)

Esta técnica aunque resulte la que más se acopla a los sistemas que están en vigencia, requiere un procesamiento digital de la señal muy riguroso y complejo. Al no contar con la información real de profundidad, el emular o crear esta información disminuye calidad y precisión. A pesar de todo ello, este es el método más usado en la actualidad, dado que permite una compatibilidad total con los sistemas existentes, que lo único que deben incorporar es la etapa de conversión a 3D.

Esta conversión se puede hacer a dos niveles, en la producción y en la recepción; el último caso es lo que actualmente se tiene como 3DTV, la cual toma el cuadro en 2D, lo convierte a 3D y lo presenta en anáglifo polarizado o multiplexado en el tiempo, sin embargo cuando este proceso se realiza en la producción se puede tener mayor calidad.

En principio los algoritmos que permiten la conversión de 2D a 3D, crean un mapa de profundidad de una secuencia de video 2D. A continuación se expone el desarrollo de un algoritmo que permite la conversión, siguiendo una serie de

procesos independientes, se debe aclarar que no todas las secuencias de video pueden convertirse a 3D, el progreso del estudio de estas técnicas apuntan a que se intervenga manualmente lo menos posible al video.

2.1.5.1 Descripción del algoritmo de conversión

El algoritmo que se propone a continuación consta de 6 procesos (ver Figura 2.11), este ha sido muy usado para la producción de películas 3D, por lo que involucra resultados de buena calidad.

Este método se efectúa para cada tamaño y tipo de pantalla donde se reproduzca, es así que es necesario caracterizar el monitor que va a reproducir la imagen, esta caracterización permite evaluar el mínimo y el máximo valor de disparidad que se tiene para la imagen, posteriormente se realiza la estimación del movimiento basado en el algoritmo bidireccional KLT (Kanade-Lucas-Tomasi) basado en la segmentación del color.

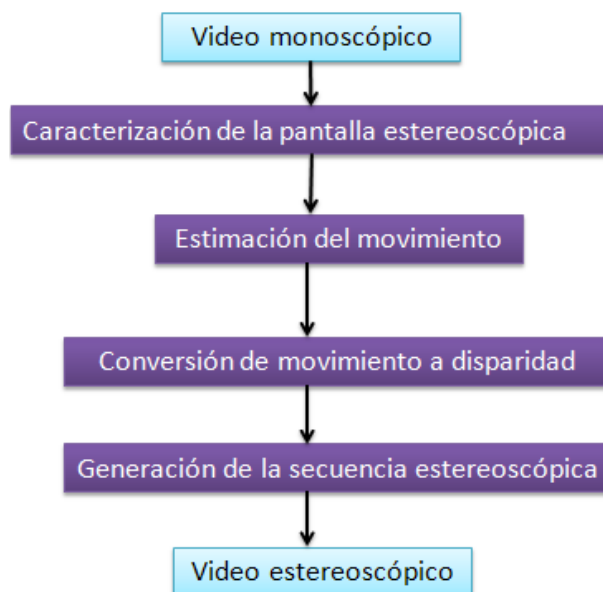


Figura 2. 11 Algoritmo de conversión 2D-3D.

La estimación del movimiento se realiza mediante 5 pasos: Segmentación del color, etiquetado, extracción de los límites, selección de las características y seguimiento bidireccional KLT. Esta estimación del movimiento permite obtener

valores de disparidad, es decir el mapa de profundidad de la escena, con lo cual se puede generar la información necesaria para la secuencia estereoscópica, y con ello obtener como resultado el video 3D para TV.

2.1.6 ADQUISICIÓN CON SISTEMA GENÉRICO PARA PRODUCCIÓN 3D

De las anteriores técnicas de adquisición se puede concluir que cada una ofrece información 3D a su modo y de una manera diferente, es así que desde la primera etapa los sistemas de TV digital deberían definir el tipo de adquisición que se debe usar. No obstante se puede establecer un sistema genérico que permitiría obtener toda la información que se conoce para producir 3D, el sistema que se propone es una solución más compleja que las anteriores, sin embargo no es inalcanzable y en algunos casos puede resultar una solución viable y conveniente.

De lo analizado, la información que es útil para producir 3D es: imagen izquierda, imagen derecha, información multivista e información de profundidad, las tres primeras no son más que cuadros captados por lentes estéreo, para la información de profundidad se puede utilizar la Z-Cam ó TOF, que ayudaría con esta percepción; con estas consideraciones se puede generar un sistema de adquisición genérico como el ilustrado en la Figura 2.12.

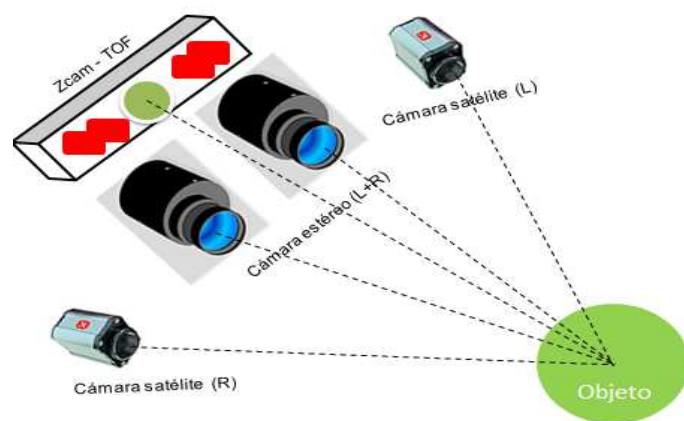


Figura 2. 12 Adquisición genérica de video 3D.

Se puede considerar a este sistema como una adquisición estéreo mejorada, porque sobre el par de lentes de la cámara estéreo se le da robustez con cámaras satélites que permiten tener algunas ventajas de los sistemas multivista,

generando un par de vistas más que pueden regenerar una imagen 3D más precisa. Además se incluye un módulo TOF opcional para reforzar al sistema con información de profundidad más completa, en esta se incluyen otros cuadros que permiten más precisión como son capas de oclusión de profundidad y de imagen.

Con esta configuración se tiene acceso a la siguiente información (ver Figura 2.13):

- Información del cuadro izquierdo y derecho
- Información de profundidad (L y R)
- Capa de oclusión de profundidad (L y R)
- Capa de oclusión de la imagen(L y R)

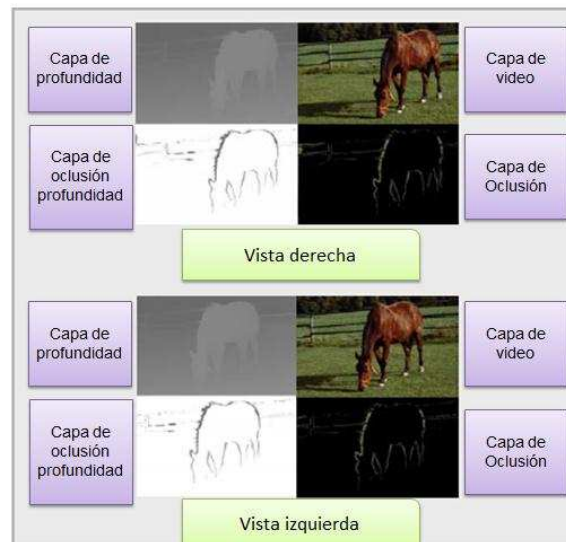


Figura 2. 13 Cuadros resultantes de la configuración genérica.

Se observa un total de 8 imágenes que permiten versatilidad en el procesamiento y a la vez, precisión y calidad en la reproducción. Para el sistema ISDB-T, se debe tomar en cuenta las restricciones en compresión y transmisión por lo que se debe procesar esta información de manera óptima para adaptarla a este estándar de TV digital. Más adelante se analizarán los métodos de compresión y formatos de transmisión óptimos para la TV3D, lo cual permitirá definir qué tipo de procesamiento es óptimo para los diferentes cuadros.

Todos los procesos de adquisición son la base de la producción del video, a esto se debe agregar que antes de codificar y transmitir estas señales, es necesario y muy útil añadir procesos de edición y efectos al video para corregir algunos defectos o dar alguna característica especial. Esta edición parte desde la convencional para televisión monoscópica, ajuste de brillo, color, tonalidad, etc. En el caso de 3DTV, se corrige la disparidad, la paralaje, el tamaño de los objetos etc.

2.2 FORMATOS DE VIDEO PARA 3DTV

Al analizar la adquisición del video, es notoria la versatilidad, y el número de opciones existentes para capturar las escenas y toda la información que de allí se puede obtener. Los formatos establecidos para la 3DTV han ido evolucionando desde los más simples hasta aquellos con procesos exclusivos y complejos. Un formato se puede definir como la descripción de la información que se adquiere y se procesa, así como la forma de obtener la tridimensionalidad en el video para el estándar de TDT.

El reporte ITU-R BT.2160-1, contiene una clasificación interesante de los formatos de TV3D, además de establecer a priori niveles de compatibilidad y generaciones de este servicio, los formatos descritos en la Tabla 2.2 están ubicados bajo el esquema de clasificación de este reporte UIT; posteriormente se realizará un análisis de cada uno para establecer la óptimos para el estándar ISDB-T, y con ello desarrollar el o los formatos que representen solución de manera detallada.

	<i>FORMATO</i>	<i>NIVEL DE COMPATIBILIDAD</i>
1° Generación de 3DTV(Se presenta a las imágenes en un perfil estereoscópico plano)	OCA (Optimized Chromatic Anaglyph): Anáglifo cromático optimizado	No requiere ninguna clase de equipos o receptores nuevos, solo las gafas (filtros) cromáticas anáglifos
	CSV (Conventional Stereo Video): Proces. independiente de las señales video con técnicas convencionales (Simulcast, Multi Stream).	Se requiere actualización básica y pantallas 3D

	CFC (Conventional HD Frame Compatible): Compatible con el cuadro HD (Imágenes L,R en el cuadro HD)	No requiere actualización al Set top box HD tradicional, pero requiere pantallas 3D
	CFC+ (Conventional HD Frame Compatible plus): Compatible con el cuadro HD y con MPEG-4 H.264 resolución mejorada (AVC,SVC)	
	V+D (Video + Depth) : Video más información de profundidad.	
	MVC (Multiview Video Coding): Imágenes izquierda y derecha codificado con MVC	
2º Generación de 3DTV(Se presenta a las imágenes en un perfil multivista)	MVD (Multiview Video Coding + Depth) : Formato MVC+D Codificación multivista + profundidad	
LDV (Layered Depth Video): Formato LDV Video en profundidad por capas		
DES (Depth Enhanced Video) : Formato DES, Profundidad estéreo elevada		
3º Generación de 3DTV(Se presenta a las imágenes en un perfil volumétrico del objeto)	<p>Aún no existen formatos desarrollados, sin embargo la holografía cumpliría este objetivo</p>	

Tabla 2. 2 Formatos de 3DTV.

En definitiva no importa cuál formato se utilice para transmitir la información 3D en televisión digital, el desarrollo de la tecnología de displays permite convertir esa información a dos o más cuadros que se pueden visualizar por cualquier método que el fabricante elija.

Lo más importante es encontrar un formato óptimo y que además se ajuste fácilmente al estándar ISDB-T, este análisis debe considerar que este estándar tiene alta calidad y movilidad, por lo tanto el formato elegido debe mantener estas exigencias y no representar cambios profundos en la infraestructura de este estándar.

2.2.1 FORMATO ANÁGLIFO CROMÁTICO OPTIMIZADO (OCA).

En el capítulo 1 se describió a esta técnica como una de la más antiguas y con mayores restricciones por su baja calidad; es notorio que en este sistema se tiene completa compatibilidad ya que este se reproduce en las pantallas convencionales y el procesamiento para generar 3D - OCA es sencillo; OCA se lo puede considerar como un sistema admisible ya que ha sido optimizado para no perder calidad en el color ni grandes alteraciones en el video original, es de fácil manipulación y de bajo costo. Este sistema no representa ninguna variación en la infraestructura de transmisión del estándar ISDB-T ya que la información se envía como un cuadro SD o HD. Para generar este tipo de imágenes se requieren las dos vistas de una escena por lo cual el método óptimo de adquisición será el descrito en 2.1.2, el sistema para el formato OCA se describe en la Figura 2.14.

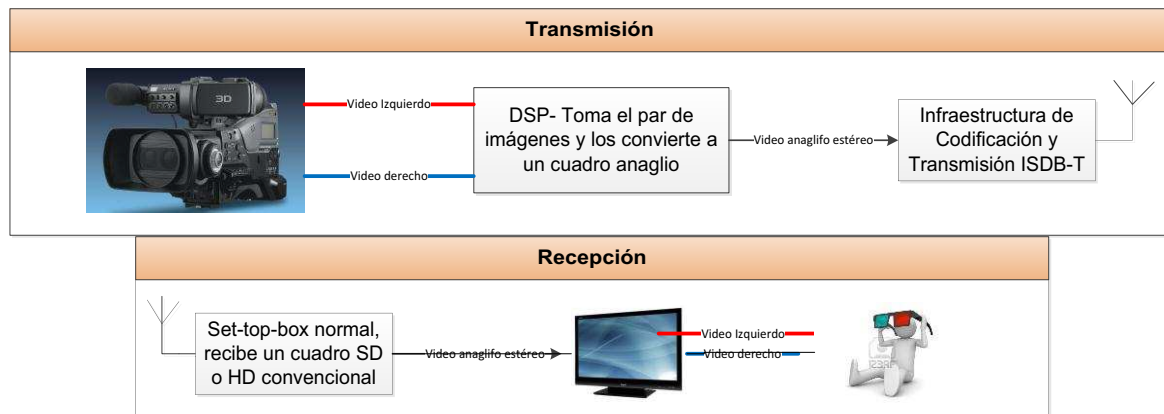


Figura 2. 14 Resumen del sistema anáglifo en ISDB-T.

Si bien se piensa que este sistema tiene muchas falencias de calidad, las cuales no justifican la gran ventaja de la compatibilidad, algunas empresas han dedicado estudios minuciosos para el desarrollo de sistemas anáglifos cromáticos de altísima calidad; poniendo gran atención en procesar las dos vistas de manera

precisa y priorizando la calidad sobre todo. Una de estas empresas es ColorCode3D[®], la cual cuenta con una gama completa de productos para este sistema, en la Figura 2.15 se presentan una imagen, extracto de un video anáglifo, que demuestra la calidad que este sistema puede llegar a tener:



Figura 2. 15 Imagen Anáglifo de alta calidad.

Hoy en día el sistema anáglifo optimizado se utiliza en receptores convencionales en los cuales se hace la conversión 2D-3D y se presenta al video en este formato.

2.2.2 PROCESAMIENTO INDEPENDIENTE DE LAS SEÑALES DE VIDEO (CSV).

En este formato se procura tratar y procesar las señales que puedan existir para 3D de manera independiente, es así que si se tienen señales izquierda y derecha, cada una de ellas se someterá a su estándar de compresión, a su codificación y será incluso transmitida considerando que es una señal 2D, esto supone efectivamente el gasto del doble de recursos para una información 3D, por lo tanto el sistema se limita a la existencia de dos imágenes. Dentro de este formato podemos encontrar la técnica que desembocará en Dual-Stream, y en algunos casos considerarse un multi-Stream, y para otros Simulcast. En ambos las señales que se adquieren independientemente, tienen su línea de procesamiento independiente dentro del sistema y son transmitidas o bien sincronizadamente

multiplexadas o de manera simultánea, ocupando el doble de recursos técnicos y en recepción se utiliza dos plataformas independientes una para cada señal.

Existe una variante al formato CSV llamada MRS (Mixed Resolution Stereo), en este caso alguno de los cuadros L ó R reduce su resolución (típicamente al 25%) y acompaña al otro cuadro que está en su resolución completa. Este formato, alternativo a CSV, ahorra de manera considerable los recursos de transmisión, ya que no se necesita el 200 % de recursos sino alrededor del 130% de lo utilizado en transmisiones monoscópicas.

Este formato representa la solución natural a la producción y transmisión de señales estéreo de video, el formato es conveniente y aplicable para 2 vistas solamente, izquierda y derecha. A cada una de las vistas se le aplicarán codificadores que se usan actualmente, como los estandarizados H.264. La desventaja principal de este formato es que no se utiliza en lo absoluto la relación y la dependencia entre las vistas, que para otros formatos representa un elemento vital al momento de reducir los recursos técnicos utilizados. Por otro lado la ventaja existente es la fácil recuperación de una de las vistas para la compatibilidad 2D.

Dado que no se utilizan técnicas avanzadas para la codificación de este video se conoce a este como CSV, o video estéreo convencional.

2.2.2.1 Concepto de Dual Stream

En este caso se generan dos flujos de las señales y posteriormente se someten a una etapa de multiplexación donde se estructura una secuencia de transmisión y para que así se genere el par de cuadros de video de manera ordenada, este formato es compatible con 2D debido a que si solo ese desea un cuadro, es fácil descartar el otro y diferenciar cuál es el que se desea ver. Inicialmente se contempla el dual Streaming como el enviar los mismos flujos para dos fines diferentes, pero si se trata de señales estereoscópicas, se aprovecha esta condición para enviar en dos flujos ligeramente distintitos y que estos sean procesados de manera independiente. Este formato es bastante aplicable con

técnicas de compresión y transmisión actuales, ya que se maneja cada señal de manera independiente.

Si bien es cierto el dual Streaming es concebido como una estructura de transmisión de datos, está muy ligado al formato de adquisición de las imágenes que serán enviadas; es por ello que el Dual Streaming establece el formato para el estéreo convencional, es decir si se plantea transmitir dos señales de video, lo natural es que se adquieran dos señales de video y que estas se mantengan independientes en su procesamiento (ver Figura 2.16).

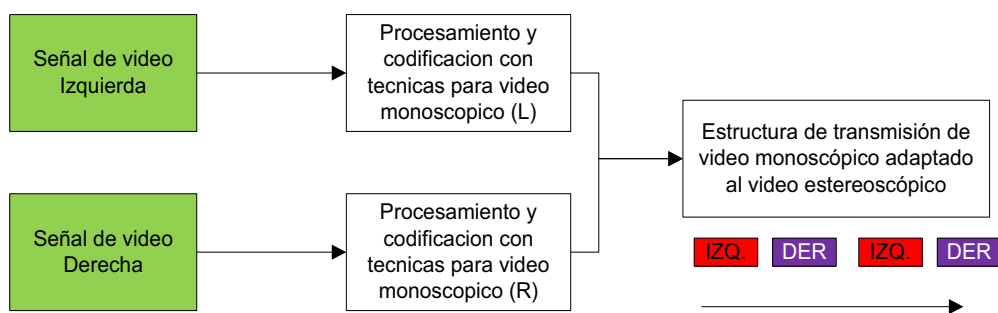


Figura 2. 16 Descripción del Dual-Streaming para TV3D.

Como se observa en la figura 2.16, se estructura un empaquetamiento de las dos señales en cuestión, para que estas puedan ser transmitidas y a cada una de ellas tengan un procesamiento independiente. La estructura de transmisión es común para las señales de video L y R.

2.2.2.2 Transmisión Simultánea (Simulcast)

Aunque el procesar las señales independientemente es lo más simple, es recomendable simplificar el sistema, el esquema de transmisión simultánea plantea la existencia de dos señales individuales L y R que serán transmitidas en diferentes canales disponibles (ver Figura 2.17), es decir como dos señales diferentes; sin duda los recursos necesarios son elevados pero la infraestructura para video monoscópico no se altera, solo se aumenta.

Cabe mencionar que es un sistema poco viable, ya que ocupa el doble de ancho de banda, y que en recursos de HD estarían ocupados más del 80% de los 13 segmentos disponibles.

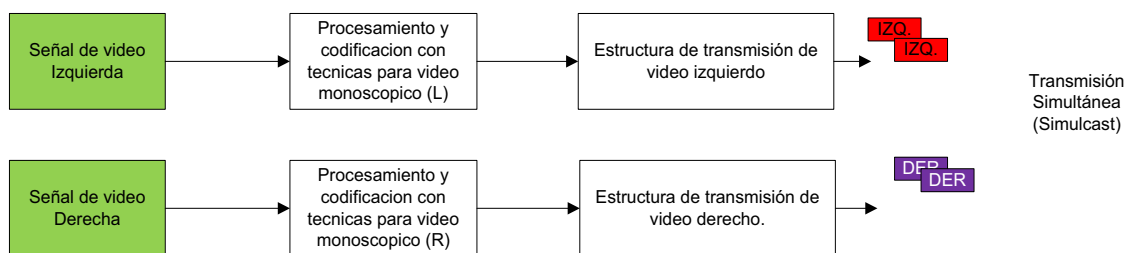


Figura 2. 17 Descripción del Simulcast para TV3D.

2.2.3 FORMATO COMPATIBLE CON EL CUADRO HD (CFC)

El conjunto de formatos que se expone a continuación, ha tenido una evolución significativa, y comenzó por la idea más sencilla, que fue enviar los cuadros izquierdo y derecho por separado de manera alternada es decir como una transmisión “doble” de las escenas. Después siguió evolucionando hasta obtener mayor eficiencia ubicando a las dos imágenes en el mismo campo. Dado que la base de este formato es el cuadro HD, el estándar HDMI propuso la idea de implementarlo y definió formatos bases y formatos alternativos, dejando muchas opciones para realizar la producción de imágenes 3D. Además se propone una efectiva solución a la compatibilidad, el fundamento básico dice que estos formatos se basan en la existencia de 2 vistas (izquierda y derecha) que aprovechando el número de líneas del cuadro, se ubican dentro del mismo plano. El hecho de que se ubiquen dos imágenes en el lugar que ocuparía una supone una pérdida de calidad, sin embargo la evolución de las técnicas de compresión ha permitido que se mantenga la buena calidad y se comprima con estándares avanzados. Entre los formatos avanzados HDMI, la versión 1.4a establece que se puede sustituir la información de uno de los cuadros por información de profundidad, haciendo que la adquisición de esta información pueda ser aprovechada. Este formato se presenta como una solución sencilla a la compatibilidad, ya que se aprovecha las cualidades del canal de alta definición para ubicar en él, el máximo de información posible. HDMI 1.4 establece las normas para la comunicación de imágenes HD a través de la interfaz con el mismo nombre, y en su versión ampliada 1.4a establece los formatos para ubicar imágenes L+R en el cuadro HD. Para el estándar ISDB-T se puede considerar un

canal HD que ocupa 6 segmentos del canal, y por lo tanto el canal de 3DTV ocuparía los mismos recursos.

El hecho de introducir dos imágenes en un cuadro HD implica la pérdida de resolución, lo cual es una desventaja; por otro lado la gran ventaja de este formato es la elevada compatibilidad que se obtiene. Dado que los equipos en ISDB-T soportan y son compatibles con HDMI 1.4a, esta serie de formatos pueden ser aplicados. De igual forma se definen los tipos de barrido de línea para cada estructura 3D, ya sea progresivo o entrelazado, manteniendo los principios de cada uno de estos métodos. Las estructuras 3D en HDMI 1.4a deberían procurar transmitir dos señales de video en completa resolución 1080p, una para cada ojo, es así que se establecen nuevas dimensiones del cuadro HD para 3D, todas ellas se basan en imágenes L+R, sin embargo también es válido para las estructuras actuales de 1920x1080p, claro que reduciendo su resolución. Estas nuevas dimensiones del cuadro HD establecen los siguientes conceptos:

- *Full HD 2D (FHD2D) / Half HD 3D (HHD3D)*: Es la de menor calidad para reproducción 3D, mantiene las dimensiones convencionales del FHD2D es decir 1920x1080p.
- *Full HD 3D (FHD3D)*: Es la que conserva la calidad de cada cuadro, por lo tanto las dimensiones del cuadro, aumentan de acuerdo a la configuración.

A continuación se presentan todas las configuraciones posibles y para cada dimensión del cuadro HD, cabe recalcar que la producción de estos contenidos dependerá de la acogida que tengan en los receptores. La información y datos detallados para las tramas y cuadros se exponen en el Anexo 1.

2.2.3.1 Formato de Empaquetamiento de Cuadro

2.2.3.1.1 Empaquetamiento de cuadro para formato progresivo

El formato de empaquetamiento progresivo consiste en que en lugar de que los cuadros L ó R ocupen el cuadro HD, se presenten íntegramente con todas sus líneas de manera continua, es decir que no exista división alguna entre líneas pares e impares, por ejemplo si se decide colocar el cuadro L a la derecha y el R a la izquierda del cuadro HD, estos cuadros tendrán todas sus líneas completas

en el lugar que ocupan. Para este formato se ubica el cuadro izquierdo en la parte superior, posteriormente un espacio en blanco y luego el cuadro derecho, se establece una nueva dimensión del cuadro HD (ver Figura 2.18), este espacio en blanco viene dado por el pulso de borrado que el cuadro tiene originalmente y que es de 45 líneas. Es muy similar, por la disposición de los cuadros, al formato Top and Bottom, pero se diferencia por la estructura detallada del cuadro.

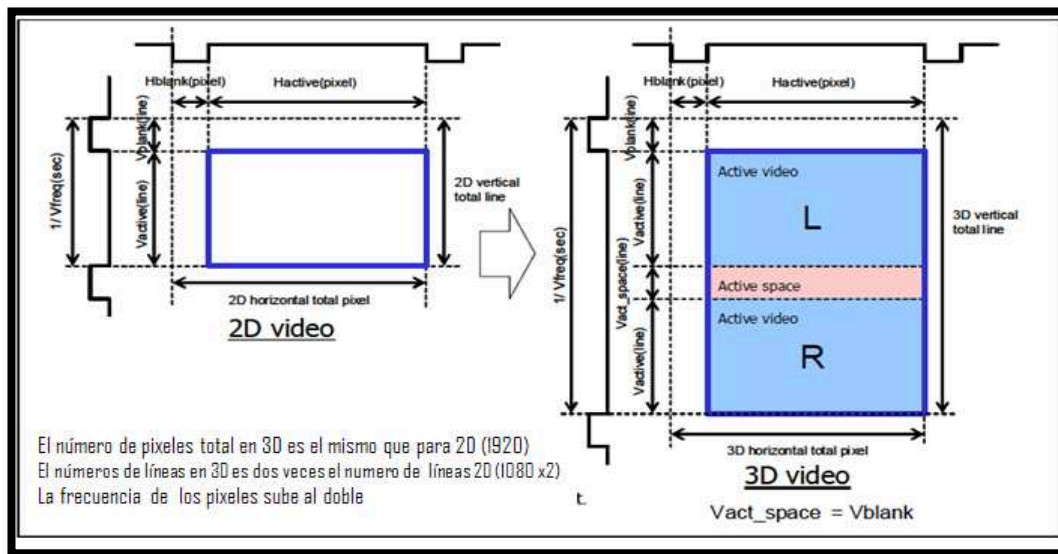


Figura 2. 18 Empaquetamiento de cuadro para formato progresivo.³⁵

El espacio “Active space” es característico de este formato, y con este el tamaño total del cuadro está dado por:

$$Total\ de\ lineas = V_{borrado} + Lineas\ Cuadro\ izq. + V_{active} + Lineas\ Cuadro\ der.$$

$$Total\ de\ lineas = 45 + 1080 + 45 + 1080 = 2250\ lineas \quad \mathbf{Ec\ 2.15}$$

2.2.3.1.2 Empaquetamiento de cuadro para formato entrelazado

Para este caso se debe extraer la información de líneas pares e impares de cada cuadro, es decir se tienen en total cuatro conjuntos de líneas, líneas pares del cuadro izquierdo, líneas impares del cuadro izquierdo, líneas pares del cuadro derecho y líneas impares del cuadro derecho, estos se ubican en el cuadro HD de

³⁵ Fuente: HDMI Specification V1.4a extracto “Señalización 3D”; Hitachi, Panasonic, Philips, Sony, Technicolor, Toshiba, Marzo 2010. (Las Figuras de la 2.18 a la 2.27 se refieren a esta misma fuente)

manera intercalada de la siguiente forma: líneas pares del cuadro derecho, líneas pares del cuadro izquierdo, líneas impares del cuadro izquierdo y líneas impares del cuadro derecho (ver Figura 2.19).

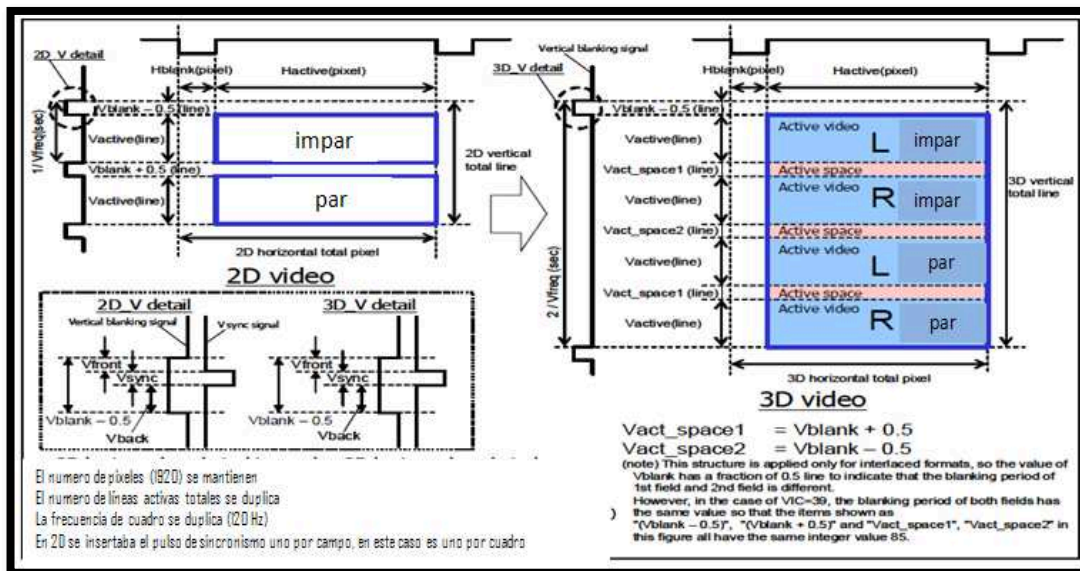


Figura 2. 19 Empaquetamiento de cuadro para formato entrelazado.

El número total de líneas viene dado por:

$$Tot.de\ líneas = 3 \times Active\ space + lin.\ der.\ impar + lin.\ izq.\ impar + lin.\ der.\ par + lin.\ izq.\ par + Vborrado$$

$$Total\ de\ líneas = 4 \times \left(\frac{45}{4}\right) + \left(\frac{1080}{2}\right) + \left(\frac{1080}{2}\right) + \left(\frac{1080}{2}\right) + \left(\frac{1080}{2}\right) = 2205 \quad \mathbf{Ec\ 2.16}$$

2.2.3.2 Configuración Side by Side Half (SbS)

En esta configuración se propone ubicar los cuadros izquierdo y derecho uno a continuación de otro horizontalmente (ver Figura 2.20). En este caso el SbS del tipo Half (HHD3D) los cuadros originales HD reducen su resolución a la mitad en el eje horizontal, para al final obtener un cuadro HD de dimensiones normales, en este caso el cuadro derecho se ubica a la derecha y el izquierdo en la izquierda. La señalización para detectar este tipo de cuadros se realiza mediante el estándar HDMI1.4a, el cual establece que dentro de la trama de control y señalización enviada existe un campo para especificar si es transmisión 3D, en este caso se usa el campo 3d_structure de la trama en el valor 1000 y el campo VIC en el valor 5 para frecuencia de 60 Hz, de igual forma existen bits señalizadores para

empaquetamiento progresivo o entrelazado, para establecer el tamaño de los pulsos de sincronismo, borrado, etc. (La descripción detallada de las tramas en HDMI 1.4a se encuentra en el Anexo 1).

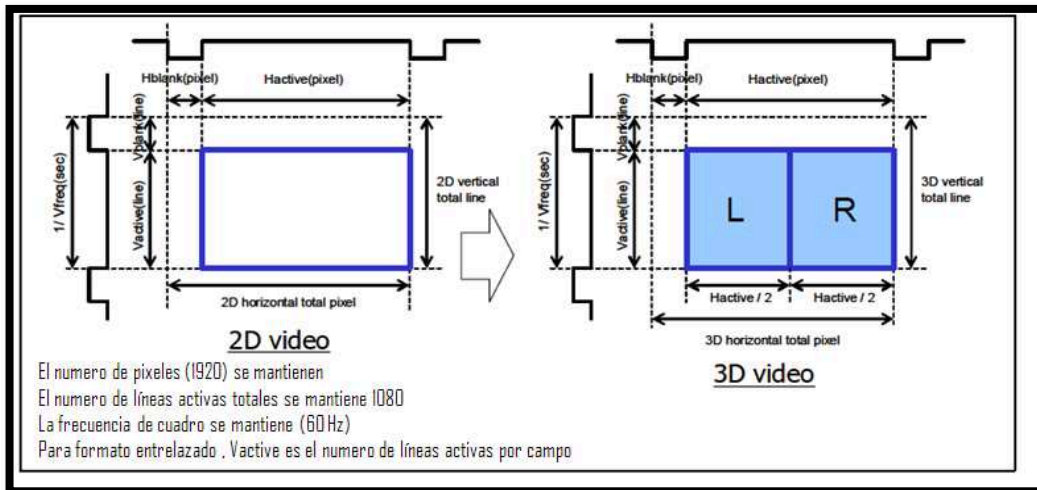


Figura 2. 20 Empaquetamiento de cuadro para formato entrelazado.

2.2.3.3 Configuración Top and Bottom (TaB)

En este caso las imágenes izquierda y derecha son submuestreadas a la mitad de su resolución en el eje vertical y posteriormente ubicadas como muestra la Figura 2.21, para este caso la señalización del 3D_Structure está en 0110 y de igual forma existe varias posibilidades, de acuerdo a la frecuencia de cuadro, en progresivo y en entrelazado.

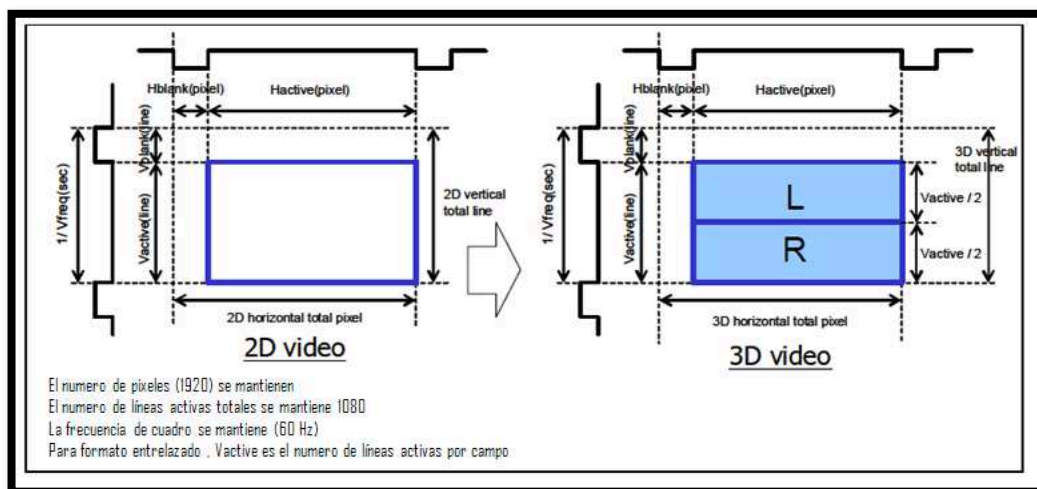


Figura 2. 21 Configuración Top and Bottom (TaB).

2.2.3.4 Configuración quincunx matrix (Tablero de ajedrez)

Este formato forma parte del grupo de formatos extendidos de HDMI 1.4a, es decir no son muy utilizados, pero también proponen una solución para la disposición de la información estereoscópica, en este caso se propone dividir a la imagen en todos sus pixeles e ir alternando en forma de tablero de ajedrez entre el pixel del cuadro derecho y el pixel del cuadro izquierdo (ver Figura 2.22), en este caso cada imagen también se submuestra a la mitad de su resolución para poder ocupar un cuadro HD de dimensiones convencionales.

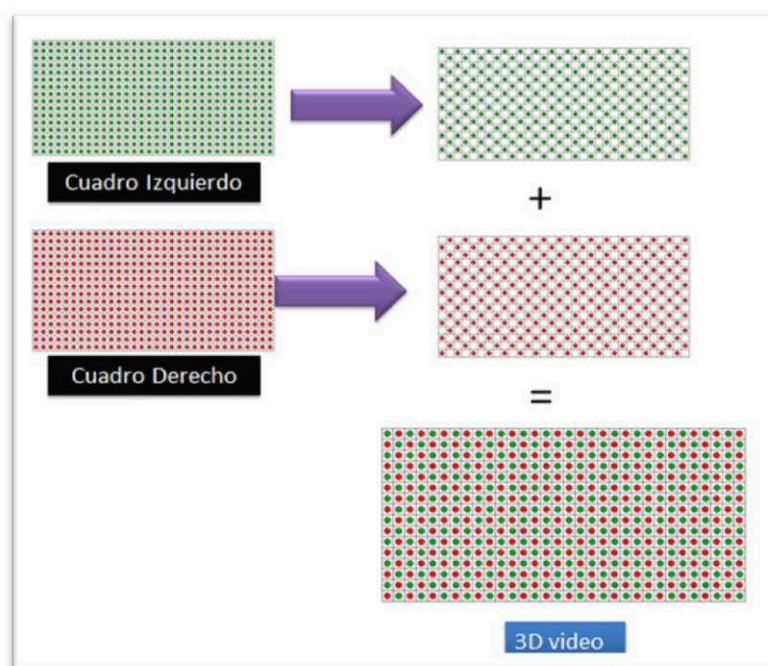


Figura 2. 22 Configuración Tablero de ajedrez.

En este formato se mantiene el número de líneas originales, es poco usado para la estructuración de las imágenes antes de la transmisión. Por ello no se requiere un estudio minucioso; existen dos posibilidades para estos cuadros, cualquiera de los dos izquierdo o derecho puede ocupar el primer pixel del cuadro HD.

2.2.3.5 Configuración campos alternados.

Este es uno más de los formatos extendidos de 3d_Structure para HDMI, es una configuración que se basa en el formato entrelazado, dado que intercala en la disposición del cuadro HD (ver Figura 2.23), los campos pares e impares para

información estereoscópica izquierda y derecha. Para este caso se mantiene el número de píxeles por línea, el número de líneas efectivas se duplica al igual que la frecuencia.

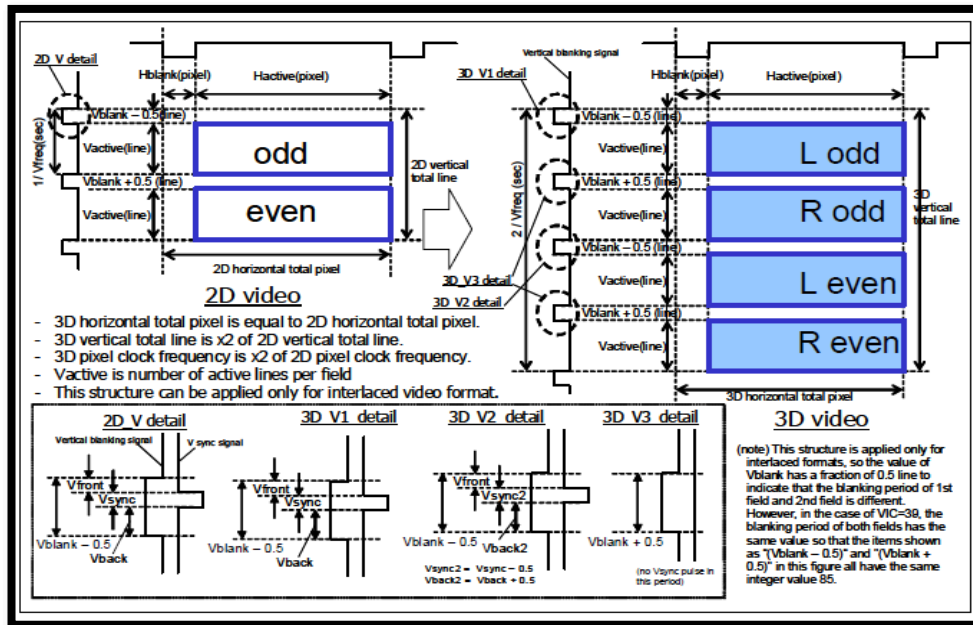


Figura 2. 23 Configuración de Campos alternados.

2.2.3.6 Configuración líneas alternadas.

Esta es otra propuesta que se admite solamente bajo el formato progresivo, se ubican las líneas de los cuadros izquierdo y derecho de manera alternada, una línea para el cuadro izquierdo la siguiente para el derecho y así sucesivamente hasta completar toda la información (ver Figura 2.24), aquí el número de píxeles por línea se mantiene en el nuevo cuadro, el número de líneas efectivas se duplica y el pulso de borrado tiene el doble del tamaño original.

2.2.3.7 Configuración Side by Side (Full).

Aquí se maneja la disposición de los cuadros igual al SbS half, pero se omite el proceso de submuestreo horizontal a la mitad de la resolución, por ello se duplica el número de píxeles por línea, y el número de líneas se mantiene, configurando un cuadro de 3840 x 1080 (ver Figura 2.25).

Para esa configuración se pueden ubicar las líneas de los cuadros de forma progresiva o de forma entrelazada.

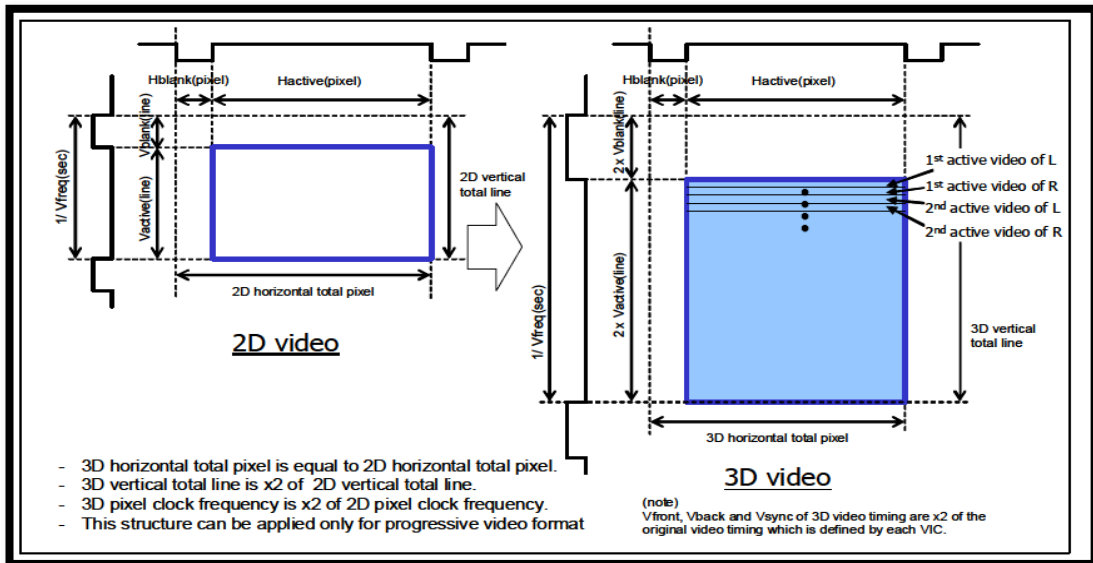


Figura 2. 24 Configuración Líneas alternadas.

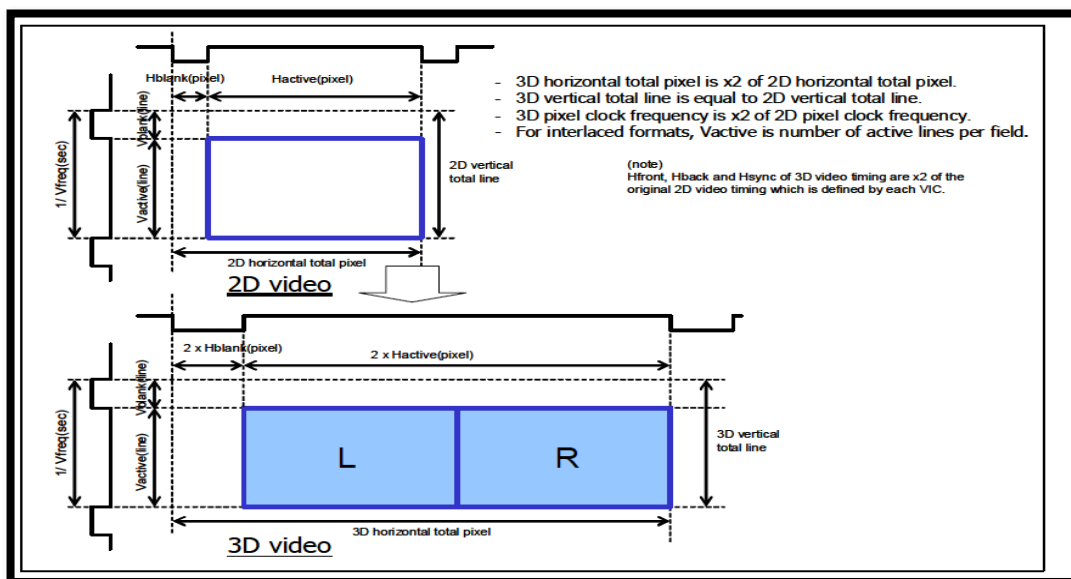


Figura 2. 25 Configuración SbS Full

Existen otros tipos de formatos que son compatibles con el cuadro HD y también entran en la categoría del formato V+D (Video + Depth), aún no son totalmente explotados como tal, pero HDMI 1.4a propone que en el cuadro HD se estructure la información del video y la de la profundidad.

2.2.3.8 Configuración “L + Profundidad” (Video L + Depth)

Esta es la versión de V+D que HDMI propone, para este caso se considera al igual que el audio, que el canal izquierdo lleva la información monoscópica del video y que basta para interpretar el plano. Además se ubica el plano o mapa de profundidad de manera sucesiva basado en el formato de empaquetamiento progresivo, en este caso se duplica el número de líneas efectivas y la información de profundidad tiene una resolución igual a la del video L, lo que no aprovecha al máximo los recursos que se tienen para HD (ver Figura 2.26), sin embargo resulta una solución cuando la información que se obtiene es de profundidad y no se puede reconstruir una segunda vista a partir de ello.

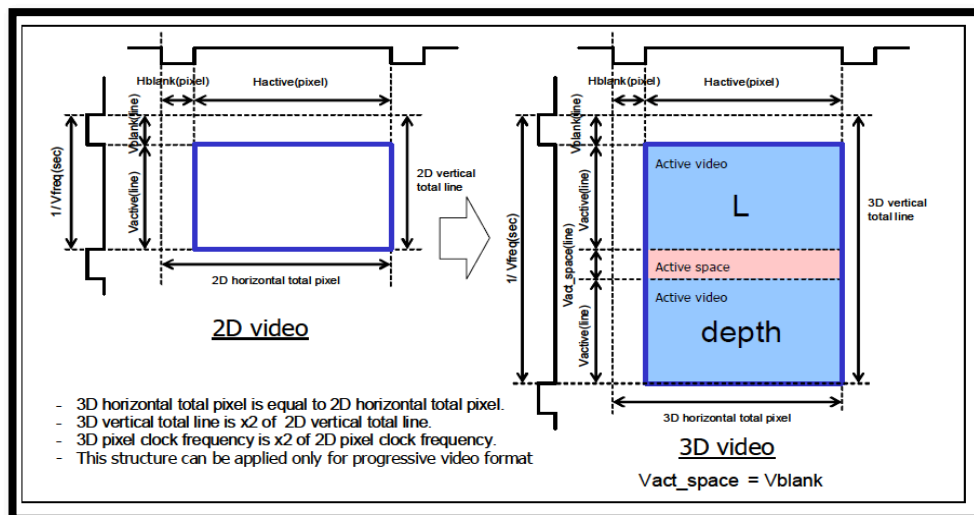


Figura 2. 26 Configuración L + Profundidad

2.2.3.9 Configuración L + D + Gfx + Gfx-L (Left + Depth + Graphics + Graphics – Left)

Este formato propone una solución de mayor calidad en la producción del video, ya que establece que además de enviar la información del video y de la profundidad, se necesita obtener la información gráfica de los componentes de la escena y adicionalmente un cuadro donde se plasme la diferencia entre los gráficos y la profundidad (ver Figura 2.27). Todo esto, para tener herramientas más precisas, y reproducir en recepción una mejor imagen y con sentido de profundidad más exacto; la desventaja de este método está en que el

procesamiento inicial del video debe ser más minucioso para poder lograr obtener toda esta información, además de una infraestructura de adquisición más compleja.

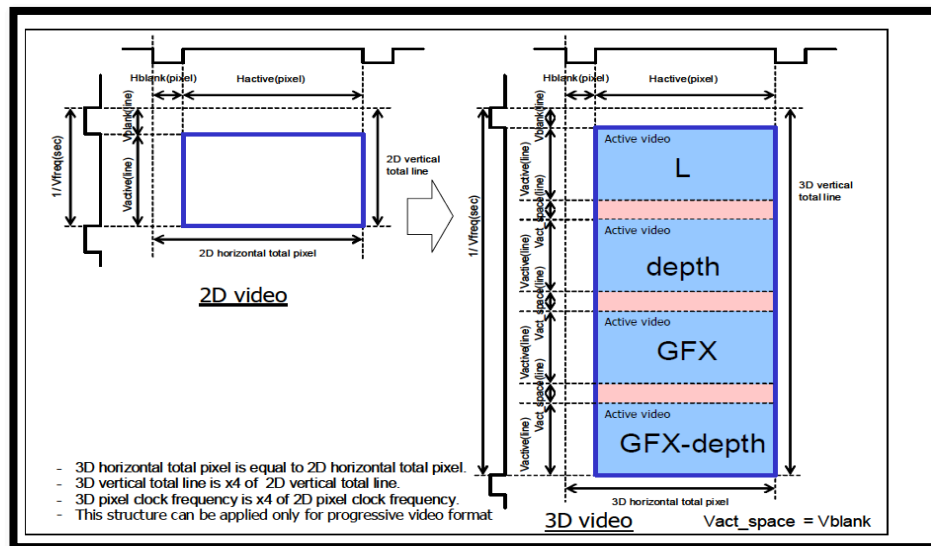


Figura 2. 27 Configuración L + D + Gfx + Gfx-L

2.2.4 FORMATO VIDEO + PROFUNDIDAD (V+D)

En las estructuras expuestas anteriormente, ya se aborda la posibilidad de que la profundidad sea parte de la información necesaria para una señal de TV estereoscópica. Es fundamental entender claramente que significa la señal de profundidad, qué contiene y cómo se la obtiene.

La herramienta que representa la información de la profundidad es el mapa de profundidad. La normativa ISO/IEC 23002-3 (MPEG-C Parte 3) define a este como un formato válido para los sistemas de TV3D y establece los modos de producción y compresión para estas imágenes. Para obtener un mapa de profundidad existen varias técnicas; a continuación se describen los procesos en el lado transmisor de este tipo de señales.

Existen 3 formas de obtener las señales necesarias para este formato, de todas estas se desprende el mapa de profundidad que luego será codificado independientemente y se estructura junto con el video en el conjunto de bits a ser transmitido.

2.2.4.1 El mapa de profundidad

El mapa de profundidad es una representación visual de las escenas de video, la cual contiene valores discretos que representan la profundidad de cada pixel del cuadro original. Las imágenes son representadas por cuadros en escalas de grises donde cada gris representa una distancia z , y se ha establecido que para los objetos cercanos se asigne el color blanco y para el más alejado el color negro (ver Figura 2.28). Para el video, la profundidad se representa por una señal de video monocromática de 8 bits, con estos 255 valores se obtienen buena calidad para la posterior interpretación de la profundidad de una escena. El mapa de profundidad consta simplemente de una matriz que contienen N valores asociados a los pixeles individuales del flujo de video regular. Esta señal puede ser procesada con técnicas específicas para señales de luminancia, o más modernamente con las técnicas de codificación que el video de la televisión digital utiliza. Esta al ser una señal auxiliar de video requiere ser señalizada, para lo cual se introduce un byte que determina que la señal es de este tipo. Esta señalización es compatible con varios sistemas, incluyendo ISDB-T. El estándar de codificación H.264 permite la codificación de una señal auxiliar de video que en este caso sería la de profundidad. La ISO/IEC 23002-3 establece dos tipos de señales auxiliares de video, la profundidad y el paralaje (el cual se puede ver como la representación hiperbólica de la profundidad). Estas dos señales ponen el byte del campo video auxiliar en los valores de 0x10 y 0x11 respectivamente. En el caso de que se pueda generar más información se pueden asignar más valores a estos bits. Los valores Z_p de cada pixel representados en este mapa vienen dados por un valor m de N -bits:

$$Z_p = \frac{m}{2^N} (k_{cercano} W + k_{lejano} W) - k_{lejano} W \quad \text{Ec 2.17}$$

W representa el ancho de la pantalla (en unidades de longitud) en recepción, esta información no es necesaria transmitirla ni se puede transmitir tampoco, se puede entender que W es un control del rango de profundidad. $K_{cercano}$ y K_{lejano} representan el rango relativo de profundidad respectivamente detrás y delante de la pantalla, los valores k se expresan en unidades inversas de longitud y

representan el valor más cercano y más lejano al que se encuentra un objeto, es decir el rango de operación del mapa de profundidad.

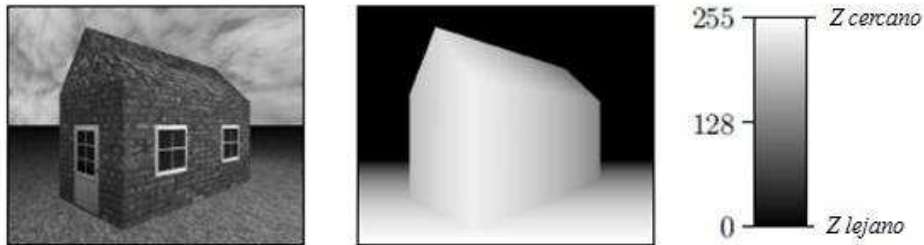


Figura 2. 28 Descripción del mapa de profundidad

2.2.4.2 Generación del mapa de profundidad a partir de 2D

La generación de este tipo de información, forma parte del formato de conversión 2D a 3D, es decir aquí se genera un mapa de profundidad utilizando algún algoritmo que estime el valor de la componente Z de cada pixel. Algunas técnicas crean un mapa de profundidad a través de la estimación de la segunda vista, y posteriormente emplean la técnica como si se hubiera obtenido la imagen desde dos cámaras. Existen una gran variedad de propuestas para obtener este mapa de profundidad, y todos estiman la componente Z de diferente manera, incluso existen técnicas para realizar transformación en tiempo real. Uno de ellos se expone en 2.1.5.

Para la generación de este mapa, algunas técnicas parten del análisis del vector del movimiento el cual se obtiene fácilmente a través del H.264 (MPEG-4), y con este vector construyen el mapa de profundidad. A continuación se expone otro sistema de creación del mapa de profundidad el cual plantea extraer un mapa a partir de analizar el movimiento de la escena.

La Figura 2.29 muestra que se parte de analizar la escena 2D, de la cual se desprenden dos aspectos, la profundidad de las señales y la estructura del movimiento, ambas relacionadas entre sí. Posteriormente se ejecutan dos cadenas de procesados para cada una de estas señales, en el caso de la profundidad, enfoque oclusión y textura son aspectos que determinan la exactitud del mapa; y por otro lado en la estructura del movimiento los algoritmos procuran

obtener características geométricas del video que permitan reconstruir la escena tridimensionalmente.

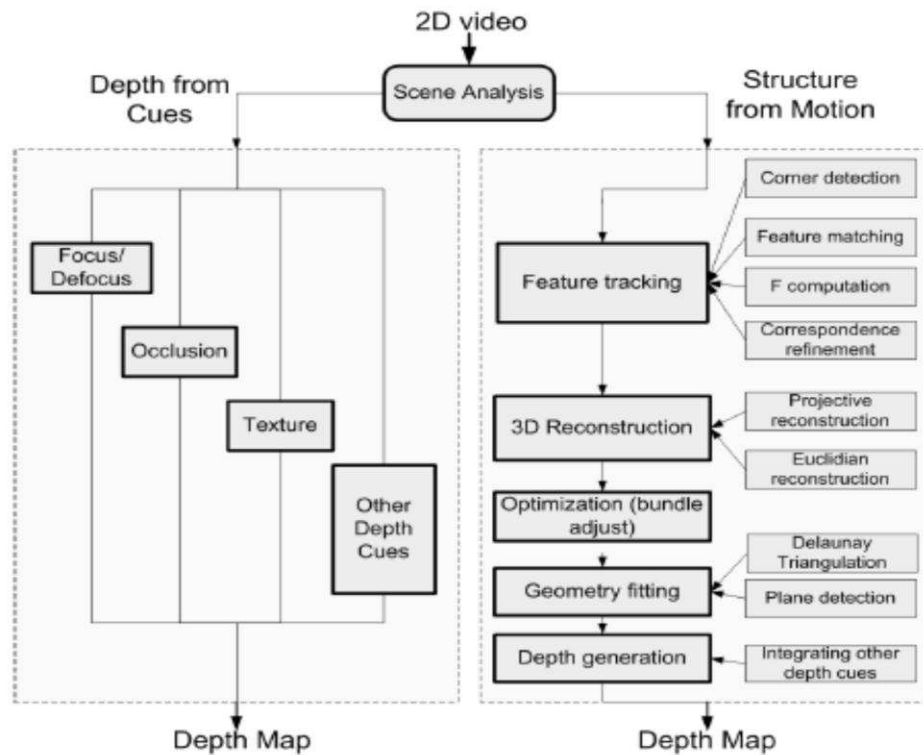


Figura 2. 29 Generación del mapa de profundidad a partir de 2D

2.2.4.3 Generación del mapa de profundidad a través de Cámara infrarroja

Las Zcam funcionan a partir de la luz infrarroja que detecta la posición de los objetos, por lo que también se denominan cámaras infrarrojas. Estas cámaras obtienen dos señales de video, el plano 2D y el mapa de profundidad, este tipo de cámaras generan pulsos infrarrojos que se reflejan por los objetos y dependiendo de dicha reflexión se obtiene el valor de profundidad de cada objeto de la imagen. Estas cámaras constan de un CCD ultra rápido y una fuente de luz infrarroja. Los sistemas de creación de mapa de profundidad por rayos infrarrojos procuran generar imágenes de alta calidad y no solo es la adquisición sino el proceso para generar el mapa en sí. Después de la adquisición, es clave en el proceso corregir y poner a punto factores como brillo, textura y segmentación de la imagen. En primera instancia la cámara emite y recibe pulsos de luz infrarroja, los cuales viajan junto con la luz que representa el plano, es por ello que cada

información debe ir a su correspondiente módulo de procesamiento. La Figura 2.30 presenta los componentes principales de estas cámaras, se observa que el sensor de imagen y el sensor de profundidad se encuentran sobre la misma estructura.

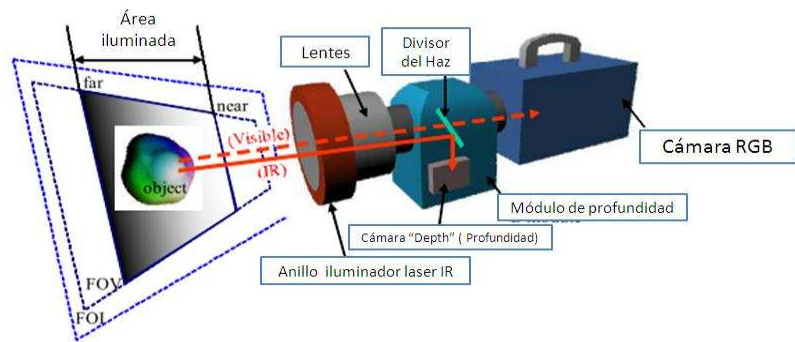


Figura 2. 30 Estructura de una Cámara de Profundidad (Z-Cam).

El principio de funcionamiento de estas cámaras, se basa en la detección de la profundidad, la cual se hace combinando la intensidad modulada de los iluminadores, con la cámara CCD equipada con un disparador ultrarrápido. El diagrama de la Figura 2.31 muestra la trayectoria que debe realizar la luz y los rayos IR de los objetos que la cámara debe captar:

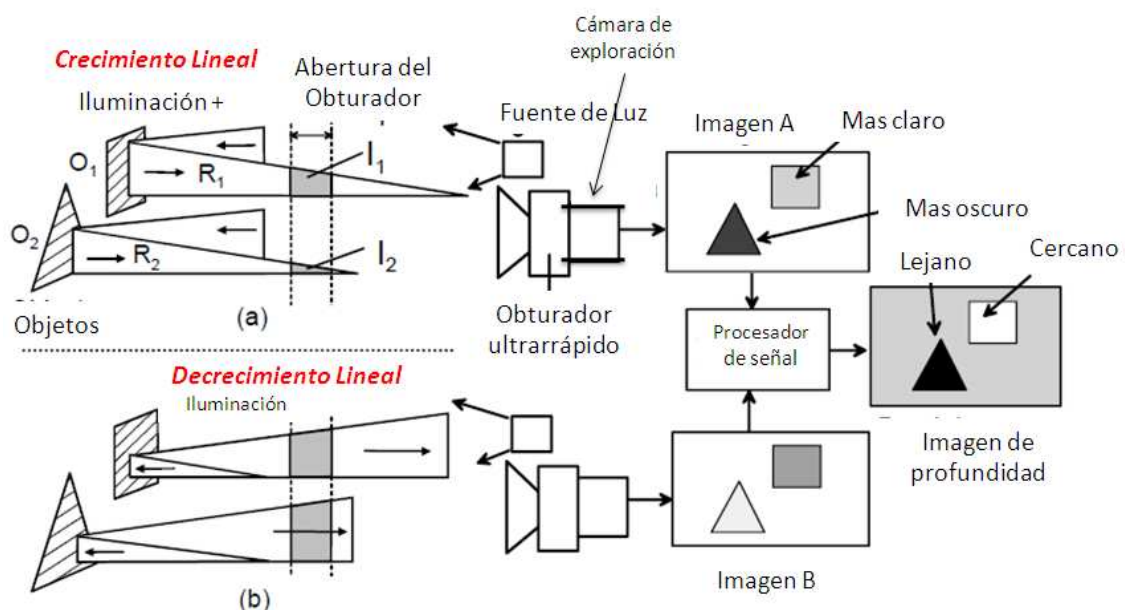


Figura 2. 31 Funcionamiento de las cámaras infrarrojas.

En la parte a) de la Figura 2.31 se observa el caso en el cual se da una captura de la distribución espacial de la intensidad de la iluminación que crece linealmente con el tiempo, la parte superior del diagrama muestra la captura del objeto más cercano O_1 (cuadrado); la intensidad instantánea I_1 de la luz reflejada regresa a la cámara y es detectada por el obturador ultrarrápido. El resto del trayecto de las imágenes es trasladado al procesador de señal el cual discrimina los elementos cercanos y lejanos relativamente. Varios modelos de cámaras basan su funcionamiento en bloques funcionales similares y utilizan fuentes infrarrojas similares de características parecidas (ver Tabla 2.3), entre los parámetros más comunes, para este tipo de medios de adquisición, se tienen:

Arreglo de los leds IR	Longitud de onda	850 nm
	Potencia luminosa promedio	1 W
	Frecuencia de modulación	10 a 50 MHz
Intensificador de Imagen	Tiempo de compuerta	1 a 20 nS
	Frecuencia de repetición	10 a 50 MHz
Cámara CCD de exploración	Numero efectivo de pixeles	1280x720 (29,91 Hz)
		853x480 (59,94 Hz)
Imagen de profundidad	Señal de salida	HD-SDI
	Resolución de profundidad	1,7 cm a una distancia de 2 m.

Tabla 2. 3 Características técnicas de las cámaras infrarrojas.

Con estas especificaciones trabajan muchos modelos de cámaras, en la Figura 2.32 se describe el proceso para que la cámara capte la señal V+D.

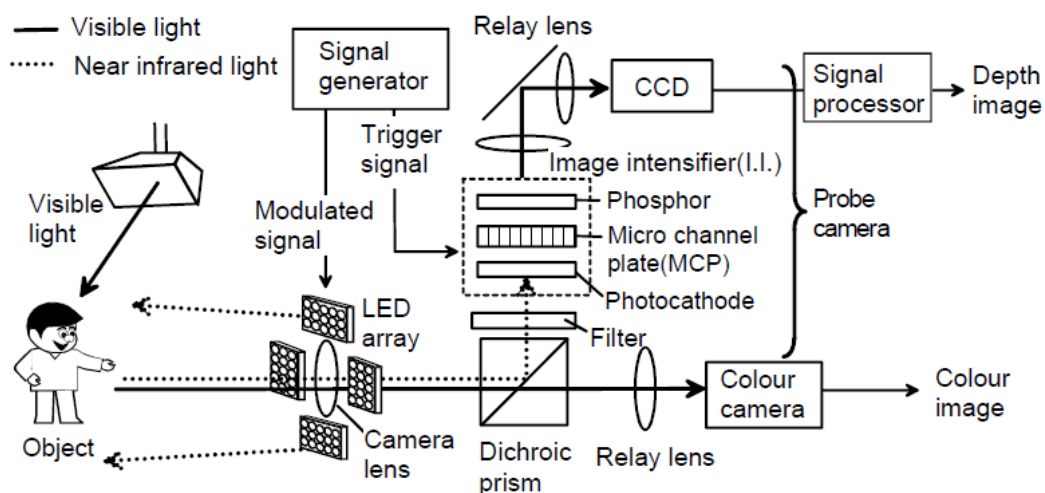


Figura 2. 32 Adquisición de una imagen con la Z-Cam.

Se observan claramente 2 señales ópticas la de la luz visible que provoca la imagen misma RGB del objeto, y por otro lado la información de profundidad. Los lentes de la cámara reciben las señales de color e infrarroja, estas son separadas por el prisma dicróico, el cual hace que todas las señales infrarrojas tengan un procesamiento singular. La parte de color llega a través de lentes de relevo y llegan a una cámara de tecnología convencional y se obtiene la imagen a color.

Por otro lado el tratamiento que debe tener la señal infrarroja es más complejo, ya que se somete, después de ser separada de la de color, a un filtro y posteriormente a un intensificador de imagen (I.I), el cual consta de un fotocátodo que se activa a través de un patrón que emite el generador de señal; el fósforo es un elemento capaz de crear una imagen más intensa de los puntos reflejados del objeto. Parte de la intensificación consiste en multiplicar a través del MCP (Placa micro-cana) la carga de electrones por 10^3 o 10^4 creando una imagen óptica en el fósforo. Esta señal intensificada pasa por lentes direccionadores que lo referencian a un CCD, el cual detecta esta luz la hace señal eléctrica que el procesador de señal se encarga de digitalizar y generar la información de profundidad. La acción del obturador ultrarrápido del I.I (Intensificador de Imagen), que actúa en el orden de los nanosegundos, es desempeñada aplicando pulsos cortos de un voltaje bias entre el fotocátodo y el MCP. La apertura del obturador se repite a la misma tasa que la frecuencia de la modulación de luz para conseguir una mejor relación señal a ruido SNR, posteriormente la imagen al final llegara a la cámara CCD, que suele ser de alta resolución HD.

La relación señal a ruido depende mucho del número de pixeles de la señal de profundidad, es así que, si incrementa el número de pixeles de la profundidad detectada por el CCD, la señal de corriente detectada por un pixel reduce. De ahí que la SNR tiene la siguiente relación con los parámetros de la cámara:

$$SNR \propto \frac{\sqrt{\eta I_0 T_0 A \tau}}{d} \quad \mathbf{Ec\ 2.18}$$

Donde η es la eficiencia cuántica del fotocátodo (GaAs) ubicado en el I.I, I_0 es la potencia lumínica, T_0 es la transmitancia óptica, A es el área de un pixel, d es la

distancia del objeto y τ es el tiempo de la imagen. Se conoce que el tamaño del pixel de una cámara HDTV es menos de la quinta parte que para una cámara SDTV y por lo tanto la SNR es proporcional a estas dimensiones. Con el objetivo de obtener una buena relación señal a ruido y considerando los factores que intervienen, se puede realizar lo siguiente: mejorar la eficiencia cuántica del I.I, optimizar el sistema óptico e incluso mejorar el arreglo de leds de alta potencia.

2.2.4.4 Generación del mapa de profundidad a través de Cámara Estéreo

Si bien es cierto la técnica más adecuada para generar un mapa de profundidad es a través de un transductor diseñado para ello; se han desarrollado diversos algoritmos que parten de que la mayoría de sistemas de 3DTV trabajan con cámaras estereo por la facilidad y porque las cámaras con dos lentes son más populares que las cámaras infrarrojas para profundidad. Por lo tanto la atención ya no está en los lentes ni en el proceso de adquisición, sino mas bien en la línea base del procesamiento de las señales izquierda y derecha para generar la información V+D. Con este desarrollo las cámaras estereoscópicas se vuelven más y más populares e indispensables si se desea realizar producción de TV.

Las condiciones geométricas para los lentes son iguales a las definidas en 2.1.2 y por lo tanto la parte minuciosa y detallada se encuentra en el tratamiento que obtengan estas señales. El avance de estos algoritmos ha permitido no solo generar el mapa de profundidad sino también reconstruir vistas adicionales a las dos ya existentes, con esto se entra en el campo de la multivista que como se mencionó anteriormente genera una apreciación de altísima calidad de una imagen con un realismo superior, esta generación de múltiples vistas toma como base el mapa de profundidad. En esta sección se analizará solo hasta la generación de dicho mapa, el tratamiento posterior será tomado en cuenta en próximas secciones. La gran ventaja de este tipo de procesamiento es que para la representación muestra gran versatilidad al ser compatible con varias técnicas.

El procesamiento que el par de imágenes tendrá se muestra en la Figura 2.33.

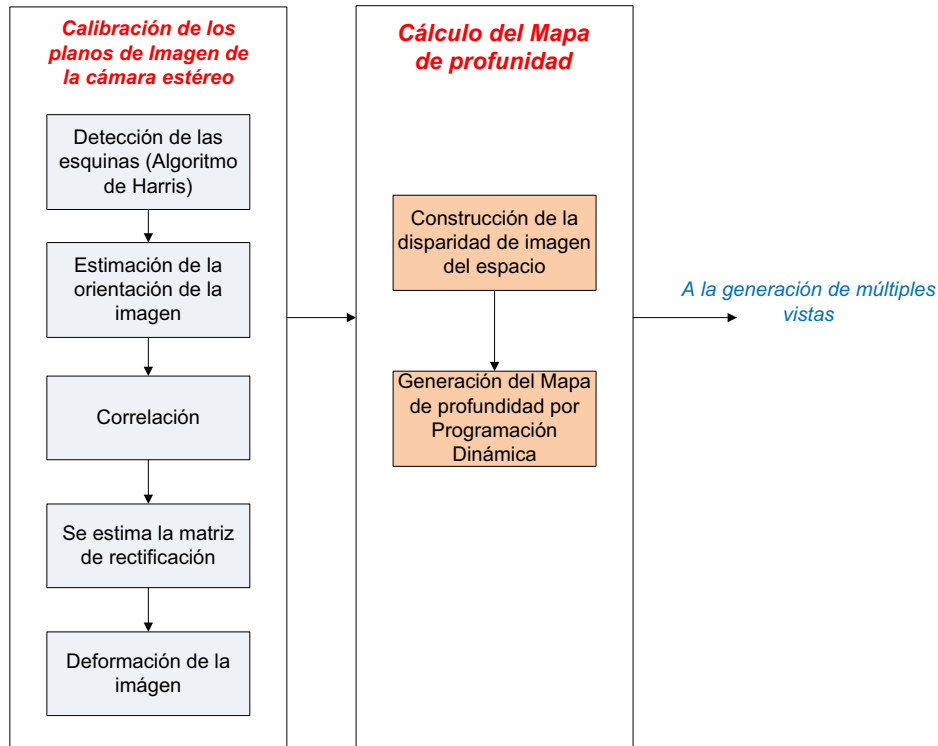


Figura 2. 33 Generación del mapa de profundidad con cámara estéreo.

Se identifican claramente los dos procesos para la generación del mapa, el primero consiste básicamente en preparar a la imagen para que se pueda extraer el mapa a partir de la disparidad. Los 5 procesos que la primera etapa comprende, procuran corregir los errores de los píxeles de la imagen para que posteriormente la detección de la disparidad sea correcta; a continuación se describe cada proceso y como la imagen varía en cada uno de ellos:

2.2.4.4.1 Calibración de los planos de la Imagen.

La calibración de estos planos consiste en encontrar la correspondencia de las imágenes y rectificar una de ellas para que encuentren en restricción epipolar³⁶. Para calibrar las dos imágenes primeramente se emplea el algoritmo de Harris para detectar los píxeles esquinas, debido a que estos píxeles tienen perfecta correlación entre las imágenes L y R, posteriormente una matriz de rectificación es estimada a través de la minimización de la localización de las distancias de los puntos correspondientes antes obtenidos; para que posteriormente una imagen

³⁶ Condición que establece que para todo punto de una imagen de un par estereoscópico, existe una línea epipolar en la otra imagen sobre la que está su homólogo.

sea rectificada, y las coordenadas verticales de los pixeles correspondientes sean idénticos o muy cercanos en ambas imágenes (ver Figura 2.34).

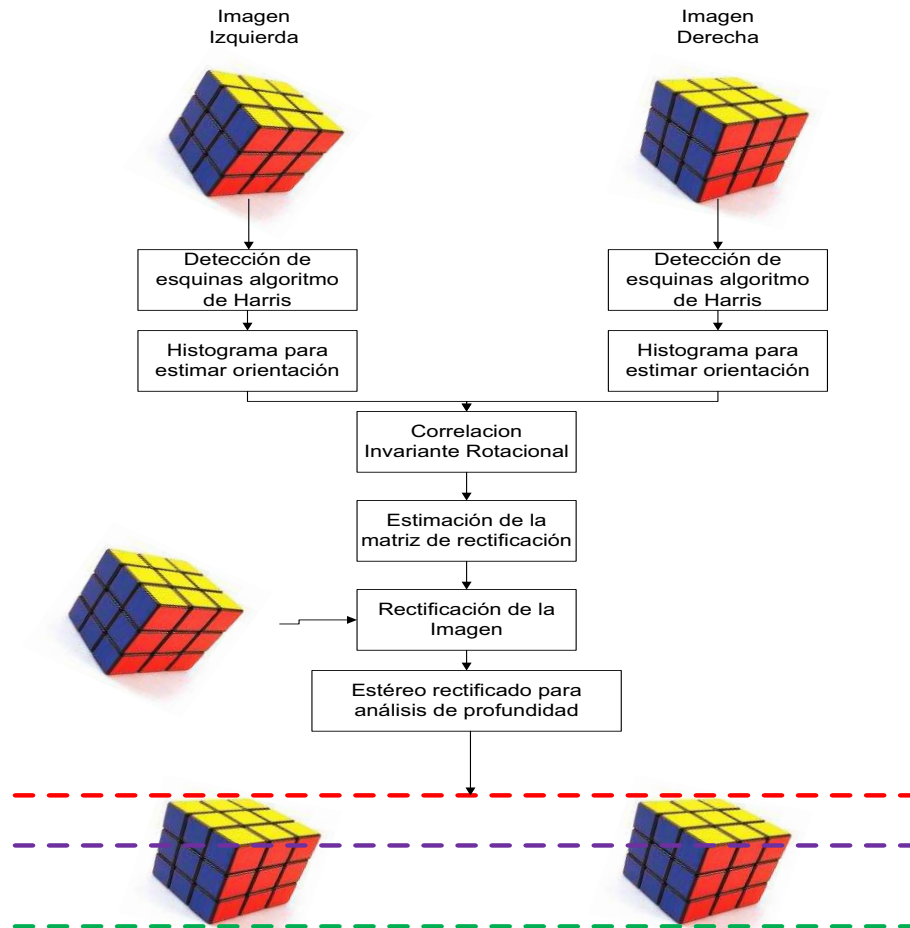


Figura 2. 34 Generación del mapa de profundidad con cámara estéreo.

En el diagrama se identifican claramente 5 etapas: la detección de esquinas permite, a través del algoritmo de Harris determinar los pixeles que representan las esquinas de las imágenes (ver Figura 2.35).

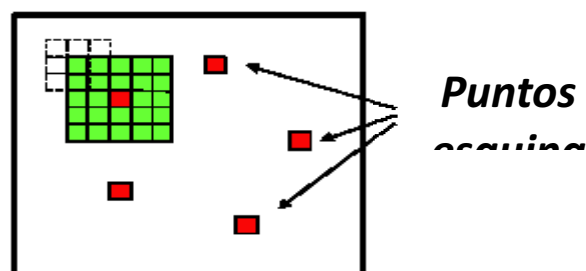


Figura 2. 35 Puntos de los pixeles esquina.

Posteriormente se estima la orientación de la imagen, en este proceso se evalúa el gradiente de cada pixel esquina determinado anteriormente, esto se hace con la derivada de los pixeles vecinos en una ventana de tamaño $n \times n$. En la Figura 2.35 se muestran 5 pixeles esquina y cada uno de ellos se evalúa individualmente, existiendo para cada uno de ellos derivadas horizontales G_x y verticales G_y de todos sus pixeles vecinos enmarcados en la ventana antes señalada, este gradiente se minimiza y permite obtener el histograma.

Como tercera etapa se descartan aquellos pixeles esquina que sean poco fiables, con ello se logra la correlación invariante rotacional, solamente los pocos pixeles esquina que poseen alta correlación entre las dos imágenes son tomados en cuenta para el siguiente proceso de establecer la función que permitirá rectificar estos pixeles (ver Figura 2.36), esta función se expresa como una matriz de rectificación y al final ya se puede establecer la nueva imagen y con ello el estéreo mejorado dispuesto al siguiente proceso.

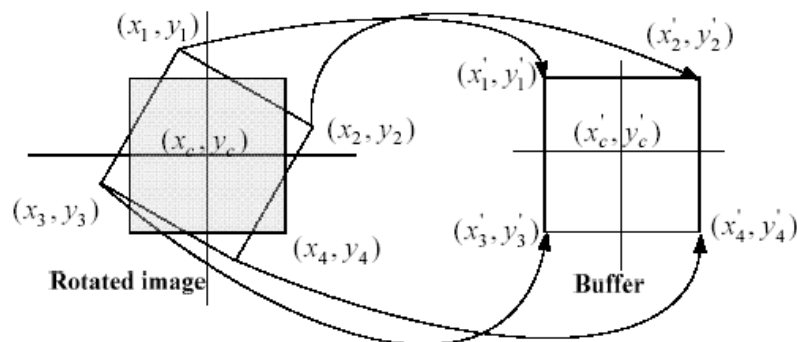


Figura 2. 36 Rectificación de rotación de un pixel.

2.2.4.4.2 Cálculo del mapa de profundidad.

Una vez realizada esta calibración a las imágenes, el procesamiento para obtener el mapa de profundidad ya es posible, en este caso el mapa de profundidad nace como el análisis de la disparidad de las dos imágenes; el proceso mediante el cual se calcula la disparidad de cada pixel de un par de imágenes se denomina problema de correspondencia estéreo. Se han propuesto algunas técnicas para resolver este problema, el más sencillo propone hacer una comparación pixel a pixel de las imágenes y, tomando una imagen como referencia, establecer la

diferencia de cada punto, y consecuentemente la disparidad; desafortunadamente esta sencilla solución no brinda los resultados esperados. El método que se propone como una solución más eficaz es el de obtener una disparidad de todos los píxeles en un esquema global de minimización de energía. Para estimar la disparidad de un píxel, el método usa completamente las disparidades útiles de los píxeles previamente estimados. En primera instancia se presenta el problema de la disparidad de los píxeles, es decir para un píxel de la imagen izquierda que se representa en la pantalla como (x,y) , correspondería uno en la derecha de la forma $(x-d,y)$; se observa disparidad vertical nula y un desplazamiento d en la coordenada horizontal. Si la imagen se define en las dimensiones $c \times h$, entonces existirían $cxhxd$ posibles estimaciones de disparidad, con ello se puede formar la matriz de disparidad que será usada posteriormente para el cálculo del mapa de disparidad.

El procesamiento, se torna menos complicado si se utilizan herramientas computacionales precisas, con ello se vuelve más fuerte el sistema de adquisición a través de dos cámaras, ya que no solo permite las vistas independientes sino también el mapa de profundidad, que es una alternativa válida para generar contenidos 3DTV.

2.2.5 FORMATO MULTI-VIEW VIDEO CODING (MVC)

Este formato es robusto debido a que posee su propio estándar de compresión y codificación, el Multi-view Video Coding es una extensión que se hizo para H.264 (MPEG-4), y en este ya se contempla la posibilidad de que desde la adquisición se obtengan no solo dos, sino múltiples vistas de una escena, y que estas sean comprimidas y transmitidas como un perfil multivista

Con la aparición de esta extensión a la norma, no solo se fortalece el concepto del 3D basado en la existencia de dos imágenes como en el sistema visual, sino también surge el concepto de FVV (Free View-point Video), en el cual se adquieren n vistas, y se presentan al espectador n vistas de las escenas (ver Figura 2.37). La explicación de esta extensión a la norma se explicará posteriormente en la sección de codificación y compresión.

MVC logra superar algunas dificultades técnicas para transmitir el video 3D, entre estas dificultades se presentan las siguientes; las representaciones FVV y las del video 3D convencional requieren que se transmitan señales múltiples de video y además, todas ellas sincronizadas, por eso grandes cantidades de datos deberán ser comprimidos con eficiencia. Además existe una gran dependencia estadística entre las vistas existentes que debe ser interpretada y utilizada por el estándar.

El hecho de que este estándar sea una nueva extensión al ya existente H.264, no impide que exista la compatibilidad para transmitir el flujo de bits con el AVC (Advanced Video Coding).

Está claro que este formato representa una generalización para los sistemas 3DTV debido a que requiere el tipo de adquisición presentado en 2.1.3, sin embargo las ventajas que presenta este sistema lo ha vuelto muy atractivo y en particular el caso cuando $n=2$ vistas es hoy en día muy aceptado, y se conoce como estéreo de alto desempeño, utilizado en los sistemas Blu-Ray 3D.

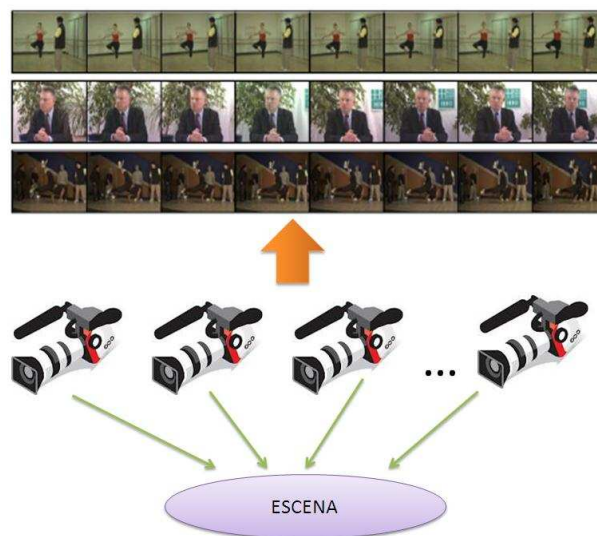


Figura 2. 37 Formato de 3D MVC (Perfil Multi-vista)

La figura 2.37 muestra sistemas multivista en los cuales se generan n imágenes, aquí es donde actúa MVC, ya que cuenta con estándar propio para predecir corregir y comprimir las n imágenes tomando en cuenta la dependencia entre todas las vistas.

Es importante destacar también que MVC ha sido usado en otros formatos, así como se aprovechan algunas funcionalidades de este estándar, es decir se pueden generar otro tipo de señales de video o incluso señales adicionales que son soportadas también por MVC.

2.2.6 FORMATO MVD (MVC + D)

La búsqueda de cada vez más calidad ha llevado a extender formatos ya existentes, es así que al MVC, a pesar de que en su concepción requiere n vistas lo que lo haría de gran calidad, se le añade una información importante que ya ha sido estudiada, la profundidad. Este formato es considerado dentro de la categoría de los formatos avanzados para producción de video 3D y se establece la existencia de n vistas, cada una de ellas con su respectivo mapa de profundidad (ver Figura 2.38). Al igual que para MVC, la particularización de este formato en $n=2$ significa un estéreo de alto desempeño.

La gran ventaja de este formato es que al existir múltiples vistas y reforzadas con la información de profundidad, la presentación del video al espectador tiene más calidad y mejor desempeño; incluso el espectador podrá observar cada una de las n vistas de manera independiente en 3D. Debido a que cada una tiene su mapa de profundidad. Además el procesamiento digital de la señal se simplifica al evitar el cálculo de algunos parámetros de profundidad.

Dada la existencia de múltiples vistas, se da una interpolación de vistas intermedias virtuales a lo largo de todas las cámaras. Para el procesamiento de estas señales, el formato se basa en la plataforma MVC para el procesamiento de las n vistas.

Como consecuencia de adquirir tantas imágenes, puede decirse que en recepción la tarea para representar este video al usuario con técnicas avanzadas, es más simple y puede llegar a resultados excelentes, se ha demostrado que este formato permite el ajuste del paralaje cero, lo cual resulta muy útil para el empleo de pantallas autoestereoscópicas. Aún con la complejidad en la adquisición el formato es conveniente para los usuarios que mantengan el uso de gafas y para

los sistemas anteriores a este, es decir se guarda compatibilidad con el estéreo tradicional.

La información de profundidad de la escena puede ser obtenida a través de una Z-CAM o por procesamiento de un par de imágenes, incluso por conversión 2D a mapa de profundidad.

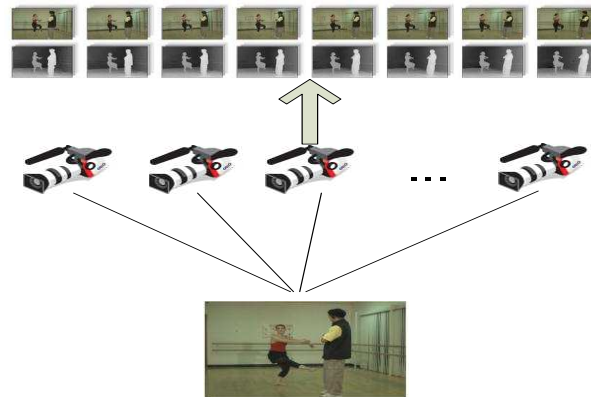


Figura 2. 38 Formato de 3D MVD (Perfil Multi-vista + Profundidad)

Se observa que se requiere de un número n de vistas elevado, sin embargo el número de vistas para la transmisión puede ser reducida por la interpolación de vistas (ver Figura 2.39); dado que las vistas son muy similares. En la realidad MVD transmite sobre el canal solamente una vista con la información completa (con sus respectivos datos residuales) y posteriormente las vistas laterales no transmitidas son generadas por la síntesis de las vistas así como la transferencia desde la vista central a las vistas laterales. El problema que surge con esto, es que no necesariamente existe cada pixel para cada vista, lo cual desemboca en espacios vacíos en el momento de proyectar la vista central, este proceso se conoce como disoclusión. El método para contrarrestar esta disoclusión, es pre-procesar el video de profundidad para permitir la reducción de las discontinuidades en los datos de profundidad en una forma en la cual se procure disminuir la disoclusión.

Dentro de las ventajas que este formato presenta, se encuentran: la presentación favorece a los sistemas que no utilizan gafas 3D, se permite el escalamiento de la profundidad, algunas vistas son interpoladas y se permite el ajuste del punto de

cero paralaje. Las desventajas que se pueden notar es que no surge como estándar para los formatos 3D en el cine y que para n muy elevado el sistema se vuelve ineficiente.



Figura 2. 39 Interpolación de vistas virtuales en MVD.

2.2.7 FORMATO LAYERED DEPTH VIDEO (LDV).

Este al igual que MVD, está dentro de los llamados formatos avanzados para video 3D. Para este formato se generan más informaciones de profundidad para la imagen que se adquiere; al igual que el anterior es un formato multivista por lo que pueden existir n informaciones visuales por señal de video, en este caso se aprovecha mucho la interpolación dado que a partir de m vistas se interpolan el resto de las n vistas ($m < n$), por lo general $2 \leq n \leq 8$. Como es de imaginarse la cantidad de datos que se obtienen es bastante grande por lo cual la compresión que se realice es fundamental y determinante. La producción extra de información hace que se explote de mejor forma la correlación intervista además de permitir una eficiente reconstrucción de la escena. El método nace con LDI (Layered Depth Image) y se expande hasta aplicarlo no solo a imágenes sino en video también. La información que se asocia por cada vista de la imagen es: la imagen de video monoscópica, el mapa de profundidad asociado, una capa para representar el fondo del video (se obtiene de la oclusión del video original), y una capa de fondo para la profundidad, tal como lo muestra la Figura 2.40.

A lo largo de todas las técnicas de reconstrucción de imágenes 3D, LDV es una de las más eficientes en el momento de sintetizar las vistas en escenas 3D complejas.



Figura 2. 40 Información obtenida de LDV.

LDV, que surge como la extensión de LDI, se considera conveniente cuando se provee video 3D con una vista completa o una vista central como la capa principal y como extensiones adicionales a estas, se incluye la información extra antes descrita. Por todo esto se considera que LDV se deriva y que al mismo tiempo representa una alternativa para MVD. En LDV se elimina el problema de la disoclusión, este formato considera representaciones de datos complejas y multidimensionales, lo cual permite el almacenamiento de la profundidad y la información de color para los píxeles que se ocuyen en la vista central.

En este formato puede surgir un inconveniente para generar la capa de oclusión, por lo que el método para crearla es un tanto complejo, para generar esta capa ya se tendrán calibradas las vistas completas y la capa de profundidad que se puede obtener por varios métodos, después de la alineación del color, todos los píxeles que han perdido el valor de profundidad son asignados al fondo de la información de profundidad de los alrededores. Para dar volumen a los píxeles de la imagen de color de acuerdo a los valores de profundidad, pueden ser generadas vistas adicionales a las convencionales izquierda y derecha.

Al moverse a lo largo de la vista analizada nuevas aéreas se tornan visibles donde la información de la textura de la imagen no está disponible para la vista original, esta área llamada oclusión se puede crear por interpolación de la textura

original creando así, un efecto de plano texturizado para los bordes de los objetos con lo cual se puede ya extraer las capas de oclusión necesarias.

Originalmente para el formato LDV existen 3 vistas, ilustradas en la Figura 2.41.



Figura 2. 41 Vistas originales adquiridas en LDV

Deformando la vista central hacia las vistas izquierda y derecha (ver Figura 2.42).

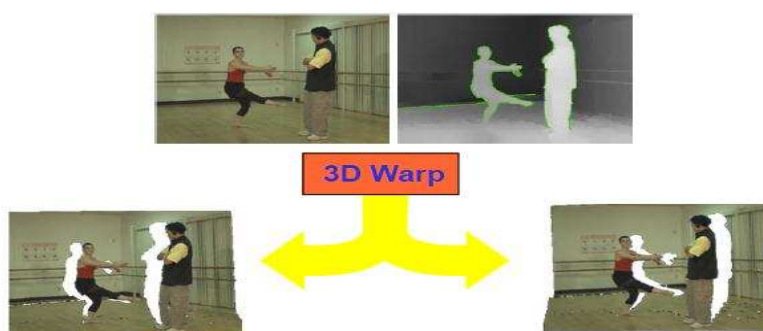


Figura 2. 42 Deformación de las vistas originales adquiridas en LDV.

En la Figura 2.43 se observa el proceso que se dio en las vistas para este formato.

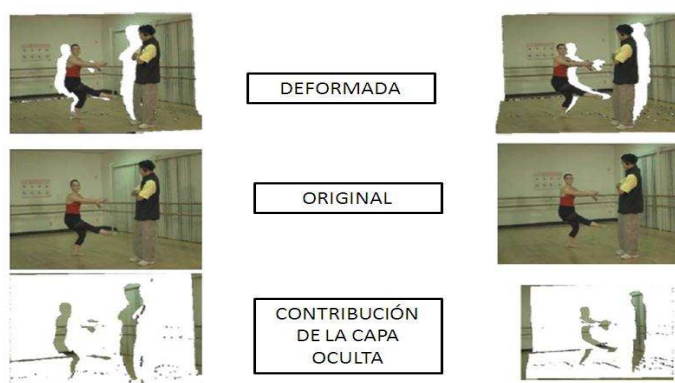


Figura 2. 43 Procesamiento de las vistas originales.

Ya con la contribución de esta capa oculta se puede generar la información de oclusión, llegando al final a obtener los 4 cuadros necesarios tal como la Figura 2.40 muestra.

La aplicación de este formato parte de variadas técnicas de adquisición de los cuadros, se pueden plantear sistemas desde $n=2$ lentes estereoscópicas hasta usar el sistema genérico de adquisición descrito en 2.1.6, en general se debe procurar obtener la información de profundidad lo más exactamente posible, por lo que siempre será recomendable contar con un sistema de adquisición que brinde la más completa información necesaria para una óptima recepción.

2.2.8 FORMATO DEPTH ENHANCED VIDEO (DES)

Este es el último de los formatos avanzados y desarrollado a la actualidad para el video 3D; en este se establece que se pueden generar 2 o más vistas ($n \geq 2$) acompañadas de información de profundidad y capas de oclusión (ver Figura 2.44); la información es similar a LDV sin embargo la diferencia es que LDV transmite una vista completa y crea el resto, en este caso cada una de las cámaras capta el video de color con su respectiva extensión de la información.

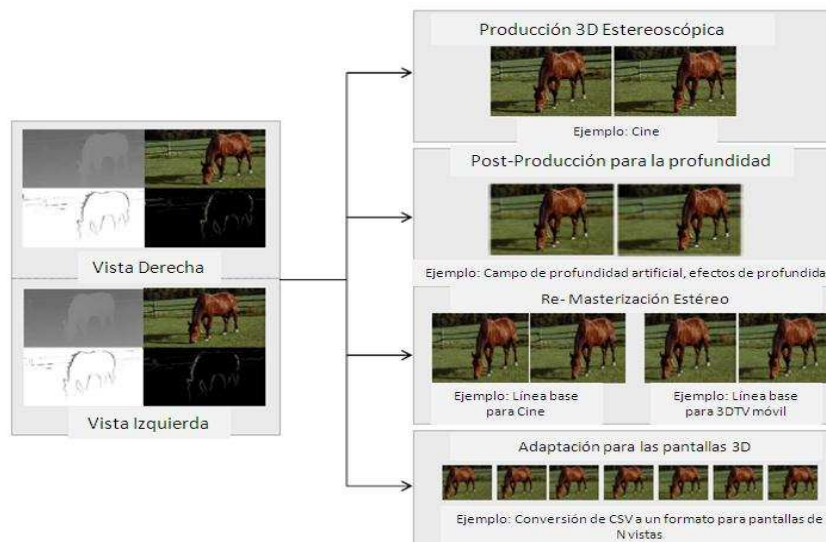


Figura 2. 44 Información en DES para $n=2$ vistas.

En general todos los formatos, en los cuales se utilice a la profundidad como información para crear el 3D dan flexibilidad para que se pueda manejar de mejor

forma toda la información, por lo tanto se vuelve un poco más sencillo el manejar conceptos como movilidad, efectos de profundidad, etc.

Está claro que el CSV, los formatos compatibles con HD y el MVC son formatos más fáciles de producir en la actualidad, sin embargo el Depth Enhancer Video se presenta como un formato genérico para video 3D³⁷. Este formato extiende al estéreo convencional ya que genera una serie de datos extras a las originales izquierda y derecha de un sistema estereoscópico tradicional, el extender el estéreo tiene como ventaja, dejar la posibilidad de que los formatos y esquemas anteriores puedan adaptarse. Dentro de las informaciones adicionales se manejan señales de profundidad adicionales y capas de oclusión, lo cual brinda una funcionalidad mayor por ejemplo se crean facilidades para la post producción, además se puede sintetizar varias vistas así como adaptar la línea base con el fin de presentar en una pantalla de TV o ser proyectado en el cine. DES establece que existan 2 flujos de video, con un mapa de profundidad y dos capas de oclusión cada uno. La figura 2.44 muestra el contenido de la información de DES, así como las posibilidades que el formato brinda al tener toda esta información. Se dice que este es un formato genérico porque el sistema genérico de adquisición brinda todas las señales necesarias de manera precisa y sencilla.

MVD, LDV y DES fueron presentados como formatos avanzados para esquemas multivista y se mencionó como ventaja sobresaliente la viabilidad para que se usen en recepción pantallas autoestereoscópicas al usuario.

2.3 PROCESAMIENTO DE LAS SEÑALES DE VIDEO ADQUIRIDAS

De acuerdo a la cadena de transmisión del video 3D, expuesto al inicio de este capítulo en la Figura 2.1, la adquisición, la forma (formato) en que se lo hace, y la información que se genera constituyen el primer bloque en aquella secuencia, el siguiente paso a este tiene que ver con el pre procesamiento, el cual depende de cada formato, por ejemplo si se maneja profundidad, se referirá a la estimación

³⁷ A. Smolic, K. Müller, P. Merkle, P. Kauff, and T. Wiegand, "An Overview of Available and Emerging 3D Video Formats and Depth Enhanced Stereo as Efficient Generic Solution", Proc. Picture Coding Symposium (PCS) 2009, Chicago, IL, USA, May 2009.)

del mapa de profundidad, dado que ya se han desarrollado los formatos anteriormente. A continuación se describe qué tipo de pre procesamiento requiere cada formato, posterior a ello se realizará una breve descripción de dichos procedimientos.

El procesamiento no suele ser una etapa complicada y en la mayoría de los casos, solo requiere un desarrollo de software que realice todas estas operaciones, la pagina web www.youtube.com, ha implementado en su portal, el servicio para videos 3D, en este se puede hacer una manipulación completa de los videos, cambiar los filtros coloreados, cambiar el esquema para un CSV, establecer SbS o con líneas alternadas, etc.

A continuación la Tabla 2.4 muestra un resumen de cómo las señales se procesan antes de la etapa de codificación y transmisión.

<u>Formato</u>	<u>Procesamiento</u>	<u>Descripción</u>
OCA: Anáglifo cromático optimizado.	Conversión L+R a anáglifo cromático.	Se toman dos señales visuales de video (izquierdo + derecho) y se procesa para adquirir un cuadro en anáglifo cromático
CSV: Video estéreo convencional.	Procesamiento convencional para video monoscópico	Dado que las señales se tratan al igual que el video 2D, aquí se incluyen técnicas ya conocidas como mejoramiento de la imagen, corrección del color brillo, efectos visuales, etc.
CFC y CFC+ : Compatible con el cuadro HD.	Procesamiento de mejoramiento de imágenes HD	Al igual que CSV, se mejora pero en este caso las técnicas se aplican a los cuadros HD.
V+D: Video + Profundidad.	Estimación del mapa de profundidad	Para cualquier sistema de adquisición de video + depth el proceso de estimación varía e incluye etapas de correlación de imágenes, detección de esquinas, etc.
MVC, MVD, LDV, DES.	Estimación del mapa de profundidad, mejoramiento de imágenes creación / interpolación de vistas.	Dependiendo de lo que contenga cada uno de estos formatos avanzados, se utiliza las herramientas que se tienen para los anteriores, es exclusivo de estos formatos el procesamiento para la señal multivista

Tabla 2. 4 Pre procesamiento requerido por cada formato.

2.3.1 PROCESAMIENTO DE L+R HACIA ANÁGLIFO CROMÁTICO OPTIMIZADO

Este proceso es vital para sistemas OCA, aquí se genera una única señal que contiene, mediante la técnica cromática, la información correspondiente a las imágenes izquierda y derecha (ver Figura 2.45); como se menciona en la descripción del formato este procesamiento es sencillo, y hoy en día las técnicas permiten tener alta calidad y una distorsión cromática de la imagen que prácticamente tiende a cero. Los equipos y software que logran este efecto utilizan algoritmos conocidos y desarrollados desde hace muchos años, por lo cual es una técnica muy simple a tal punto que las aplicaciones desarrolladas para crear este tipo de imágenes y videos son de libre acceso; sin embargo el software y el equipo necesario para generar anáglifo optimizado (de alta calidad), aún es desarrollado y se ofrece solo en compañías especializadas en este tipo de procesamiento de video. Se han desarrollado muchas herramientas para el procesamiento del video L+R a anáglifo cromático, entre ellas Matlab, Stereo Movie Maker, y Youtube que son las más conocidas y de fácil acceso, con todas estas herramientas el procesamiento no resulta complejo y hace a este sistema muy conveniente aún sin tener una aceptabilidad completa.



Figura 2. 45 Conversión L+R hacia anáglifo cromático.

2.3.2 PROCESAMIENTO DE MEJORAMIENTO DE IMAGEN, TÉCNICAS CONVENCIONALES EN VIDEO MONOSCÓPICO

Tanto para CSV, como para CFC y CFC+, lo importante es mejorar la apariencia y características de las vistas que involucran la información tridimensional. Para la mayoría de los procedimientos que mejoran la imagen, se utilizan algoritmos comunes como son, detección de bordes, cálculo de distancias, relación entre

pixeles, etc. Además de corregir posibles errores en la adquisición eventualmente se puede añadir en esta etapa, efectos de color o textura para las imágenes.

Entre ellos se encuentran:

- Desenfoque
- Desenfoque gaussiano
- Brillo
- Tonalidad
- Filtros de Imagen
- Corrección del color

2.3.3 PROCESAMIENTO DE ESTIMACIÓN DE PROFUNDIDAD, INTERPOLACIÓN DE VISTAS, CREACIÓN DE MAPAS.

Desde el formato V+D hasta los formatos avanzados, es importante procesar las imágenes de manera que generen mapas de profundidad, interpolen nuevas vistas y creen una correspondencia entre información de profundidad y video 2D. El procesamiento adicional para estos formatos es el señalado en los puntos en los que fueron desarrollados dichos formatos. Dado que el procedimiento depende del tipo de adquisición y de la información que se quiera generar, los algoritmos y las técnicas son variables; en algunos casos este procesamiento se lo hace en la adquisición y en otros casos se lo hace en una etapa posterior.

En todo este procesamiento es muy normal también, corregir algún defecto del video en adquisición, de tal manera que la señal a codificarse sea la deseada, y sea la que se procure recuperar posteriormente en recepción.

2.4 CODIFICACIÓN Y COMPRESIÓN DE LAS SEÑALES DE VIDEO ESTEREOSCÓPICO

En las secciones 2.1 y 2.2 se pudo notar la versatilidad y las posibles formas de crear información para el video estereoscópico, con todas ellas se establecieron los formatos, que en definitiva establecen qué información verdaderamente se transmitirá y se difundirá a los usuarios, es por ello que en esta sección se realiza un análisis de los algoritmos (estándares) de codificación compresión que se utilizarán dependiendo del formato que se utilice y teniendo en cuenta si dicha compresión está dentro de los esquemas que ISDB-T contempla. A continuación se describen el estándar de compresión y codificación a utilizarse. Es importante

considerar dos aspectos importantes: primero, que ISDB-T trabaja con codificación MPEG-4 y segundo, soporta HDTV.

Cabe señalar, que parte importante de MPEG-4 es el H.264 o parte 10 que rige la codificación del video, es por ello que es posible referirse al estándar como H.264, MPEG-4 (parte 10) ó ISO-IEC 14496-2. Para cada formato la Tabla 2.5 muestra la codificación adecuada.

FORMATO	ESQUEMA DE CODIFICACIÓN - COMPRESIÓN	APLICABLE	OBSERVACIÓN
OCA	Cualquier MPEG	SI	Resulta aplicable si se trata al video como 2D, sería de mayor calidad en HDTV
CSV	Cualquier MPEG en Simulcast	SI	Aplicable solo si a los dos canales se les aplica MPEG-4
CFC	MPEG-2	NO	Esquema de codificación anterior.
CFC+	H.264 (AVC, SVC)	SI	Aplicable con las actualizaciones necesarias al MPEG-4
V+D.	MPEG-Parte 3 , MPEG-4, AVC, SVC , Platelet	SI	En todos los casos se debe actualizar a la extensión reciente de H.264, es decir, los otros esquemas no se encuentran estandarizados
MVC	MPEG-4 , MVC	SI	
MVD	MVC adaptado, DIBR	SI	
LDV	LDV (3DTV Solutions), H.264 (MVC)	SI	
DES	MVC adaptado, Platelet	SI	

Tabla 2. 5 Esquema de codificación según el formato.

La razón para que prácticamente todos los formatos sean aplicables, radica en la versatilidad y gran desempeño del estándar MPEG-4, si bien es cierto MPEG-2 llegó a desarrollar una extensión para perfil multivisión, en el cual se contempla el caso de múltiples vistas y de cómo explotar la relación entre ellas para codificar y comprimir, no llegó a ser el estándar de vanguardia y llegó a su máximo desarrollo en dicha extensión.

Las ventajas de MPEG-4 hacen que este sirva de base para desarrollar modelos más eficientes ante la gran cantidad de información que los formatos de 3DTV requieren.

A continuación se realiza una descripción de MPEG-4 y de cómo este se adapta a los diferentes formatos, dependiendo la clase y cantidad de información necesarias.

2.4.1 MPEG-4 APLICABLE EN OCA, CSV Y CFC+

La recomendación ITU-T H.264, o más conocida como MPEG-4, tiene grandes ventajas y buen desempeño; si bien es cierto H.264 tiene varias extensiones aplicables a 3DTV, el MPEG-4 convencional también se puede aplicar a estos formatos, porque el video a transmitirse se lo trata como video 2D de características especiales.

2.4.1.1 Características generales del estándar H.264, MPEG-4 Parte 10.

La recomendación ITU-T H.264-AVC (Advanced Video Coding) representa una evolución a los estándares existentes para la codificación de video como son H.261, H.262, H.263. Este estándar surge ante la necesidad de procesar grandes cantidades de información digital de video y más específicamente como una solución eficiente al servicio de Televisión Digital, flujo de video por IP y las video-comunicaciones. Este estándar presenta flexibilidad total ante la variedad de redes y tecnologías existentes.

H.264 nace en 2003 y después de 10 ediciones, el 9 de marzo de 2010 se tiene la más reciente y extendida versión de H.264. En una de estas revisiones, la 2005-03, se incluyen las modificaciones al estándar de codificación para añadir cuatro nuevos perfiles referidos principalmente a las 4:2:2 y 4:4:4, con ello se mejora la calidad del video y se da la posibilidad de que el estándar se extienda hacia un rango más amplio de aplicaciones.

El estándar ha evolucionado por más de 6 años y en cada revisión se incluyeron mejoras que permitieron el uso de nuevas aplicaciones y nuevas demandas en el procesamiento de video, estas extensiones se muestran en la Tabla 2.6.

<u>FECHA</u>	<u>ESTÁNDAR</u>	<u>DESCRIPCIÓN</u>
2003-05-30	ITU-T H.264	Se aprueba por el SG-16, el estándar evolucionando a los existentes h.261, h.262, h.263.
2004-05-07	ITU-T H.264 (2003) Cor. 1	Corrección 1 al estándar por el SG-16
2005-03-01	ITU-T H.264	revisión aprobada que contiene perfiles High 4:2:2 y High 4:4:4
2006-06-13	ITU-T H.264 (2005) Amend. 1	1º rectificación al estándar, se incluye el soporte de nuevos espacios de color y se elimina el High 4:4:4
2007-04-06	ITU-T H.264 (2005) Amend. 2	Contiene extensiones “ Nuevos perfiles para aplicaciones profesionales (High 10 Intra, High 4:2:2 Intra, High 4:4:4 Intra, CAVLC 4:4:4 Intra, y perfil predictivo High 4:4:4 “ y se crean dos tipos de SEI (Supplemental Enhanced Information), el mensaje para un post-filtrado y un mensaje para información que contenga el mapa de tonalidad.
2007-11-22	Enmienda 3 "Scalable video coding" (SVC)	Mejoras del AVC, se incluyen 3 perfiles Scalable-baseline, Scalable-High y Scalable High Intraprofiles
2009-03-16	Nueva edición del H.264	Incluye MVC; como herramienta de codificación para múltiples vistas
2010-03-09	Actual H.264	Contiene una especificación de MVC, para dos vistas y lo mejora como Stereo High Profile incluyendo su SEI.

Tabla 2. 6 Evolución del estándar H.264

H.264 para hacer más fácil la estandarización ha establecido desde el inicio, y en cada una de sus enmiendas, niveles y perfiles, los cuales permiten dar una escala y una clasificación de las diferentes tasas de flujos de bits posibles. Se han desarrollado 7 diferentes perfiles con lo cual se cubre las varias aplicaciones existentes, desde redes inalámbricas hasta aplicaciones de cine digital.

Cada perfil especifica un subconjunto de velocidades de transmisión, de la sintaxis correspondiente y de los límites a los que se deben acoplar los decodificadores existentes para cada perfil. En la primera versión existen 3 perfiles: Perfil línea base, perfil principal, y perfil extendido. El primero es aplicable a servicios de video-comunicaciones en tiempo real tales como video conferencias y el video teléfono. El perfil principal está diseñado para el

almacenamiento de información digital y para servicios de broadcasting de televisión. El perfil extendido se aplica a los servicios multimedia sobre Internet. Existen además tres perfiles de alto nivel (High-Profile) definidos en las extensiones de los rangos de fidelidad para aplicaciones tales como contenido-contribución, contenido-distribución, y edición de estudio y post-procesamiento, estos perfiles son: Alto (High), Alto 10 (High 10), Alto 4:2:2 (High 4:2:2) y Alto 4:4:4 (High 4:4:4). El High Profile es capaz de soportar cuantización de 8 bits con esquema de muestreo 4:2:0 para aplicaciones de alta resolución. El perfil High 10 se basa en el mismo esquema 4:2:0, pero cuantiza las muestras con 10 o más bits, el perfil High 4:2:2 trabaja en su esquema de muestro para 10 o más bits. En el caso del perfil alto en 4:4:4 se asignan 12 bits por muestra, y se utiliza la transformada de color residual integral para codificar la señal RGB. Además de los perfiles, se definen también los niveles para cada perfil existente, los niveles corresponden generalmente al poder de procesamiento y capacidad de memoria del códec, cada nivel soporta una determinada dimensión de la imagen (SDTV, HDTV). Como se observa los perfiles tienen puntos en común y puntos divergentes, a continuación se hace una descripción más detallada de los mismos para cada perfil.

2.4.1.1.1 Elementos comunes en todos los perfiles.

Para todos los perfiles se realiza la predicción de los cuadros de la misma forma, se establecen dos tipos de cuadros, los cuadros I (cuadro intra) y los cuadros P (cuadros predictivos), los cuales dependiendo si el formato requiere más tipos de cuadros, conforman el GOP (Group of Pictures). Los cuadros I se refieren a la predicción realizada dentro del mismo cuadro I, es decir basándose en las muestras del mismo cuadro. Para los cuadros P, se hace una inter-predicción, basándose en los cuadros previamente decodificados, usando mayoritariamente un vector de movimiento y un índice de referencia para predecir los valores de las muestras de cada bloque.

Otro elemento común en todos los perfiles es el CAVLC (Context-based Adaptive Variable Length Coding), es decir una codificación de longitud variable para mejorar la entropía del código.

2.4.1.1.2 Perfil de línea Base

Los elementos exclusivos de este perfil son:

- *Orden flexible para los macrobloques:* No necesariamente pueden estar estos en orden, existe un mapa el cual asigna los macrobloques a los grupos de cuadros.
- *Orden arbitrario de los cuadros:* Puede darse la alternancia en el orden entre bloques precedentes y posteriores.
- *Redundancia de datos:* Los datos son procesados en partes, y estas porciones de datos pertenecen a los datos del código redundante obtenido por las diferencias o semejanzas de la tasa de codificación, en comparación con los datos comparados previamente en la misma porción de datos.

2.4.1.1.3 Perfil Principal

En este perfil existe otro tipo de cuadros, los cuadros B, la codificación de este grupo de cuadros se da usando inter-predicción con imágenes de referencia decodificadas previamente usando a lo sumo dos vectores de movimiento e índices de referencia para predecir los valores de muestreo de cada bloque.

Para el perfil principal se utiliza la predicción ponderada, debido a que se aplica un factor de ponderación a las muestras de datos de predicción de movimiento en el grupo de cuadros P ó B.

El código para mejorar la entropía en este caso es CABAC (Context-adaptive binary arithmetic coding), el cual consigue flujos de bits 10 % más pequeños.

2.4.1.1.4 Perfil Extendido

Este perfil incluye todas las características del perfil de línea base, y añade otros elementos como:

- *Porciones SP:* Es un caso especial de codificación para los cuadros P haciendo más eficiente la conmutación entre varios flujos de bits.
- *Porciones SI:* Similar al caso SP, pero este se da en cuadros I.

- *Particiones de Datos:* Los datos codificados son ubicados en particiones de datos separados, cada partición puede ser ubicada en una capa diferente.
- *Maneja cuadros B.*
- *Maneja predicción ponderada*

2.4.1.1.5 *Perfiles Altos (High-Profiles)*

Este perfil incluye todas las características del perfil principal, es decir, cuadros B, predicción ponderada, CABAC; y además tiene las siguientes características:

- *Tamaño del bloque adaptativo:* Puede ser 4x4 o 8x8, para muestreo de luminancia.
- *Matrices escaladas de cuantización:* Diferentes escalas de acuerdo a frecuencias específicas asociadas con los coeficientes de la transformada en el proceso de cuantización para optimizar la calidad subjetiva.

2.4.1.1.6 *Estructura por capas*

Como ya se hace referencia en las características antes descritas, H.264 trabaja por capas, y estas son 2: Network Abstraction Layer (NAL) o capa de abstracción de red y Video Coding Layer (VCL) o capa de codificación de video. La capa NAL abstrae a la VCL de tal forma que sea apropiado el transporte de datos sobre una variedad de canales de comunicación, una unidad NAL especifica al mismo tiempo tanto flujo de bytes como formatos basados en paquetes. Una unidad VCL contiene los datos codificados del video, el cual consiste en secuencias de video, imágenes, porciones de cuadros (I, B, P) y macrobloques. La secuencia de video, los cuadros o campos están destinados a tres esquemas de muestreo, uno para luminancia y dos para crominancia (RGB). El estándar puede trabajar en formato progresivo y entrelazado.

Las imágenes son divididas en slices (porciones), un slice es una secuencia de macrobloques de tamaño flexible.

2.4.1.1.7 *Algoritmo de codificación*

El algoritmo de codificación de H.264 consta de varios procesos, los cuales son descritos de manera detallada en la Figura 2.46.

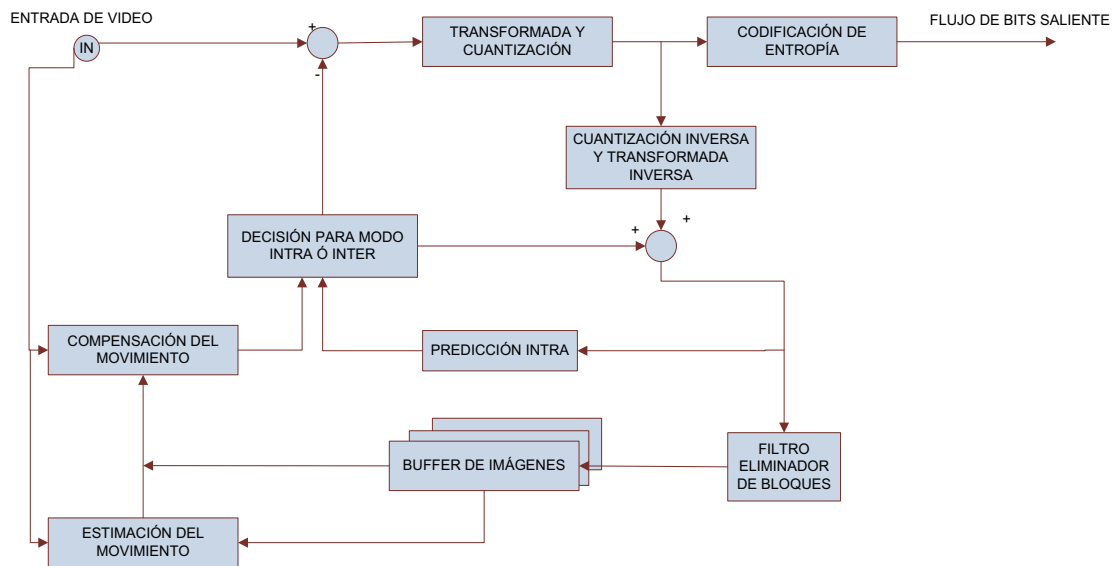


Figura 2. 46 Diagrama de bloques de la codificación H.264.

Como se observa en la Figura 2.46 el codificador tiene la posibilidad de escoger entre modos de predicción Intra e Inter para determinada región en las imágenes. La predicción se obtiene al eliminar los bloques y filtrar la señal proveniente de las imágenes reconstruidas. El vector de movimiento y los modos de predicción intra pueden ser especificados para una variedad de tamaños de bloque en la imagen, dentro del bloque de codificación de entropía H.264 establece dos métodos dependiendo el perfil: CAVLC y CABAC.

2.4.1.1.8 Predicción Intra-Frame

H.264 usa el método de predicción intra-código en macrobloques de tipo I, para reducir la gran cantidad de bits codificados que la señal de entrada genera por sí misma. Este tipo de predicción es más eficiente para el caso de secuencias de video con poco movimiento. El método se basa en intentar predecir el bloque actual que se está procesando, extrapolando los píxeles vecinos de bloques adyacentes en una serie de direcciones definidas. Para codificar un bloque o un macro bloque bajo este intra-codificación, un bloque de predicción se forma en el bloque reconstruido previamente, la señal residual entre el bloque en procesamiento y el de la predicción es codificada. Para las muestras de luminancia el bloque de predicción puede estar formado por sub-bloques de 4x4, bloques de 8x8, o macrobloques de 16x16 en nueve direcciones la predicción;

para la crominancia puede darse en bloques de 8x8 para 4:2:0, 8x16 en 4:2:2, y 16x16 en 4:4:4 en cuatro direcciones de predicción; para cada macro-bloque y para cada una de sus componentes de color se determina el modo de predicción a usarse y el conjunto de direcciones de predicción.

2.4.1.1.9 Predicción Inter-Frame

La predicción Inter-Frame está destinada a reducir la correlación temporal con ayuda de las herramientas de compensación y estimación del movimiento. Como es conocido en este estándar, la imagen puede ser dividida en macrobloques o en bloques más pequeños. Un macro-bloque de 16x16 muestras de luminancia se puede dividir en bloques más pequeños de a lo sumo 4x4. Para el modo 16x16 existen cuatro casos: 16x16, 16x8, 8x16 y 8x8. En el modo 8x8, los casos posibles son: 8x8, 8x4, 4x8 o 4x4. El bloque más pequeño requiere una mayor cantidad de bits, para la señal, el vector de movimiento y los datos extra del tipo de partición; de ahí que la elección del tamaño de la partición depende de las características de la señal de video de entrada, por lo general grandes particiones son apropiadas para áreas homogéneas en un cuadro, y las particiones de tamaño pequeño son apropiadas para áreas muy detalladas.

El proceso de predicción Inter de un bloque en muestreo involucra eventualmente también la selección de las imágenes a ser usadas como imágenes de referencia de un conjunto de imágenes decodificadas previamente almacenadas, las imágenes de referencia para la compensación del movimiento son almacenadas en el buffer de imágenes. En lo que se refiere a la imagen en procesamiento, las imágenes que le preceden y le suceden a la misma se almacenan en ese orden en el Buffer.

2.4.1.1.10 Transformación y cuantificación

Sin duda en este aspecto se ha avanzado mucho en relación a los anteriores formatos, en este caso se usa una Transformada entera a diferencia de las DCT que antes se usaban, con ello se reduce, además de con el filtro, el efecto visual de bloque. H.264 utiliza una transformación adaptativa con tamaños de bloque 4x4 y 8x8 (para los perfiles High).

Para mejorar la eficiencia en la compresión, el estándar emplea una estructura jerárquica para la transformación, en la cual los coeficientes DC (Direct Cosine), de las transformaciones 4x4 vecinas para las señales de luminancia, están agrupadas en bloques de 4x4 y transformadas nuevamente usando la transformada Hadamard³⁸, el estándar especifica la transformada Hadamard 4x4 para los coeficientes DC de la luminancia y solo para el modo Intra 16x16, y la transformada Hadamard 2x2 para los coeficientes DC de la crominancia.

Por lo general este proceso de transformación y cuantificación requiere muchas multiplicaciones haciéndolo muy complejo en el momento de la implementación. Para evitar esto, se usa en H.264, la transformada Hadamard con valores enteros haciendo más rápido y eficiente el proceso.

2.4.1.1.11 Proceso de codificación

Este proceso consta de una serie de pasos que establecen el orden de la codificación, con ello también es fácilmente identificable lo que se debe hacer en recepción para decodificar las señales.

- *Paso 1:* Se aplica la transformada entera hacia adelante, para DCT de una entrada de luminancia de 4x4 la formula resulta en :

$$X = HFH^T \text{ Con } H = \begin{bmatrix} a & a & a & a \\ b & c & -c & -b \\ a & -a & -a & a \\ c & -b & b & -c \end{bmatrix} \text{ siendo las variables a, b, c: Ec 2.19}$$

$$a = \frac{1}{2} \quad b = \sqrt{\frac{1}{2}} \cos\left(\frac{\pi}{8}\right) \quad c = \sqrt{\frac{1}{2}} \cos\left(\frac{3\pi}{8}\right) \text{ Ec 2.20}$$

Estos valores de a, b, c pueden ser simplificados por facilidad, así se tienen estos nuevos valores:

$$a = \frac{1}{2} \quad b = \sqrt{\frac{2}{5}} \quad c = \frac{1}{2} \text{ Ec 2.21}$$

³⁸ Transformada de imágenes sumamente rápida, que consiste en un desarrollo en serie de funciones básicas cuyos valores son +1 o -1, en términos matemáticos la transformada H viene dada por: $H(u) = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \cdot (-1)^{\sum_{i=0}^{u-1} b_i(n) b_i(u)}$

La multiplicación en el proceso de transformación se evita al integrarla con la cuantización, de esta forma se obtiene:

$$X = \bar{H}F\bar{H}^T \otimes SF \quad \text{Ec.22} \quad \bar{H} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \text{ y } SF = \begin{bmatrix} a^2 & \frac{ab}{2} & a^2 & \frac{ab}{2} \\ \frac{ab}{2} & \frac{b^2}{4} & \frac{ab}{2} & \frac{b^2}{4} \\ a^2 & \frac{ab}{2} & a^2 & \frac{ab}{2} \\ \frac{ab}{2} & \frac{b^2}{4} & \frac{ab}{2} & \frac{b^2}{4} \end{bmatrix} \quad \text{Ec 2.23}$$

La multiplicación indicada en la expresión de X se la hace elemento a elemento en las matrices señaladas.

- *Paso 2:* Post-escalamiento y cuantización, la matriz SF (post-escalamiento) permite obtener una señal Y transformada y cuantizada. El tamaño del paso en la cuantización se denomina $Qstep$.

$$Y_{ij} = X_{ij} \text{entero} \left(\frac{SF_{ij}}{Qstep} \right) \quad \text{Ec 2.24}$$

H.264 define un total de 52 valores de paso para la cuantización.

2.4.1.1.12 Codificación de entropía

Anterior a este estándar la forma de codificar era usando VLC's (Variable Length Codes), en el cual la longitud de la palabra código tenía que ver con la probabilidad de ocurrencia; para H.264 los VLC's funcionan de forma diferente. Todos los elementos de sintaxis a excepción de los datos residuales, son codificados con el código Golomb Exponencial³⁹. Para los casos residuales en cambio se usa CAVLC, y para el perfil principal y para los altos perfiles se utiliza el CABAC, el cual presenta mayor eficiencia pero mayor complejidad comparándolo a CAVLC.

- **CAVLC:** Código adaptativo de longitud variable basado en el contexto, después de la transformación y la cuantización la probabilidad que los coeficientes sean 0, +1 ó -1 es muy alta. CAVLC maneja los 0, +1 ó -1 de

³⁹Un código exponencial Golomb de orden k es un tipo de código universal, donde el parámetro k es un entero no negativo, para H.264 se utiliza $k=0$.

diferente modo con los niveles de los coeficientes, el total de número de ceros y +1 ó -1 son codificados, y para otros coeficientes se codifica con su nivel.

- **CABAC:** Código adaptativo aritmético binario basado en el contexto, se alcanza un buen nivel de compresión, el modelo de probabilidad para cada símbolo se basa en 3 procedimientos (ver Figura 2.47), los cuales son: Binarización, modelo de contexto y Código aritmético binario.

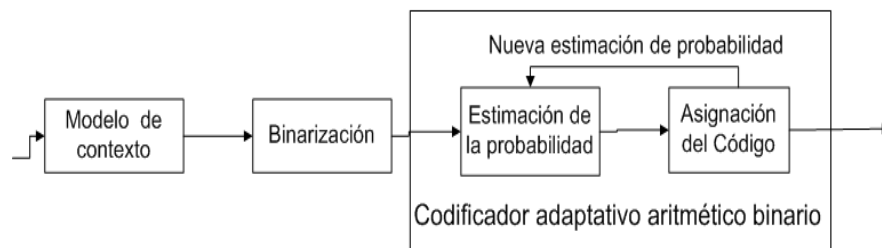


Figura 2. 47 Procesos de CABAC

2.4.1.1.13 Tasa de transmisión.

Con todos los argumentos presentados para cada perfil, la tasa de transmisión original del video digital se reduce en cada etapa y proceso que el códec efectúa, por ejemplo el filtro de suavizado de bloques ya disminuye en 5 a 10% la tasa binaria. Sin duda la eficiencia de este código es superior al de estándares anteriores; y para determinados PSNR⁴⁰, se tienen mejores tasas binarias que con tecnologías anteriores, por ejemplo se obtiene una tasa binaria 50% menor a la que se obtiene con MPEG-2, llegando a tasas de compresión de 50:1. Dependiendo del perfil, la tasa de bits es diferente, dado que para diferentes aplicaciones se requiere mayor calidad. Lo que respecta a este trabajo es el ámbito de la TV digital SD y HD, por lo cual los perfiles de interés son el perfil principal y los perfiles High en los niveles correspondientes, con estas consideraciones las tasas binarias se encuentran entre los 64 Kbps y 150 Mbps,

⁴⁰ La relación a señal a ruido pico PSNR, define la relación entre la energía máxima de la señal y la del ruido, es muy usada para medir la calidad de una imagen reconstruida de la codificación; la PSNR de una imagen viene dada por: $PSNR = 10 \log_{10} \left(\frac{MAX_i^2}{MSE} \right)$, donde MAX denota el máximo valor que puede tomar un píxel en la imagen y MSE es el error cuadrático medio.

es posible que existan tasas mayores para servicios como SHDTV, no aplicables a nuestro estudio.

2.4.1.2 Aplicación del H.264-AVC, MPEG-4 parte 10 a la TV3D.

Teniendo en cuenta que el estándar analizado en el punto anterior es muy eficiente y presenta mejoras significativas ante los estándares anteriores, se puede decir que al aplicarse este sobre algunos formatos de TV3D se pueden tener resultados bastante aceptables. Para decidir sobre a cuáles formatos es aplicable el H.264 AVC, se debe analizar los siguientes aspectos:

- Que la señal a codificarse mantenga los parámetros básicos de una señal de video convencional.
- Que si se requiere codificar más de una señal, el conjunto Video-out, no configure una señal que demande excesivos recursos al canal de transmisión.
- Dado que H.264 da mucha importancia a la calidad para la codificación de video, las señales a tratarse deberán conservar los parámetros de calidad y mínima distorsión después de esta etapa.

Analizando los puntos anteriores, se concluye que los formatos a los que en primera instancia se les puede aplicar son: OCA, CSV y CFC+, para el caso de V+D, el video 2D también se puede codificar con MPEG-4, pero la parte de la profundidad merece un análisis especial.

Analizando el caso de los primeros 3 formatos, el estándar puede aplicarse de forma idéntica a la ya descrita, es decir debería funcionar de la siguiente manera:

2.4.1.2.1 H.264 AVC en OCA (Optimized Chromatic Anaglyph)

En este caso la naturaleza de la señal de video es similar a las señales convencionales de video 2D, por lo tanto el estándar es fácilmente aplicable sin ninguna restricción, tal como lo muestra la Figura 2.48

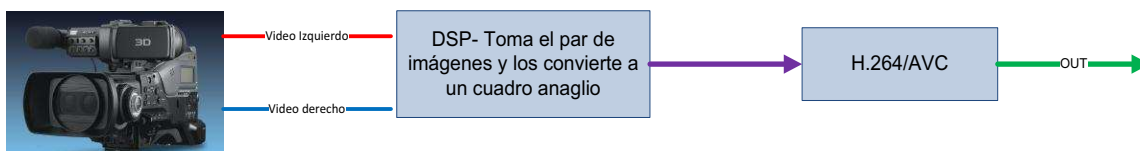


Figura 2. 48 Codificación para OCA

Cabe mencionar que se debe procurar mantener la calidad, por lo que trabajar en HD sería una idea conveniente en este formato. Para OCA, no existen señales auxiliares ni información extra para el video 3D.

2.4.1.2.2 H.264 AVC para CSV (Conventional Video Stereo).

Esta es la aplicación natural a la parte estéreo, considerando que aquí las señales de video son monoscópicas e independientes. Dentro de CSV, podemos considerar 2 casos como los formatos antes estudiados, el caso del Simulcast y del dual-streaming; con esto H.264 trabaja normalmente sobre cada uno bajo la estructura expuesta en las Figuras 2.49 y 2.50

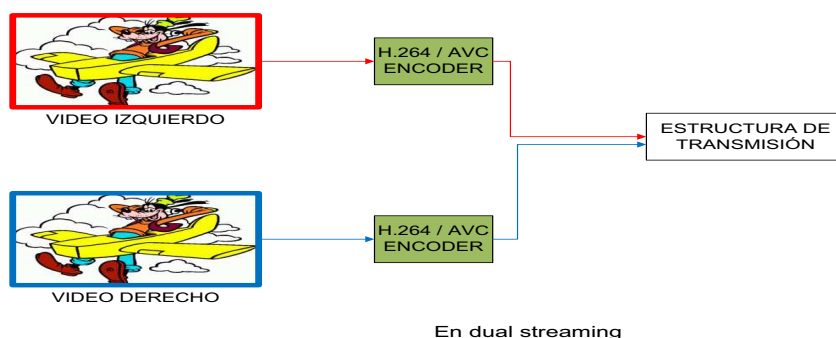


Figura 2. 49 Aplicación de H.264 en Dual-Streaming.

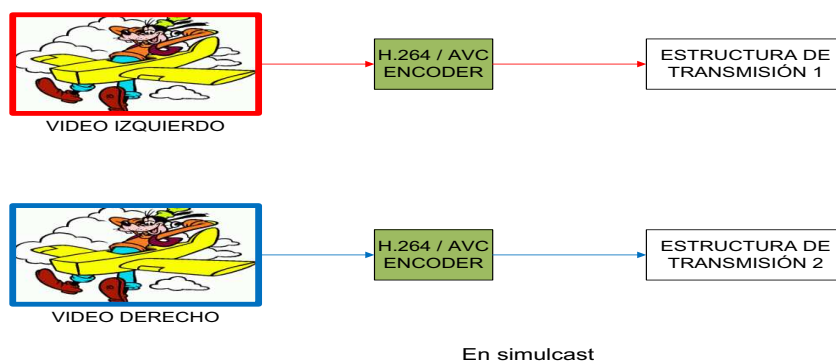


Figura 2. 50 Aplicación de H.264 en Simulcast.

En los dos casos la codificación es totalmente independiente, por lo que el codificador trabaja con señales completamente convencionales.

Las desventajas que se presentan son: para dual Streaming, se debe tener en consideración que la estructura de transmisión soporte el video-out proveniente de la codificación de dos señales, y en el caso de Simulcast, en el que la estructura de transmisión si soporta la salida del codificador, se debe considerar que el uso del canal de transmisión en este esquema de transmisión se duplica. Al ser informaciones completamente independientes estos esquemas se pueden ajustar a los formatos V+D, utilizando en lugar del video derecho a la información de profundidad.

2.4.1.2.3 H.264 AVC para CFC+ (Conventional HD Frame Compatible +)

CFC+ es el formato en el cual la información tiene una naturaleza propia para 3DTV, por lo cual la cantidad de información ya crece; como se estudió en el formato, aquí se manejan al menos las dimensiones del cuadro HD y para casos de mayor calidad, dimensiones superiores. Dado que MPEG-4 si establece algunos niveles y perfiles en los cuales encajen estos videos, la codificación si sería aplicable.

Dado que en este formato se manejan los cuadros L+R bajo algún empaquetamiento, es necesario utilizar la extensión de H.264 que define un SEI (Supplemental enhancement information), es decir un señalizador de información adicional. En este SEI se indica bajo que formato se empaqueta la información, y que tipo de información existe. Para mejorar la calidad en la codificación a esta información de video 3D se le aplica el SVC (Scalable Video Coding), el cual no es más que un anexo del H.264 para señales de mayor calidad y por supuesto compatible con los parámetros de codificación antes estudiados.

En la Figura 2.51 se ilustra como actuaría el codificador para casos de empaquetamiento en un cuadro HD. Cabe señalar que el tipo de empaquetamiento es totalmente independiente, ya que mientras se ocupe las dimensiones normadas para los cuadros HD, el video estereoscópico se codifica como señales normales, aún cuando estas lleven información estereoscópica.

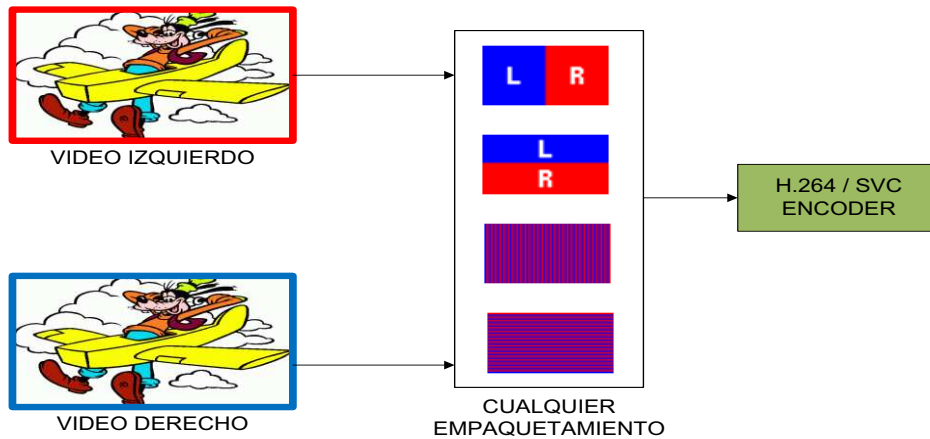


Figura 2. 51 Aplicación de H.264 en CFC+

Debido al uso del cuadro HD, es necesario utilizar la señalización correspondiente, dicha señalización se encuentra en la norma HDMI 1.4a en la porción correspondiente a 3D (Anexo 1).

2.4.2 CODIFICACIÓN PARA EL FORMATO V+D.

La idea inicial de establecer el mapa de profundidad surge del proyecto ATTEST⁴¹, y se basa en que una señal monocromática ahorraría muchos recursos de transmisión y facilidad en codificación. Existen algunas técnicas que resultan una solución viable en este formato, procurando siempre el uso eficiente de los recursos y evitar en lo posible que se demanden tasas de bits excesivamente altas.

El desarrollo de estas técnicas ha permitido que este formato sea eficiente e incluso sea aplicable a la televisión móvil 3D. Es posible también aplicar la extensión MVC de H.264 tomando en cuenta que una de las vistas puede ser el mapa de profundidad.

Hoy en día existe una nueva técnica de codificación basada en el algoritmo Platelet, sin embargo no se ajusta a los parámetros de codificación de video para ISDB-T.

⁴¹ Advanced Three-dimensional Television System Technologies, proyecto para la investigación de un sistema de difusión eficiente de televisión 3D.

2.4.2.1 MPEG-C parte 3

Este estándar conocido también como el ISO/IEC 23002-3, establece el procesamiento separado e independiente de las señales de video y de la señal de profundidad, tomando en cuenta que codificar una señal de profundidad monocromática requiere menos tasa de bits y que en recepción se debe contar con dos decodificadores para separar la reconstrucción de ambas señales.

Cabe señalar que el estándar no establece técnicas de compresión o transmisión, solo establece la metodología que puede ser aplicada para nuestro estudio con técnicas H.264/AVC.

La cadena de transmisión para este formato se representa en la Figura 2.52.

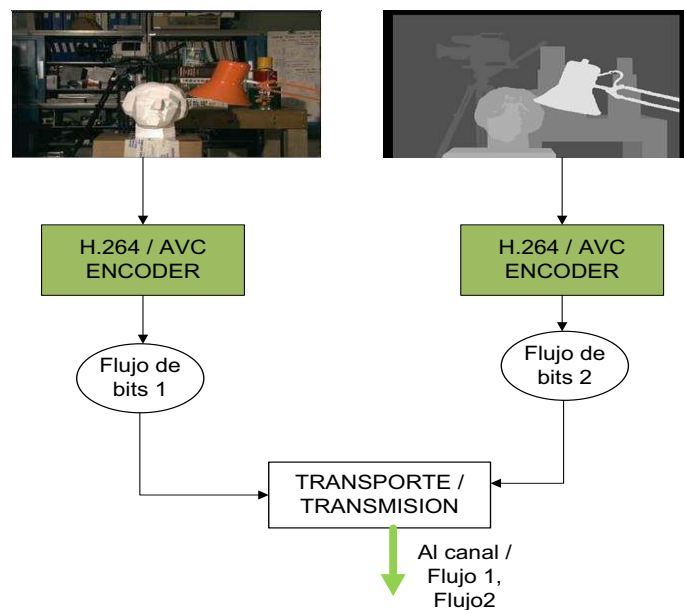


Figura 2. 52 Cadena de Transmisión para MPEG-C parte 3.

Se considera que el codificador para la profundidad tendrá como salida una tasa de bits de entre el 5 y 10% de la tasa del video 2D, siendo así la tasa total de bits en V+D usando codificación para cada señal sería en total un 110% un valor que puede resultar cuestionable para servicios donde el uso del canal sea realmente limitado.

El esquema de la figura 2.52 muestra la codificación independiente de las 2 señales, de lo cual resultarán dos flujos de bits; para la transmisión de dichos

flujos, estos deben ser multiplexados cuadro a cuadro generando un único transport-stream, el cual puede eventualmente contener parámetros adicionales del mapa de profundidad como información auxiliar.

2.4.2.2 H.264 - Imágenes auxiliares.

H.264 sostiene la posibilidad de codificar señales auxiliares de video, teniendo las mismas características de macro-bloques, de semántica y sintaxis en la codificación.

Se pueden adicionar señales de video redundantes o señales no especificadas en la norma H.264, de esto se vale la 3DTV para poder transmitir sobre MPEG-4 la señal auxiliar del mapa de profundidad (ver Figura 2.53). Esta señal debe estar debidamente señalizada según lo establece el estándar, a este tipo de señales el estándar las denomina como “Auxiliary Coded Picture”, y se definen como imágenes que suplementan a la señal primaria, normalmente esta señal se usa como redundancia y el estándar restringe a que esta sea monocromática.

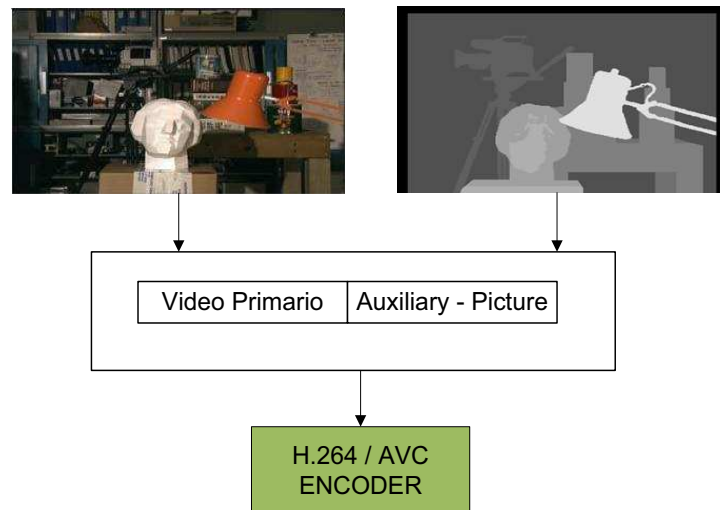


Figura 2. 53 Aplicación del codificador H.264 a las señales de imágenes auxiliares.

Esta solución permite aprovechar al máximo las ventajas de H.264 y además se brinda un gran nivel de compatibilidad con los sistemas anteriores, además de un uso eficiente en la transmisión.

2.4.3 CODIFICACIÓN PARA EL FORMATO MVC.

Esta enmienda a H.264/AVC crea la posibilidad de explotar al máximo la dependencia entre las vistas de una señal de video estereoscópico, haciendo uso de una predicción tomando en cuenta los n cuadros presentes. Para este formato se pueden especificar dos casos, el primero en el cual se generaliza la predicción para n vistas, y el segundo que se ha mencionado como estéreo de alto desempeño si $n=2$. Un nuevo SEI se añade según H.264 para indicar el uso de la predicción y codificación multivista. Un codificador MVC consiste básicamente en N codificadores simples en paralelo, cada uno de ellos utiliza estructuras de predicción temporal, donde una secuencia de imágenes sucesivas es codificada como Intra (I), Predictiva (P) o Bi-predictiva (B). Para mejorar aún más la eficiencia de la codificación se tiende a utilizar Imágenes B jerárquicas, donde la jerarquía de una imagen B es creada por la sucesión de imágenes B que resultan referencias para otras imágenes B.

En MVC los conceptos de predicción para una imagen se extienden a tal punto que una imagen que está en el proceso de codificación, puede tener tanto referencias temporales como referencias intervista para la predicción. Para explicar esta predicción MVC considerando $n=5$ y un GOP de tamaño 8, la Figura 2.54 ilustra el proceso.

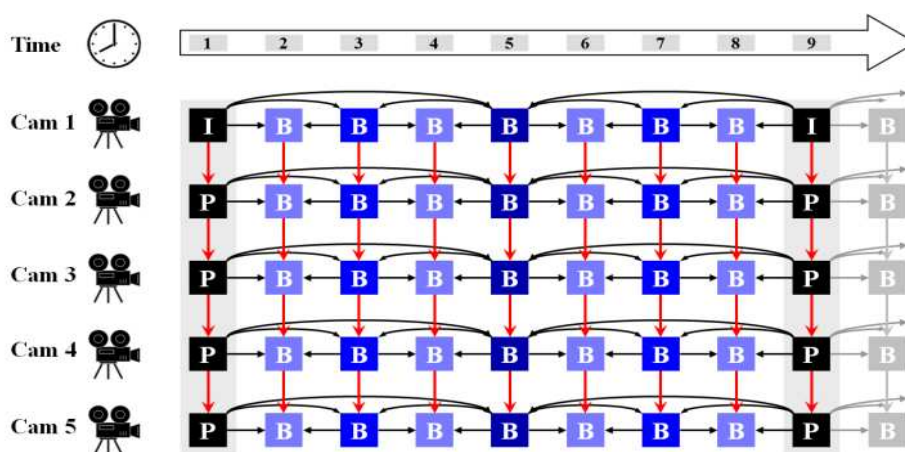


Figura 2. 54 Predicción Multivista para $n=5$ y $\text{GOP}=8$ (IPP).⁴²

⁴² Fuente: "3D Video Formats and Coding Methods"; Karsten Müller, Philipp Merkle, Gerhard Tech, and Thomas Wiegand; HHI; 2006 Berlín Alemania

Para la cámara 1, el procedimiento de predicción es idéntico al caso de una sola vista, manteniendo así una compatibilidad con sistemas anteriores, pero para el resto de vistas se hace uso de la referencia entrevista (flechas rojas) para la predicción.

Con el fin de mejorar la eficiencia existen varias estructuras de predicción desarrolladas, dos de las más importantes son IPP (inter-view prediction at key picture) y HBP (Hierarchical B Pictures), IPP funciona bajo lo expuesto en la figura 2.54.

HBP explota aún más la relación inter-vista, no solo utilizando las vistas $n+1$ para la predicción sino también las $n+2$. La Figura 2.55 muestra el proceso de predicción de las diferentes imágenes para un GOP = 8 y $n=8$ vistas.

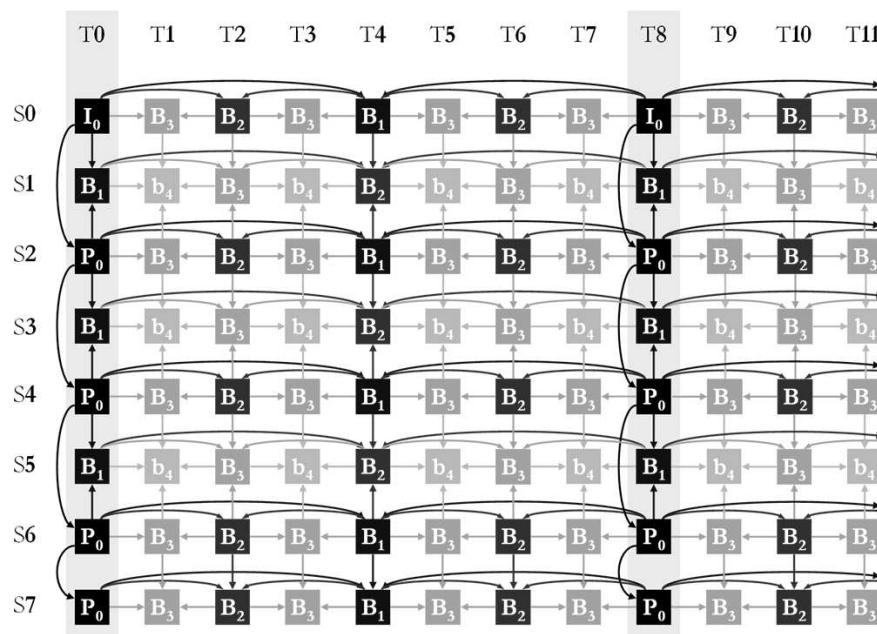


Figura 2. 55 Predicción Multivista para $n=5$ y $\text{GOP}=8$.⁴³

El codificador como tal se implementa mediante la sucesión de n codificadores (uno por cada vista) y relacionándolos con la información del resto de las vistas. Al final la codificación de cada vista forma parte del conjunto de datos de salida del codificador.

⁴³ Karsten Müller; "Representation and Coding Formats for 3D Video"; Conference Session 3DTV Content; IEEE Broadcast Technology Society; 2009 Alemania.

El diagrama de un codificador MVC, se conforma de acuerdo a lo expuesto en la Figura 2.56.

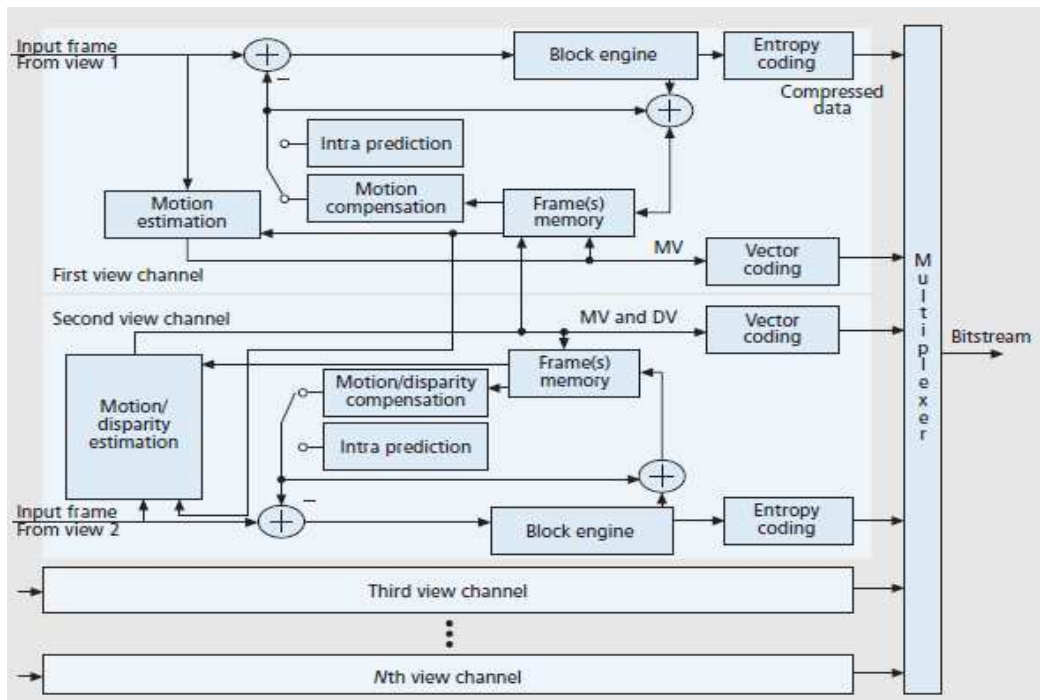


Figura 2. 56 Codificador MVC de n vistas.⁴⁴

Como se mencionó anteriormente, en principio existen n estructuras codificadoras por vista y estas se relacionan en el bloque de estimación de movimiento y en la memoria donde se almacenan las referencias. Las n vistas después del proceso se multiplexan para ordenarlas y obtener un Bitstream único de la señal 3D.

2.4.4 CODIFICACIÓN PARA EL FORMATO MVD

Para formatos avanzados como este, la codificación se vuelve compleja debido a la cantidad de información que se genera, MVD resulta una combinación de V+D y MVC. La codificación es en principio basada en la técnica MVC original, no obstante en términos prácticos se puede hablar de una codificación MVD.

La codificación en MVD funciona, dentro de un esquema global, de acuerdo a lo expuesto en la Figura 2.57

⁴⁴ Fuente: "Video Encoder Design for HD 3D Video Communication Systems"; Pei- Tsung, Ding, Chen, Chuang, Chen, Hsiao, Chien, and Chen, National Taiwan University, IEEE Com. Magazine, Abril 2010.

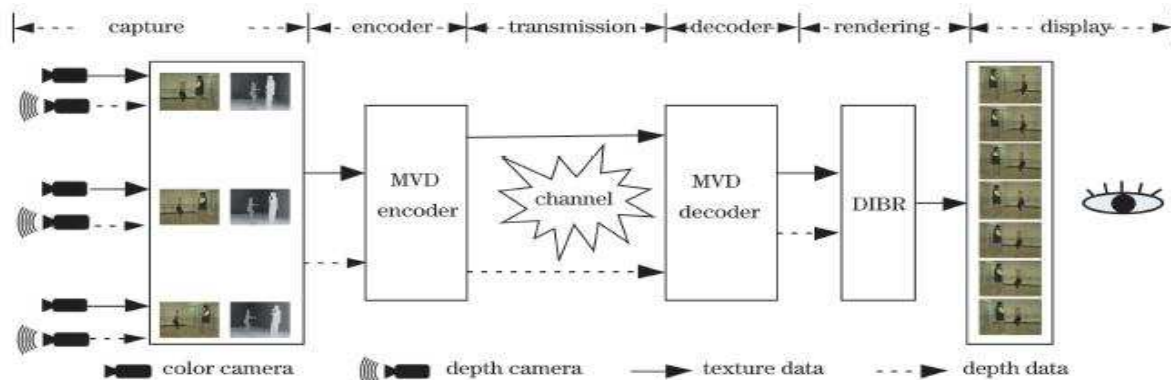


Figura 2. 57 Cadena de transmisión en MVD.⁴⁵

Se observa que el codificador genera una unida señal para la transmisión, la cual contiene la información de las n vistas y su respectiva profundidad. El modo de predicción en este caso varía ya que toma, adicionalmente, la información de profundidad. Este proceso de predicción se detalla en la Figura 2.58

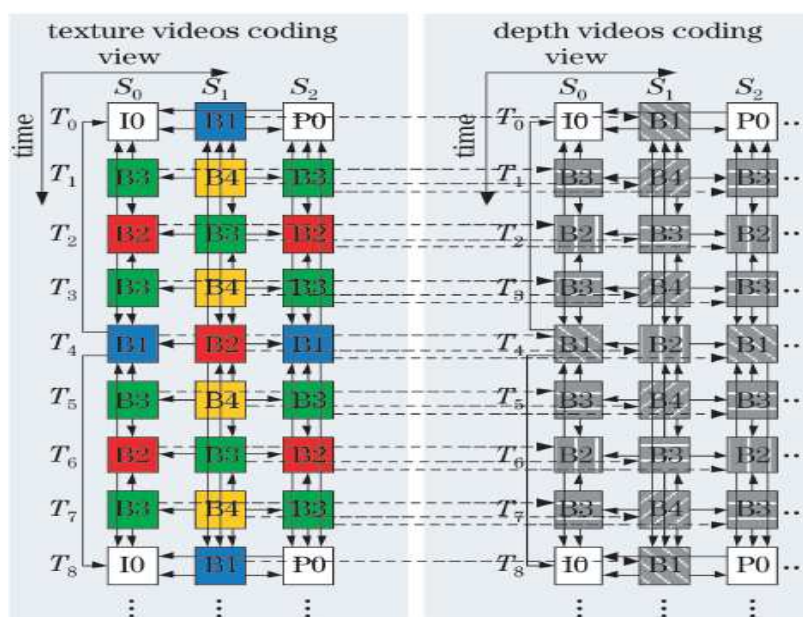


Figura 2. 58 Modo de predicción de imágenes para MVD.

La codificación y predicción son basadas en MVC, tomando en cuenta que se van a generar más señales del video, el diagrama de bloques del codificador es más complejo, tal y como lo ilustra la Figura 2.59.

⁴⁵ "Compression for 3DTV - with special focus on MPEG standards"; Dr.-Ing. Aljoscha Smolić Fraunhofer HHI Image Processing Department; Alemania 2007. (La Figura 2.58 se refiere a la misma fuente)

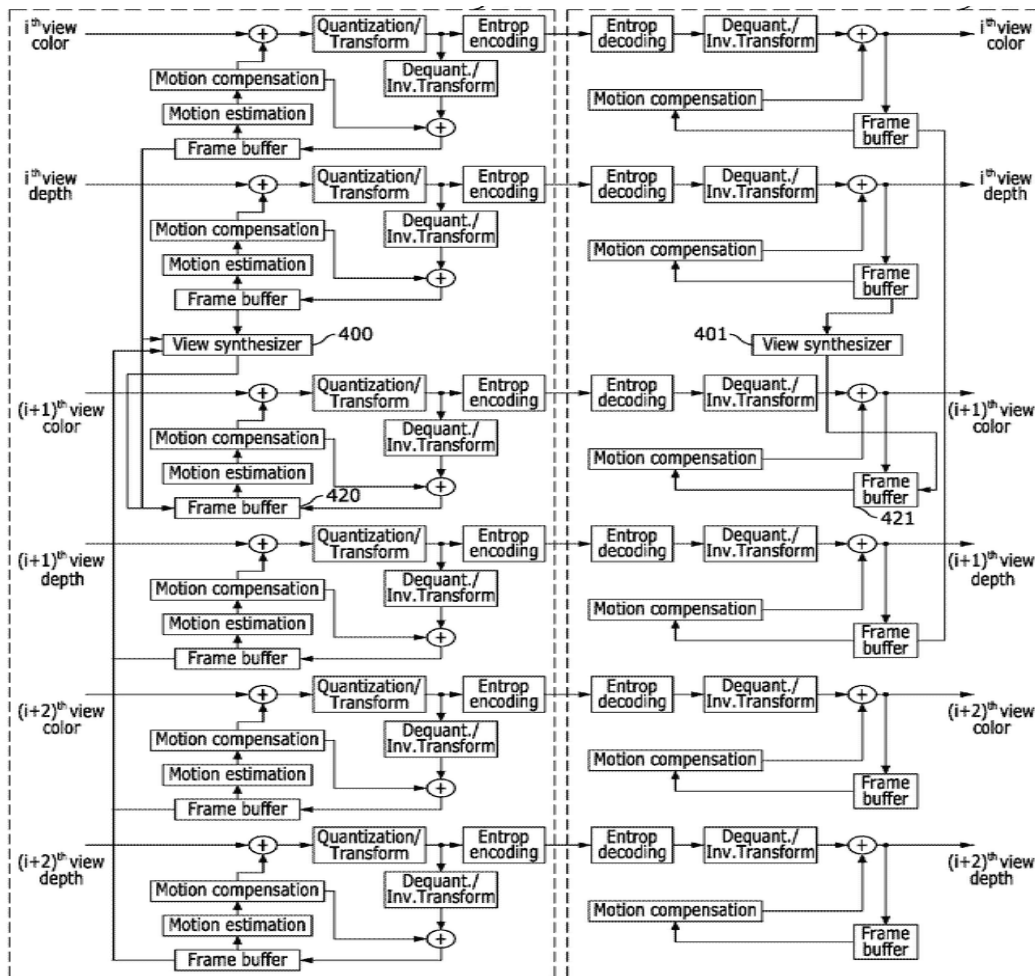


Figura 2. 59 Diagrama de bloques de un codificador MVD.⁴⁶

2.4.5 CODIFICACIÓN PARA EL FORMATO LDV

Este formato es sin duda más preciso y de mayor calidad que los anteriores, y por ello su codificación también resulta más compleja. Se puede aplicar el mismo MVC a estas señales, haciendo un proceso similar al de MVD, donde las informaciones adicionales entran como vistas extras al codificador MVC. Bajo esta premisa se puede aplicar el H.264 sin problema y se guardaría la compatibilidad con ISDB-T.

Es notorio que se puede recurrir en una complejidad mayor para los codificadores y en un flujo de bits mucho mayor al que se pueda manejar para broadcasting. Por ello la compañía 3DTV Solutions™ en pro de la calidad ha propuesto todo un

⁴⁶ Fuente: "Multi-view video plus depth representation and coding"; Philipp Merkle, Aljoscha Smolic, Karsten Müller, and Thomas Wiegand; Berlin Alemania, 2009

esquema para este formato basándose en su propio esquema de codificación LDV, donde se explota al máximo la relación entre todas las capas y las vistas permitiendo en lo posterior una efectiva reconstrucción a tasas de bits más accesibles. Sin embargo una desventaja de esta propuesta es que requiere de gran exactitud en los mapas de profundidad, de no ser así se recaerá en resultados erróneos. Para el detalle del proceso de codificación y compresión de este sistema referirse al Anexo 2 (el Anexo 2 cubre detalles de otros formatos como MVC y V+D)

2.4.6 CODIFICACIÓN PARA EL FORMATO DES.

No cabe duda que para este formato se debe manejar el mismo criterio que para LDV y MVD; a tal cantidad de información si bien es cierto MVC se puede aplicar, ya resultaría de gran complejidad su implementación, aquí también es una alternativa el algoritmo Platelet codificador como solución, aunque no esté dentro del ISDB-T.

2.4.7 TASAS DE BITS RESULTANTES DE LA CODIFICACIÓN

Dependiendo del formato elegido y del esquema de codificación, se pueden tener uno o más flujos de bits resultantes como señal de video. En esta sección se analizará la tasa de bits que cada formato genera y una comparación entre ellos, con el objetivo de determinar las características que servirán para el análisis del transporte y la transmisión de la señal.

En la mayoría de los casos se presentan inicialmente valores para resoluciones diferentes a las de una señal SDTV o HDTV, por lo cual se hace un posterior análisis y estimación de lo que se puede dar en ISDB-T.

2.4.7.1 Tasa de bits para OCA

En este formato se genera un solo flujo de video proveniente del DSP, que fusiona dos imágenes en una del tipo cromático anaglífico, por lo tanto el flujo de bits resultantes estará dentro de lo normal para señales SDTV o HDTV de mejor calidad es decir en un máximo de 4Mbps para definición estándar y 12 Mbps para alta definición.

2.4.7.2 Tasa de bits para CSV

Si bien es cierto existen dos posibilidades, ambas generan 2 flujos de datos que serán transportados y transmitidos de diferente manera, por ello en este sistema se demandan muchos recursos respecto a la tasa de bits, es decir se genera un flujo doble al convencional, lo cual significa que para definición estándar se genera alrededor de 8 Mbps y en alta definición 24 Mbps. Las tasas de transmisión pueden variar de acuerdo a la codificación de entropía realizada es así que para definición estándar las tasas de bits se enmarcan dentro de lo expuesto en la Figura 2.60.

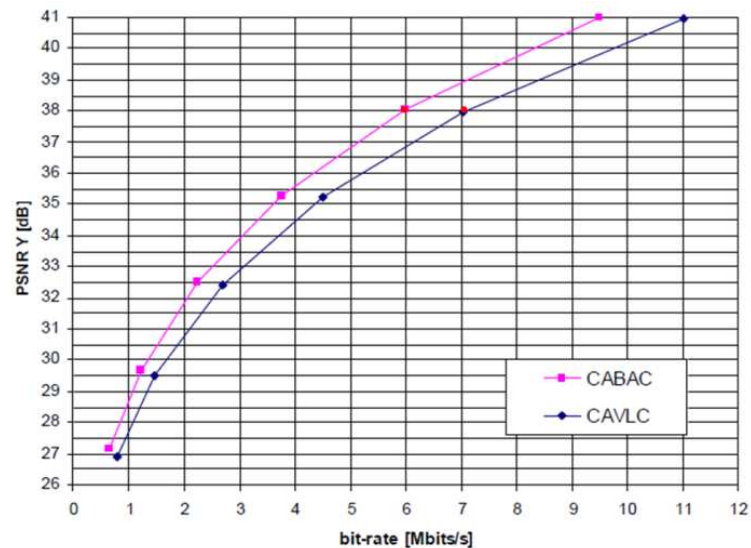


Figura 2. 60 Tasa de bits en SDTV.

2.4.7.3 Tasa de bits para CFC+

Para los cuadros HDTV, la tasa que se maneja en la adquisición, esta alrededor del 1.5Gbps, tomando en cuenta resolución 1920x1080 por pantalla y esta es renovada 30 veces por segundo, asignando 24 bits por muestra:

$$R = 1920 \times 1080 \times 30 \times 24 = 1499292000 = 1.5 \text{ Gbps} \quad \text{Ec 2.25}$$

Que para MPEG-4 se comprime en relaciones de compresión 90:1, 125:1, 180:1 obteniendo tasas de 16, 12, 8 Mbps respectivamente.

Como se estudió en el estándar de compresión, existen varios perfiles y modos de codificación para imágenes de alta definición. Por ejemplo, tomando en cuenta el perfil High y un video 720p y 1080x1920 la tasa de bits resultante se expone en las Figuras 2.61 y 2.62.

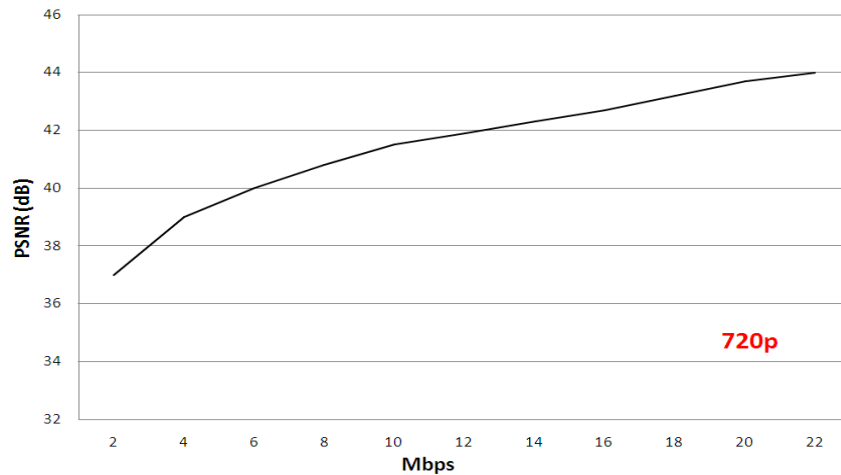


Figura 2. 61 Tasa de bits en 720p.

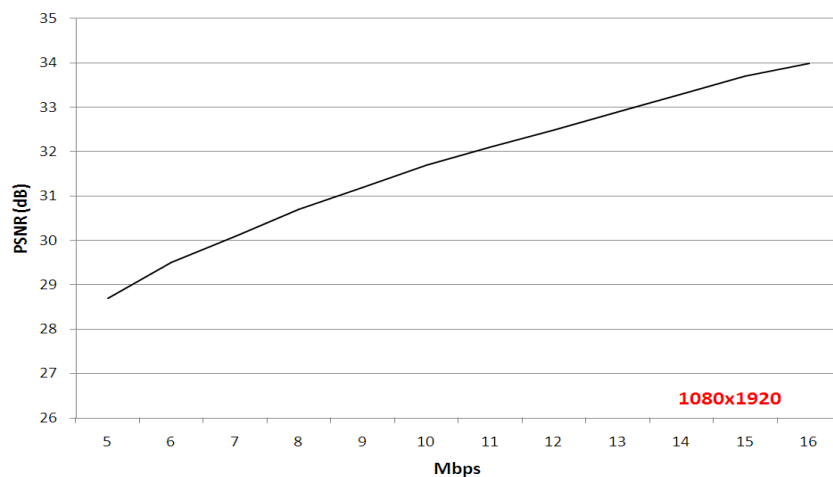


Figura 2. 62 Tasa de bits en 1080x1920.

2.4.7.4 Tasa de bits para V+D

El análisis en este caso se realiza para las dos posibilidades expuestas en 2.4.2.1 y 2.4.2.2. Es importante señalar que para este y todos los formatos que incluyan profundidad (MVD, LDV, DES), la eficiencia de las tasas de transmisión será mayor mientras los mapas de profundidad sean mejor estimados y más exactos.

2.4.7.4.1 Tasa de bits para MPEG C parte 3.

En este caso, el codificador trabaja en dos partes la primera obedece a una imagen normal que puede tener relación con los bitrate expuestos anteriormente dependiendo si esta tiene alta definición o definición estándar (Figura 2.61 y 2.62). Suponiendo una señal de video de definición estándar promedio de entre 3 y 4 Mbps para la señal 2D, el codificar su mapa de profundidad requeriría las tasas que detalla la Figura 2.63.

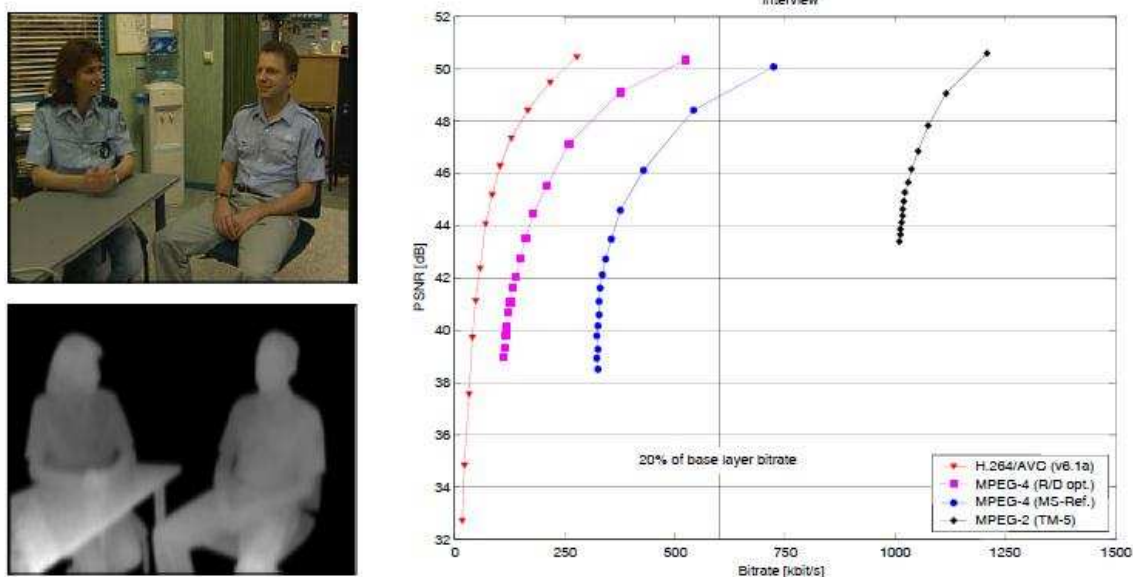


Figura 2. 63 Tasa de bits al codificar el mapa de profundidad.

En la Figura 2.63 se muestran tasas para diferentes codificaciones, incluyendo algunas versiones del MPEG-4, si se toma en cuenta la versión más reciente se observa que para un valor de PSNR de 46.29 dB se requieren alrededor de 105 Kbps, apenas un 3.5% de la señal 2D. En general se observa que esta información se codifica por debajo del 20% de la información de color, vale recalcar que se puede aplicar a señales HDTV, obviamente sabiendo que se tendrá mayor tasa de bits para el mapa de profundidad.

Estimando las tasas de bits para una señal de SDTV, que en promedio puede tener 3.5 Mbps, y considerando una PSNR de 42 dB (5% de la información de color), para una buena calidad:

$$R_{(v+d)sdv} = R_v + R_d = 3.5Mbps + 125Kbps = 3.625Mbps \quad \text{Ec 2.26}$$

Estimando las tasas de bits para una señal de HDTV, que en promedio puede tener 12 Mbps, y considerando una PSNR de 40 dB (8% de la información de color), para una buena calidad:

$$R_{(v+d)hdtv} = R_v + R_d = 12Mbps + 960Kbps = 12.960Mbps \quad \text{Ec 2.27}$$

2.4.7.4.2 Tasa de bits para video auxiliar - H.264

En este caso se genera un solo flujo de datos, la relación entre las tasas de bits del video 2D y de la profundidad tiene una proporción similar al caso anterior, por debajo del 15%, no obstante el usar un solo codificador representa una optimización y un ahorro para los sistemas de recepción. Las pruebas realizadas para este esquema se han hecho bajo la LDTV (Low Definition TV), usando el one-seg, es decir con videos de tasas que van desde 250 kbps hasta 1500 kbps aproximadamente, con una resolución de 480x272 pixeles a 30 fps, las pruebas se han realizado variando algunos parámetros del codificador como el GOP (ver Figura 2.64). Si bien es cierto no se presentan resultados en los casos de HDTV y SDTV, la relación de las tasas de bits es proporcional de acuerdo al incremento de la calidad.

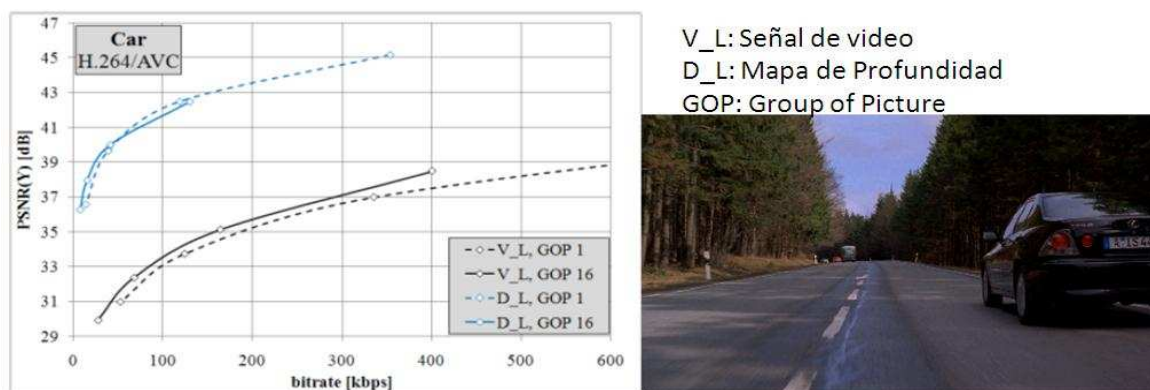


Figura 2. 64 Tasa de bits al codificar el mapa de profundidad.

En este caso se cuenta con un video LDTV, que a 38 dB de PSNR, tiene una tasa de 400 Kbps, y cuyo mapa de profundidad se codifica por debajo de los 20 Kbps. Con esto, se obtiene una señal fácilmente manejable dentro de los rangos de

videos convencionales, en general se observa que la tasa de transmisión en este formato crecerá entre un 5 y un 10%. Para una señal SDTV estimando que la profundidad toma un 7% de la información de color la tasa oscilaría entre los 3.21 Mbps y 4.28 Mbps.

Es necesario señalar que es posible utilizar la extensión MVC de H.264, para codificar profundidad, es decir haciendo que MVC trabaje con dos vistas y una de ellas sea la información de profundidad; y por otro lado codificar la profundidad basándose en Platelet. En todos los casos se genera un solo flujo de bits ya que la información de color y la de profundidad se procesan en el mismo códec H.264.

2.4.7.5 Tasa de bits para MVC

A continuación se detallará el uso de MVC en multivista con al menos información L+R, donde las exigencias son mayores pero de igual forma la calidad es superada. Si bien es cierto el MVC es adaptable a n vistas ($n > 2$). La mayor cantidad de pruebas se han realizado para $n=2$. A continuación se presentan estos resultados (ver Figura 2.65) para $n=2$ y $n=8$. Para el caso de 2 vistas los parámetros de prueba han sido: 480x272 de resolución, se presentaron 235 cuadros a 30 fps, CABAC y GOP variable al igual que el parámetro de cuantización (el video es el mismo analizado en 2.4.7.4.2).

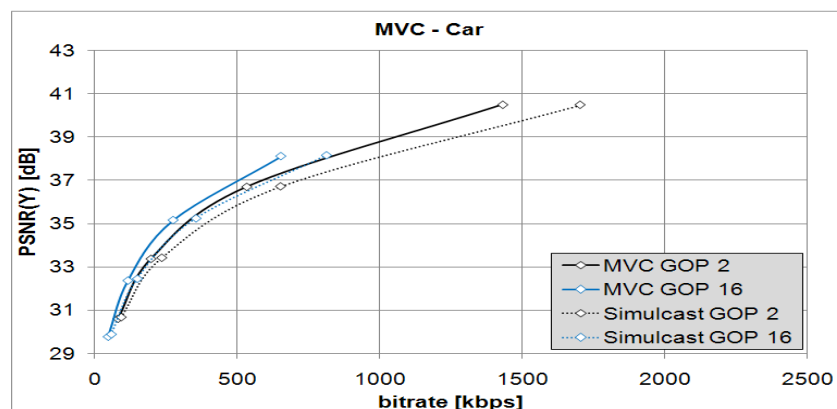


Figura 2. 65 Tasa de bits al codificar con MVC y $n=2$.

La figura muestra tanto el resultado para MVC como para Simulcast, debido a que MVC surge como una mejora para codificar 2 vistas que anteriormente se codificaron independientemente.

De la Figura 2.62 se puede observar que a 40 dB, MVC codifica dicha secuencia a 1400 Kbps, entonces:

$$480 \times 272 = 130560 \frac{\text{pixels}}{\text{cuadro}} \quad \mathbf{Ec\ 2.28}$$

$$130560 \frac{\text{pixels}}{\text{cuadro}} \times 30 \frac{\text{cuadros}}{\text{segundo}} = 3916800 \frac{\text{pixeles}}{\text{segundo}} \quad \mathbf{Ec\ 2.29}$$

$$1400 \frac{\text{Kbits}}{\text{segundo}} \times \frac{1 \text{ segundo}}{3916800 \text{ pixeles}} = 0.357434 \frac{\text{bits}}{\text{pixeles}} \quad \mathbf{Ec\ 2.30}$$

Con dicha relación, estimando la tasa de bits para una señal SDTV (720x483):

$$720 \times 483 = 347760 \frac{\text{pixels}}{\text{cuadro}} \quad \mathbf{Ec\ 2.31}$$

$$347760 \frac{\text{pixels}}{\text{cuadro}} \times 30 \frac{\text{cuadros}}{\text{segundo}} = 10432800 \frac{\text{pixeles}}{\text{segundo}} \quad \mathbf{Ec\ 2.32}$$

$$10432800 \frac{\text{pixels}}{\text{segundo}} \times 0.357434 \frac{\text{bits}}{\text{pixeles}} = 3.729 \text{ Mbps} \quad \mathbf{Ec\ 2.33}$$

De igual forma, estimando la tasa de bits para una señal HDTV (1920x1080):

$$1920 \times 1080 \times 30 = 62208000 \frac{\text{pixels}}{\text{segundo}} \quad \mathbf{Ec\ 2.34}$$

$$62208000 \frac{\text{pixels}}{\text{segundo}} \times 0.357434 \frac{\text{bits}}{\text{pixeles}} = 22.23 \text{ Mbps} \quad \mathbf{Ec\ 2.35}$$

Tomando en cuenta otros datos experimentales, la Figura 2.66 muestra las tasas en las siguientes condiciones: n=8 vistas, 640x480, 320 cuadros presentados a 30 fps, GOP=16 y CABAC; por lo tanto aplicando el mismo análisis para n=8, la estimación del bitrate en SDTV a 40dB es:

$$640 \times 480 = 307200 \frac{\text{pixels}}{\text{cuadro}} \quad \mathbf{Ec\ 2.36}$$

$$307200 \frac{\text{pixels}}{\text{cuadro}} \times 30 \frac{\text{cuadros}}{\text{segundo}} = 9216000 \frac{\text{pixeles}}{\text{segundo}} \quad \mathbf{Ec\ 2.37}$$

$$850 \frac{\text{Kbits}}{\text{segundo}} \times \frac{1 \text{ segundo}}{9216000 \text{ pixeles}} = 0.09223 \frac{\text{bits}}{\text{pixeles}} \quad \mathbf{Ec\ 2.38}$$

$$10432800 \frac{\text{pixels}}{\text{segundo}} \times 0.09223 \frac{\text{bits}}{\text{pixels}} = 1 \text{ Mbps} \quad \text{Ec 2.39}$$

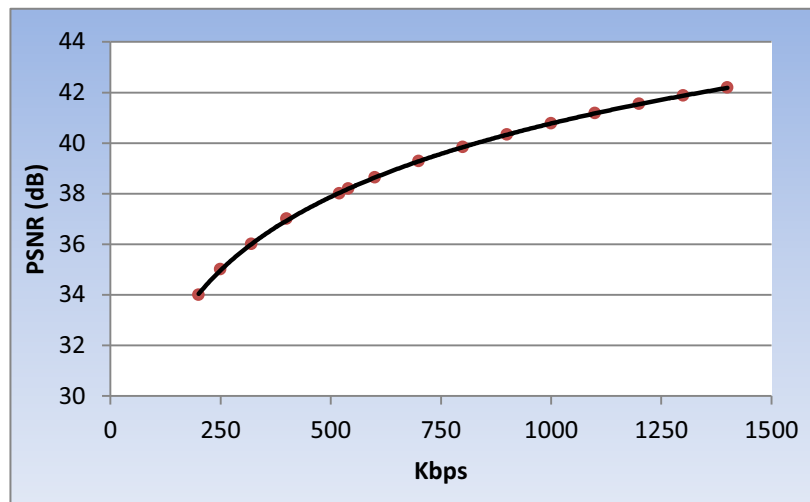


Figura 2. 66 Tasa de bits al codificar con MVC y n=8.

2.4.7.6 Tasa de bits para MVD

El análisis aquí se basa en que se obtendría una tasa superior a la del MVC en un 5% por vista de profundidad existente, sin embargo al considerar a la profundidad como una vista más en el codificador y contando con algoritmos y métodos de inter-correlación más eficientes la tasa de bits requerida no es excesiva. La Figura 2.67 muestra resultados experimentales considerando un video de 1024x768, un GOP=15, CABAC, 30 cuadros por segundo, con 100 cuadros presentados.

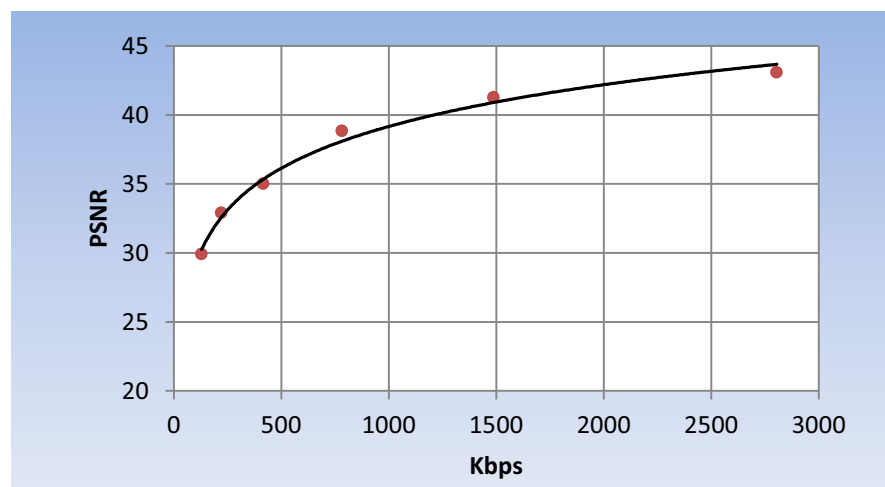


Figura 2. 67 Tasa de bits al codificar con MVD

2.4.7.7 Tasa de bits para LDV

LDV muestra un ahorro respecto a MVD debido a que la información adicional al video de color no requiere de grandes tasas de bits, y pueden no representar un exceso sobre las tasas hasta ahora estudiadas; LDV resulta una mejora y siempre las pruebas hacen referencia a la comparación de los formatos MVD y LDV. Resultados experimentales para LDV se presentan en la Figura 2.68. La tasa se enmarcara dentro del ISDB-T, por lo tanto a LDV se lo plantea codificar con H.264/MVC. Con los datos de la Figura 2.68, estimando el bitrate para un video de 1024x768 a 16.67 fps a 42 dB con la señal SDTV se tiene:

$$R = \text{Video color} + \text{Profundidad} + \text{Ocl. Color} + \text{Ocl. Profundidad}$$

$$R = 10432800 \frac{\text{pixels}}{\text{segundo}} \times (0.13 + 0.06 + 0.09 + 0.05) \frac{\text{bits}}{\text{pixel}} = 3.44 \text{ Mbps} \quad \mathbf{Ec 2.40}$$

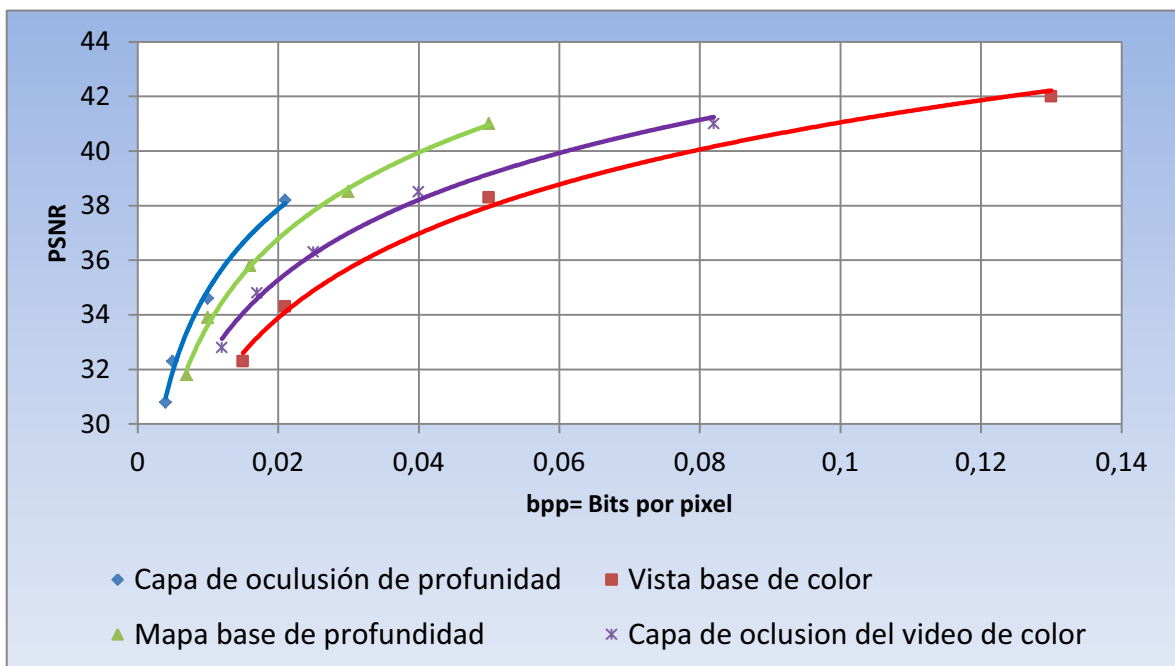


Figura 2. 68 Tasa de bits al codificar con LDV

2.4.7.8 Tasa de bits para DES

Dado que este formato se codifica mayormente con Platelet, no se realiza un estudio de sus tasas de bits, no obstante al codificarlo con MVC; el orden de sus tasas de bits es superior al de V+D en un 5 ó 10%.

2.5 TRANSMISIÓN DE LAS SEÑALES DE TV-3D EN EL ESTÁNDAR ISDB-T.

Con todo el análisis anterior, corresponde ahora estudiar el mecanismo por el cual se transmiten y se difunden las señales concernientes a una señal de televisión. Son ya conocidas las posibilidades existentes para crear efecto 3D, y en cada una las señales involucradas, por lo tanto es importante el análisis del proceso de transmisión para las señales involucradas. Dependiendo del formato y técnica elegida es necesario estudiar, cómo se ajustan las técnicas del estándar con las señales que se desean transmitir.

2.5.1 GENERACIÓN DEL TRANSPORT STREAM Y MULTIPLEXACIÓN.

De acuerdo al formato es claro que se pueden originar uno o varios flujos de señales digitales que deben configurar el Transport Stream, estos serán posteriormente codificados, modulados y transmitidos. ISDB-T establece un tipo de empaquetamiento basado en la norma MPEG-2, es decir ajustándose a la señalización, cabeceras y secuencias para audio y video.

De lo ya estudiado hay algunos formatos que originan una sola señal de video, que dentro de su concepción ya viene señalizada e identificada de acuerdo a la información que corresponda, en algunos formatos el codificador puede incluir un multiplexer, haciendo que si existen múltiples señales estas se codifiquen y se genere un único Bitstream. En otro caso cada señal eventualmente puede ser introducida al multiplexer final, todo esto dependerá de la disponibilidad de equipos en el mercado y del soporte para tal cantidad de entradas.

2.5.1.1 Empaquetamiento y generación del TS con MPEG-2

MPEG-2 establece algunas normas para generar los elementos a ser multiplexados posteriormente, tanto para audio como para video, MPEG-2 establece que el flujo de transporte TS va acompañado de su audio respectivo, por lo tanto en el caso de señales múltiples de video como es en la 3DTV se puede duplicar el audio o ahorrar ese espacio. En la Figura 2.69 se expone el empaquetamiento y las dimensiones del mismo

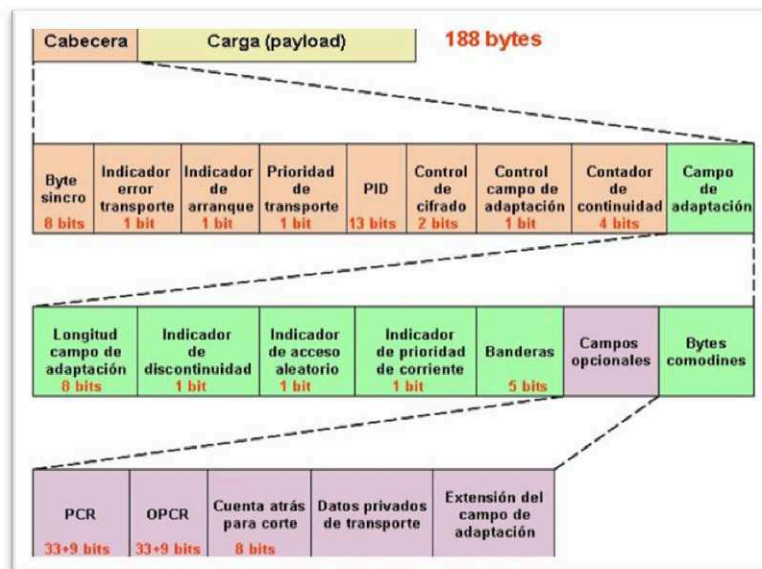


Figura 2. 69 Cabecera del Transport-Stream en MPEG-2.

Parte importante de la cabecera es el PID, (Packet Identification), los paquetes de TS pueden traer información de programas diferentes, además de datos para la reconstrucción de la información. Aparece este campo de 13 bits que permite la distinción de paquetes de diferentes Elementary Streams. De los 2^{13} valores posibles, hay 17 reservados para funciones especiales. Esto permite 8175 valores que son asignables a todos los otros ES que forman el TS. El multiplexor tiene que garantizar que cada ES tenga un único PID. La normativa MPEG no especifica qué valores de PID se tienen que dar a los ES (a excepción de los 17 mencionados). Con la cabecera incluida el Transport Stream se estructura como lo detalla la Figura 2.70.

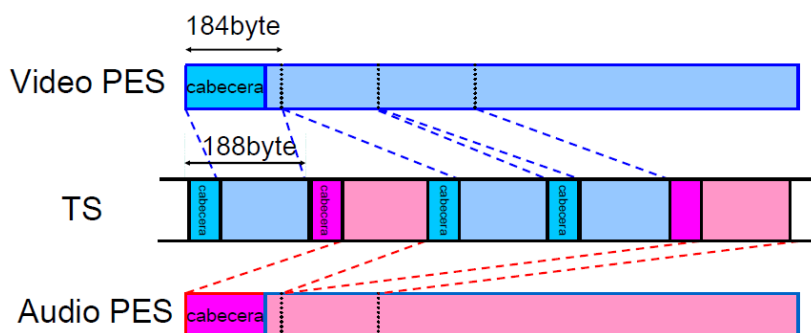


Figura 2. 70 Generación del Transport-Stream.

2.5.1.2 Multiplexación MPEG-2 para señales TV3D que contienen múltiples informaciones.

De lo ya estudiado está claro que en ciertos formatos, al utilizar más de un codificador se genera más de un flujo de de datos, y por ello estos deben ser obligatoriamente multiplexados si se quiere que se transmitan bajo la misma plataforma.

Existen otros formatos también que, si bien es cierto generan múltiples señales, también pueden contener la etapa de multiplexación en el propio codificador, ya que agregan allí su señalización y sincronía correspondiente, lo cual sería recuperado sin problema por el correspondiente decodificador. El formato que utiliza este tipo de multiplexación es el de Dual Streaming, considerando que los flujos pueden ser de señales L+R o de Video y profundidad. Con esto los formatos que después del codificador generan múltiples flujos son:

- Dual Streaming L+R
- MPEG-C parte 3
- MVC sin multiplexación en el codificador (típicamente para n=2).

En estos casos la multiplexación se estructura como lo detalla la Figura 2.71.

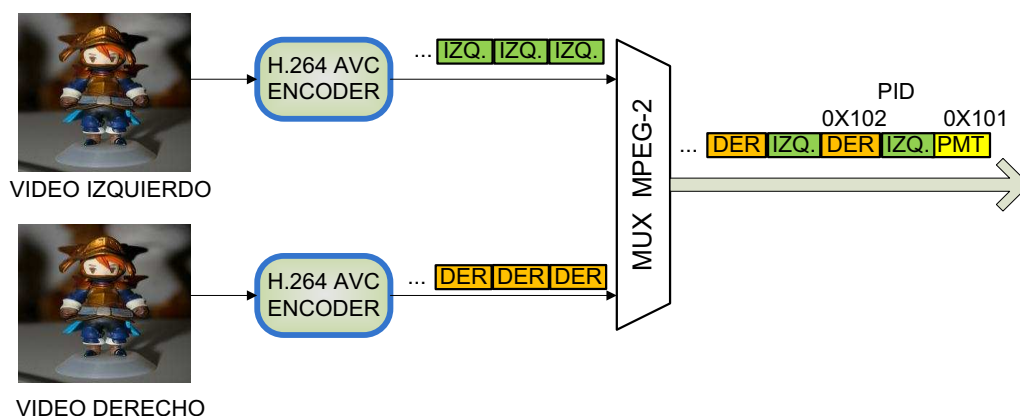


Figura 2. 71 Multiplexación de vistas en dual Streaming.

Se observa que izquierda y derecha se alternan en el MUX y se los provee de sus PID y PMT (Program Map Tables) para identificarlos y en recepción poder demultiplexarlos; también el audio se multiplexa después del video izquierdo o si

se requiere cada video puede ir acompañado del audio de forma idéntica al TS de MPEG-2 para televisión 2D.

Para el caso de MPEG-C parte 3 el esquema es idéntico y el Transport Stream se estructura de forma similar (ver Figura 2.72), la diferencia radica en que no se usa video derecho sino profundidad lo que hace que el codificador genere un cuadro de menos bitrate y el Mux alterne, entre cuadros L de mayor tamaño que los de profundidad, y los propios cuadros "depth".

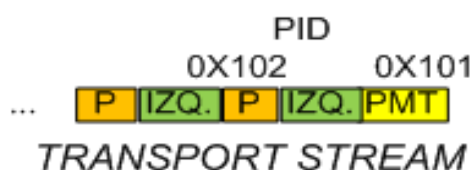


Figura 2. 72 Multiplexación de vistas izquierda + profundidad.

Para MVC el multiplexor generaliza su aplicación a n vistas, aunque lo más típico sean 2, obteniendo mejor compresión por la relación entrevista. Dado el panorama actual de la TDT, es recomendable que para señales 3D se ofrezca al multiplexor de audio, video y datos un solo flujo para no alterar la arquitectura actual. Actualmente la multiplexación se usa estableciendo grupos de entrada por canal de TV (1 video, 1 audio, datos). En nuestro caso, si la 3DTV requiere más de una señal de video para la entrada se tienen dos opciones: un Multiplexor previo para entregar un solo flujo (ver Figura 2.73), u ocupar un solo Multiplexor antes de la transmisión ocupando el espacio que ocuparían varios programas de TV (ver Figura 2.74).

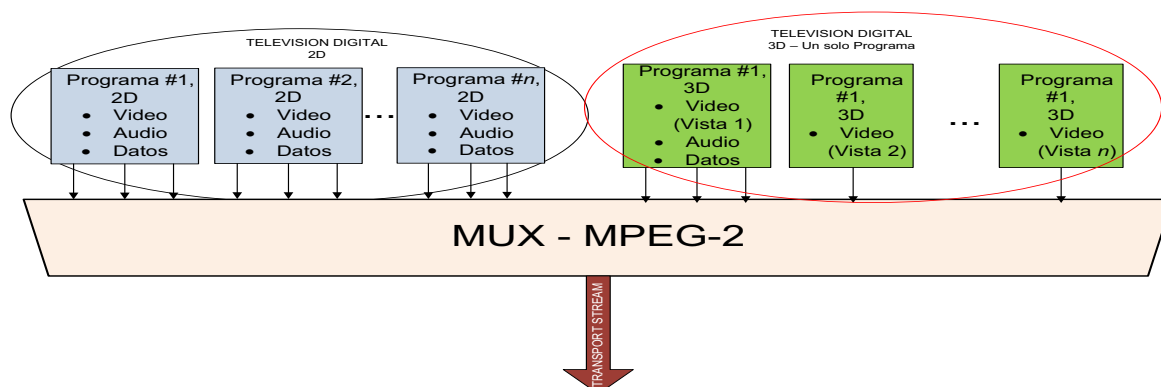


Figura 2. 73 Multiplexación de canales 2D y 3D con un solo Multiplexores.

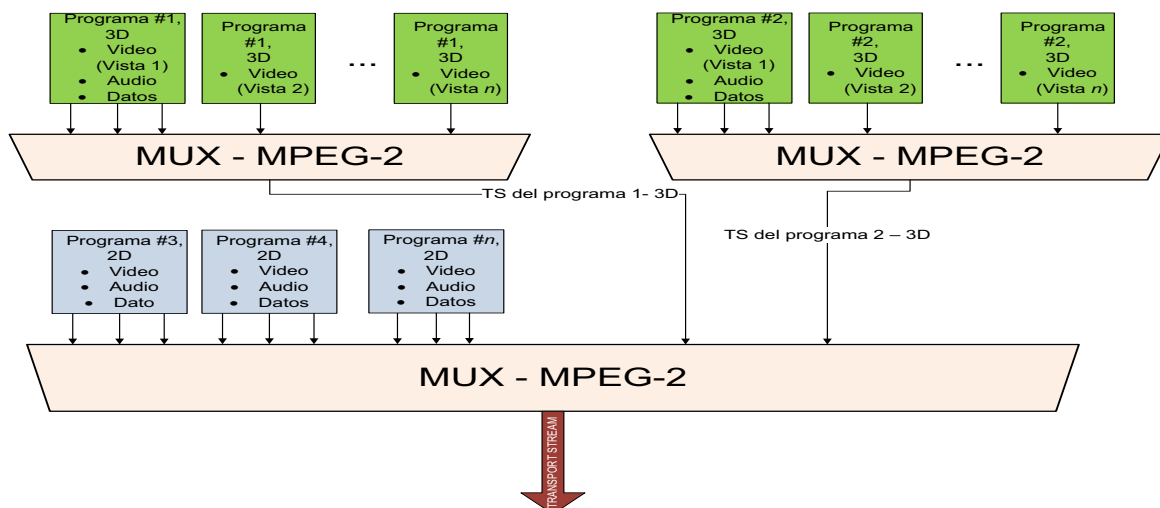


Figura 2. 74 Multiplexación de canales 2D y 3D con varios Multiplexores.

2.5.2 CODIFICACIÓN MODULACIÓN Y TRANSMISIÓN DE LA SEÑAL DE TV-3D.

Una vez generado el TS final, el estándar ISDB-T establece una serie de procedimientos hasta llegar a la modulación y difusión de la señal digital.

La secuencia de etapas después del TS son: Codificación de canal, modulación y etapa de radiofrecuencia o difusión. La codificación de canal (RS 204-188-8), y modulación con su respectivo FEC se realiza de forma idéntica a las señales monoscópicas, así como la generación de símbolos OFDM para su transmisión.

Está claro que el tipo de modulación, tasa de FEC, relación entre el tiempo útil y tiempo de guarda, al repercutir directamente con la tasa de bits que se soporta, deben ser analizados en que combinación se pueden cumplir los requerimientos que una señal TV3D exige. A continuación se analiza las características del ISDB-T que concuerdan con los requerimientos de la señal 3D para sus distintos formatos.

2.5.2.1 Características y capacidad de los 13 segmentos dentro de los 6 MHz.

Según lo analizado en el capítulo 1, por cada 6 MHz del espectro, la técnica OFDM permite ubicar 13 segmentos de los cuales 1 es exclusivo para one-seg y el resto pueden combinarse en algunas opciones de señales de diferente ancho

de banda. Cada segmento tiene un ancho de banda de 428.57 KHz y puede ser capaz de soportar diferentes velocidades de transmisión, tal y como lo especifica la Tabla 2.7.

TG/TU	FEC	QPSK-DQPSK		16-QAM		64-QAM	
1/4	1/2	3,651	0,281	7,302	0,562	10,95	0,843
	2/3	4,868	0,374	9,736	0,749	14,61	1,123
	3/4	5,477	0,421	10,95	0,843	16,43	1,264
	5/6	6,085	0,468	12,17	0,936	18,26	1,404
	7/8	6,39	0,492	12,78	0,983	19,17	1,475
1/8	1/2	4,057	0,312	8,114	0,624	12,17	0,936
	2/3	5,409	0,416	10,82	0,832	16,23	1,248
	3/4	6,085	0,468	12,17	0,936	18,26	1,404
	5/6	6,761	0,52	13,52	1,04	20,28	1,56
	7/8	7,099	0,546	14,2	1,092	21,3	1,638
1/16	1/2	4,295	0,33	8,591	0,025	12,89	0,991
	2/3	5,727	0,441	11,46	0,034	17,18	1,322
	3/4	6,443	0,496	12,89	0,038	19,33	1,487
	5/6	7,159	0,551	14,32	0,042	21,45	1,65
	7/8	7,517	0,578	15,03	0,044	22,55	1,735
1/32	1/2	4,426	0,34	8,851	0,026	13,28	1,021
	2/3	5,901	0,454	11,8	0,035	17,7	1,362
	3/4	6,638	0,511	13,28	0,039	19,92	1,532
	5/6	7,376	0,567	14,75	0,044	22,13	1,702
	7/8	7,745	0,596	15,49	0,046	23,24	1,787
	Velocidad de cada segmento en Mbps						
	Velocidad de los 13 segmentos en Mbps						

Tabla 2. 7 Tasas de transmisión en ISDB-T.

Es necesario señalar que los 13 segmentos no se constituyen puramente de señales de video sino que se compone de audio, el one-seg, datos (1.5 Mbps), etc.

De acuerdo a esta información y para señales de típicas canales, como los son SDTV de 2 a 6 Mbps y de HDTV 8 a 18 Mbps, y analizando la combinación entre FEC y modulación utilizada para obtener mejores tasas de bits, las pruebas realizadas en los países donde se establecerá este estándar han definido entre otras las siguientes opciones:

- 1 Programa HD de gran calidad en 12 seg. (18 Mbps o más) + one seg + audio + datos.
- 1 Programa HD de 8 Segmentos + 1 SD de 4 Segmentos + one seg+datos.
- 3 SD de 4 segmentos c/u + one seg + audio + datos.
- 2 HDTV de 6 segmentos c/u + one seg + audio +datos
- 6 SD de 2 segmentos c/u + one seg + audio + datos.
- 1 HDTV + 3SDTV + one seg.

Como se puede observar, todo depende de la tasa que ocupen las señales de TV, es decir, de qué calidad se requiere, para según eso acomodar las técnicas y obtener valores coherentes según la tabla 2.7; y con ello lograr utilizar los 6 MHz de forma eficiente.

2.5.2.2 Señales de TV-3D en los segmentos OFDM.

De acuerdo a las características antes expuestas compete ahora analizar la relación entre las mismas y las señales de TV3D, evidentemente esta información no puede ser tratada simplemente como una SDTV, HDTV, incluso One-Seg. Entonces corresponde analizar los potenciales Bitrate que cada formato requiere y así poder estimar el valor de los segmentos a ser utilizados.

En el caso de One-Seg, también se puede analizar qué formato permite adaptarse a un segmento en condiciones tales que se pueda transmitir a dispositivos móviles.

En la Tabla 2.8 se muestra un resumen de los bitrate de los diferentes formatos de TV-3D.

FORMATO		BIT-RATE 2D	BIT-RATE 3D	%	PSNR
OCA		SD: 3.5 Mbps	SD: 4 Mbps	14.28	36
		HD: 7 Mbps	HD: 8 Mbps	14.28	42
CSV		SD: 4 Mbps	SD: 4 Mbps x 2 = 8 Mbps	100	36
		HD: 8 Mbps	HD: 8 Mbps x 2 = 16 Mbps	100	42
CFC+		HD: 8 Mbps	HD: 8 Mbps	0	42
V+D.	MPEG-C PARTE 3.	SD: 3.5 Mbps	SD: 3.625 Mbps	10	42
		HD: 12 Mbps	HD: 12.960 Mbps	10	40
	AUX VIDEO H.264	SD: 4 Mbps	SD: 4.28 Mbps	7	40
		HD: 8 Mbps	HD: 8.56 Mbps	7	40
MVC		SD: 3 Mbps	SD: 3.729 Mbps	8	40
MVD		SD: 4 Mbps	SD: 4.635 Mbps	15	40
LDV		SD: 3 Mbps	SD: 3.44 Mbps	8.7	42
DES		SD: 4 Mbps	SD: 4.52 Mbps	8.8	42

Tabla 2. 8 Tasas de transmisión en ISDB-T para los formatos de TV3D.

Es decir, las señales 3D con formatos de calidad aumentan la tasa de un canal TV-2D entre un 10 y un 15%, por otro lado los formatos más compatibles y más fáciles de implementar pueden llegar a incrementar en un 100% los requerimientos. Con todo esto, si queremos introducir en los 13 segmentos una señal de TV-3D dependerá del formato, de la calidad y del video mismo. Por ejemplo usando 64 QAM con FEC de 3/4, que es lo que típicamente se utiliza, tendremos segmentos capaces de soportar 1.264 Mbps, con lo cual podremos transportar en 8 segmentos una señal TV3D y un canal SD extra. Así se puede variar los parámetros y, dependiendo del formato, adecuar a los segmentos la señal de video 3D.

A continuación se presentan algunas posibilidades para algunos formatos de TV 3D.

En análogo cromático optimizado:

- 1 TV3D – HD de 12 segmentos + one seg + audio + datos.

- 1 TV3D – HD de 8 Segmentos + 1 SD de 4 Segmentos + one seg+datos.
- 3 TV3D-SD de 4 segmentos c/u + one seg + audio + datos.
- 2 TV3D-HD de 6 segmentos c/u + one seg + audio +datos

En estéreo convencional (Simulcast o Dual Streaming)

- 1 TV3D – HD DE 12 segmentos + one seg + audio + datos.
- 1 TV3D – HD de 8 Segmentos + 1 SD de 4 Segmentos + one seg+datos.
- 2 TV3D-SD de 4 segmentos c/u + 1 HDTV + one seg + audio + datos.

AL estar CFC+ empaquetado en un cuadro HD se tienen las mismas posibilidades para señales de alta definición es decir 2 en los 13 segmentos, o 1 más 3 de definición estándar monoscópica.

V+D es de los formatos más adaptables, por superar en un 10% las tasas de las señales monoscópicas, por lo tanto, se tiene prácticamente la misma versatilidad que para señales 2D.

En cuanto a los llamados formatos avanzados, la versatilidad se mantiene, es decir aún en los segmentos se puede ubicar además de la señal 3D otros programas SD, lo que se debe tomar en cuenta es que en estos formatos los requerimientos aumentan de manera significativa si se aumenta el número de vistas; en todo caso se estima que una señal de SDTV en 3D ocupe 4 ó 5 segmentos, es decir 1 ó 2 más de los que se ocuparían si la señal fuere monoscópica.

2.5.2.3 Señales de TV-3D en el segmento destinado One-seg

Debido a que el one-seg es fijo para la norma, y que su modulación para alcanzar la cobertura de dispositivos móviles es de QPSK y FEC 2/3, las tasas que este servicio soporta están en el orden de los 400 Kbps, por lo que analizando los formatos se observa que CSV es inaplicable. De hecho los datos presentados en 2.4.7 son en algunos casos datos experimentales en señales LDVT, es decir que sin problema se ajustan a las tasas del estándar. Si bien es cierto que a mayor calidad se aumenta el bitrate, cabe señalar que las aplicaciones portátiles exigen

un valor menor de PSNR, por lo cual se tiene menor exigencia de tasa de transmisión, haciéndolo perfectamente aplicable la TV-3D móvil.

2.5.2.4 Transmisión de la señal de video.

Esta etapa es la parte final del proceso y ya llegado a este punto, es importante señalar que el análisis previo indicó que las señales de televisión estereoscópica tienen valores técnicamente compatibles con las etapas de codificación multiplexación, empaquetamiento, codificación de canal, etc. Dado que están dentro del estándar, la etapa final de radio frecuencia, generación de símbolos OFDM, y el modulador se atraviesa sin señalización especial.

El único caso en el cual la estructura de transmisión varía es para Simulcast, que como ya se explicó, requiere de una doble estructura para difundir las señales, es decir dos espacios diferentes e independientes en el canal de transmisión. Aún así, como dos programas se pueden transmitir en grupos de segmentos adyacentes, el caso de Simulcast puede entrar en el caso de que una estación utilice y transmita el doble de segmentos de lo que haría normalmente.

Debido a que el uso de los 6 MHz y de los segmentos no se encuentra aún regulado, es difícil establecer estrictamente el uso de los segmentos, y si las operadoras tendrán completamente el ancho de banda, o en su lugar una parte del mismo de acuerdo a la programación y la información que tengan disponible para transmitir.

2.6 EFECTO DEL SERVICIO 3DTV SOBRE LAS CARACTERÍSTICAS TÉCNICAS DE ISDB-T.

Los temas abordados en las secciones anteriores, permiten esclarecer e identificar qué parámetros del estándar ISDB-T se ven afectados directamente. De igual forma se puede detallar la forma en que el servicio interviene en las características del estándar y por lo tanto determinar la aplicabilidad del mismo; a continuación la Tabla 2.9, presenta el resumen de características para cada formato de producción 3D, detallando el efecto que tienen sobre los parámetros de ISDB-T internacional.

Para el servicio 3D									
Parámetro de ISDB-T internacional	Especificación para señales tradicionales (2D)	OCA	CSV	CFC+	V+D (MPEG-3 parte 3)	V+D (Aux Video)	MVC	MVD, LDV, DES	
Codificación - Compresión del video	H.264 - MPEG -4, se utiliza un codificador por señal de video 2D	Se procede como una señal 2D en SD ó HD	Requiere dos codificadores, uno por vista (L+R)	Un solo codificador, con cuadros HD, se requiere SVC	Requiere dos codificadores uno para V y uno para D	Un solo codificador con al menos la extensión AVC	Un solo codificador con MVC, el multistream es ineficiente	Un solo codificador con MVC, el multistream es ineficiente	
Generación del TS (Transport Stream)	Mpeg-2 combinando audio datos y video	Se procede como una señal 2D en SD ó HD	Se realiza una Mux. para intercalar L y R (Dual - Streaming)	Se procede como una señal 2D en HD	Se realiza una Mux. para intercalar el video y la profundidad	Se procede como una señal 2D en SD ó HD	Se procede como una señal 2D en SD ó HD	Un solo codificador con MVC, el multistream es ineficiente	
Multiplexación del Video	Mpeg-2 para combinar la señal de varios programas / canales	Se procede como una señal 2D en SD ó HD	Se procede como una señal 2D en SD ó HD	Se procede como una señal 2D en HD	Se procede como una señal 2D en SD ó HD	Se procede como una señal 2D en SD ó HD	Se procede como una señal 2D en HD	Un solo codificador con MVC, el multistream es ineficiente	
Bit-rate del video*	Variable dependiendo del tipo de modulación y del FEC (entre 0.28 y 1.78 Mbps por segmento)	Puede exigir tasas ligeramente mayores (un 10% mas)	Se duplica respecto a una señal 2D. (En SD 8 Mbps, 200% de aumento)	Similar a una señal HD, aumenta para UHD	Aumenta entre un 10 y 20 % a un canal HD o SD	Aumenta entre un 10 y 20 % a un canal HD o SD	Aumenta un 8% para n=2 vistas, para n superior se pueden llegar a mas de 20 Mbps	Un solo codificador con MVC, el multistream es ineficiente	
Uso de los segmentos (6 MHz)	13 segmentos que pueden usarse de acuerdo al tipo de video	Similar a canales SD ó HD	Se duplica la demanda, espacio para la mitad de canales	Se considera señales HD ó UHD	La disponibilidad se reduce muy poco	La disponibilidad se reduce muy poco	Se reduce la disponibilidad considerablemente para n elevado.	Un solo codificador con MVC, el multistream es ineficiente	
Etapas de RF (Modulación y transmisión)	Modulación: QPSK, 16-QAM, 64-QAM; se transmite mediante OFDM	Se procede como una señal 2D en SD ó HD	Para simulcast requiere dos infraestructuras independientes y sincronizadas	Se procede como una señal 2D en HD	Se procede como una señal 2D	Se procede como una señal 2D.	Se procede como una señal 2D de gran capacidad	Un solo codificador con MVC, el multistream es ineficiente	
Movilidad	One-seg exclusivo para el servicio	Se procede como una señal 2D para LDTV	No aplicable, bitrate limitado	No aplicable, bitrate limitado	Ideal por el bajo incremento de bitrate	Ideal por el bajo incremento de bitrate	No aplicable, bitrate limitado	Un solo codificador con MVC, el multistream es ineficiente	
Capacidad del Set-top-Box	En recepción de acuerdo al estándar, decodifica señales SD y HD	El Set-Top-Box no requiere funcionalidades especiales	Debe "mezclar" las vistas ó debe sincronizar dos transmisiones simultáneas	Actualizado para separar y procesar eficientemente los cuadros	Actualizado para separar V+D y procesar con DIBR	Actualizado para separar V+D y procesar con DIBR	Actualizado a MVC	Un solo codificador con MVC, el multistream es ineficiente	

* La tasa de bits es variable de acuerdo a la calidad (PSNR), la secuencia de video y la tasa fps, los valores pueden variar alrededor de los promedios

■ Formatos teóricamente ineficientes (generación cero), fáciles de aplicar sin cambios en infraestructura ni en equipos.
■ Formato de 1 Generación, alta calidad, bajo costo poco cambio de infir. Representa el primer paso en el servicio.
■ Formatos de alta calidad, ideal para ahorro de bitrate y en Rx móvil, requiere mayor costo y cambio de equipos

■ Formato de gran calidad, respaldado por el estándar en el futuro. Actualización total
■ Formatos avanzados, permiten FVV y holografía, calidad máxima, alto costo, transmisión larga.

Tabla 2. 9 Características del estándar ISDB-T para el servicio 3DTV

CAPÍTULO 3

RECEPCIÓN DE LA SEÑAL DE VIDEO 3D PARA EL ESTÁNDAR ISDB-T.

4.1 DESCRIPCIÓN GENERAL DEL SISTEMA DE RECEPCIÓN PARA VIDEO ESTEREOSCÓPICO.

Del capítulo 2 se desprenden algunas consideraciones a ser tomadas en cuenta para el proceso de recepción y reproducción del video 3D, entre ellas se encuentran:

- ¿Qué cantidad de información se transmitió?
- ¿Qué tipo de formato se utilizó y por ende qué señalizadores existen?
- El flujo de datos enviados, ¿es único o se trata de un flujo múltiple?
- ¿Bajo qué estándar se transmitió?
- ¿El Set-top Box (STB) o el IRD (Integrated Receiver Decoder) tiene las versiones actualizadas de los estándares con los que se trabajó la señal de origen?
- Considerando el Middleware GINGA, ¿es posible generar aplicaciones interactivas útiles para el servicio de televisión estereoscópica?
- Las pantallas en recepción, ¿pueden reproducir contenidos estereoscópicos?

Estas y algunas otras interrogantes deben ser explicadas, de tal manera que se pueda establecer un esquema claro y técnicamente posible para los sistemas de recepción de la TV3D en el estándar ISDB-T.

El servicio de televisión digital ha sido diseñado para que el destinatario final sea el usuario en casa, es decir no se puede tener una infraestructura de recepción complicada que haga inviable que la TV3D pueda llegar a los hogares. Con lo expuesto en el capítulo anterior, las técnicas de transmisión estudiadas deberían permitir una forma relativamente sencilla de recibir el video 3D.

La etapa de recepción para televisión estereoscópica, al igual que en todos los servicios de telecomunicaciones, consiste en aplicar los procesos contrarios a los aplicados en transmisión en el orden inverso correspondiente. Dado que todos estos procesos los hace el Set-top Box o el sintonizador del televisor, se procura que no sean procesamientos complejos que hagan que el equipo receptor sea muy costoso y muy difícil de implementar en la realidad.

La televisión digital terrestre para su recepción requiere únicamente de la instalación de un decodificador en el domicilio para su recepción. Este decodificador (IRD), que resulta necesario para los televisores convencionales, puede desaparecer como elemento físico separado si se integra al televisor, como es el caso de la totalidad de los televisores de pantalla plana de nueva generación que se están comercializando hoy en día. Con esto, la recepción de la TDT en un domicilio puede constituirse como muestra la Figura 3.1

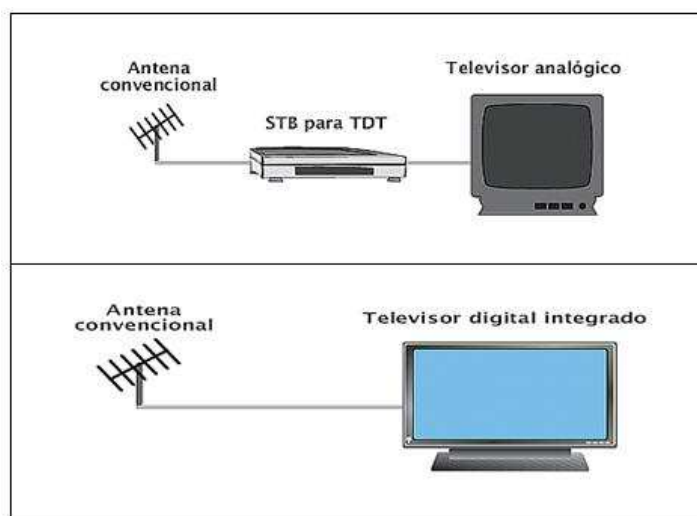


Figura 3. 1 Infraestructura de recepción de la Televisión Digital Terrestre.

En definitiva el STB o el receptor integrado al televisor, tiene la tarea de recibir y decodificar la señal proveniente de la estación tele-difusora, y presentarla en la pantalla. Es necesario analizar ahora la capacidad de este IRD con respecto a señales que representen video estereoscópico; al referirse a la tabla 2.2, notamos que para algunos casos particulares es necesaria la actualización del IRD, es decir tener un receptor con funcionalidades extras y con las últimas modificaciones que los estándares de codificación hayan presentado. Y por otro

lado, casos en los cuales el formato no establece que el Set-Top Box sufra cambios significativos.

6.1.1 DESCRIPCIÓN DEL IRD (Integrated Receiver-Decoder)

Un IRD o Set-Top-Box incluye un demodulador QPSK/QAM, un decodificador de Viterbi, corrección de errores RS, demultiplexor para las distintas señales, decodificador H.264 y conversores digital – analógico. Cada fabricante establece su mecanismo y su infraestructura de recepción en el IRD, sin embargo se mantiene una tendencia en lo que tiene que ver a los bloques funcionales de los Set-Top-Box. La Figura 3.2 presenta el diagrama de bloques de un Set-Top-Box convencional.

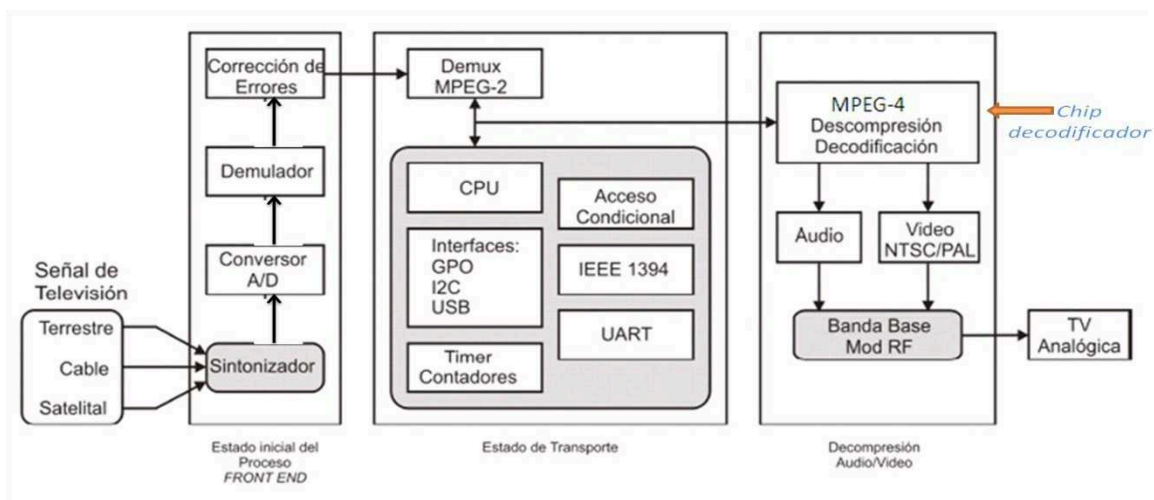


Figura 3. 2 Estructura Interna de un Set-Top-Box.⁴⁷

Son claramente identificables los bloques procesadores del video y de la conversión hacia la pantalla, una parte fundamental del IRD es el chip decoder, o chip decodificador el cual es actualizable y según el gráfico de la Figura 3.2 cumple con la norma H.264.

Investigaciones recientes y los fabricantes, hoy en día están en la labor de producir estos chips de procesamiento de video de manera que los Set-top-Box

⁴⁷ Villanueva J., Velásquez C., "Informe Preliminar: Estado del Arte de Receptores Set-Top-Box – Aplicaciones", Área de Aplicaciones Telemáticas-Universidad Nacional de Ingeniería, Lima-Perú Febrero 2010.

estén totalmente actualizados a la última versión de los estándares; por ejemplo, dado que MVC es la última actualización de H.264, se pretende que los nuevos chips decoders contengan esta extensión y hagan del Set-Top-Box un aparato más funcional y tecnológicamente más desarrollado.

6.2 PROCESO DE RECEPCIÓN DEL SERVICIO DE TELEVISIÓN ESTEREOSCÓPICA PARA ISDB-T

Con lo estudiado en el capítulo 2 ahora corresponde analizar los procesos propios de cada formato para que se represente la escena estereoscópica que en el lado transmisor se originó.

6.2.1 RECEPCIÓN DE LA SEÑAL ANÁGLIFA CROMÁTICA

Al tener la señal cromática con características similares a las del video 2D, la recepción y representación de la imagen es idéntica a la ya conocida y tradicional, lo único que cabe señalar es que lo óptimo para la representación es reproducir el video en alta definición. En este caso la descomposición de las dos vistas no se las hace en el IRD si no de manera externa mediante los lentes filtros cromáticos.

6.2.2 RECEPCIÓN DE LA SEÑAL CSV

La señal en recepción contiene información de 2 vistas, el trabajo del IRD es separarlas y enviar dichas señales a la pantalla, la cual puede mostrar el video mediante cualquier técnica conocida.

6.2.2.1 Recepción y desempquetamiento del Dual-Stream

En este caso las dos vistas vienen multiplexadas en tiempo y señalizadas como flujos que representan las 2 vistas de una escena, por lo tanto el IRD debe configurarse de tal forma que en la etapa demultiplexora identifique el señalizador y envíe las 2 imágenes hacia la pantalla la cual tomara dichas imágenes y las representará bajo cualquier técnica conocida. Es necesario señalar que de existir la comunicación entre el Set-Top-Box y la pantalla, ésta debe hacerse mediante un estándar soportado por ambos dispositivos.

6.2.2.2 Recepción y desempaquetamiento del Simulcast

El problema es que el usar un doble decoder implicaría una plataforma de recepción muy compleja, por lo tanto el IRD en este caso debe ser configurado de tal forma, que identifique que la señal sintonizada de los dos canales de TV pertenece a la información L+R y la estructure bajo un formato que la pantalla acepte para su reproducción.

6.2.3 RECEPCIÓN DE LA SEÑAL PARA CFC+

Este es uno de los más sencillos procedimientos en recuperar la señal, debido a que está soportado por el estándar HDMI, entonces la recepción y posterior reproducción en la pantalla debe enmarcarse en identificar los señalizadores que HDMI establece y considerando que tanto el IRD como la pantalla tienen este esquema vigente, la recepción requiere únicamente que el Set-Top-Box contenga actualizado y soporte a HDMI.

6.2.4 RECEPCIÓN DE LA SEÑAL PARA V+D

6.2.4.1 Recepción para MPEG-C Parte 3

Aquí se observa un comportamiento similar al estudiado en 3.2.2, solamente cambia la información a transmitirse, es importante mencionar que en este caso el Set-Top-Box, receptor (ver Figura 3.3) o el televisor debe ser capaz de recibir y sintetizar las vistas a partir de la información V+D, y en caso de no ser capaz, identificar solo la información 2D y descartar el mapa de profundidad. Esta representación de una o más vistas que se puedan generar a partir de V+D debe ser realizada en el receptor en tiempo real, la técnica que lo logra es conocida como DIBR (Depth-Image-Based Rendering).

DIBR es el proceso de sintetizar vistas virtuales a partir de una escena con o sin movimiento y su información de profundidad asociada.

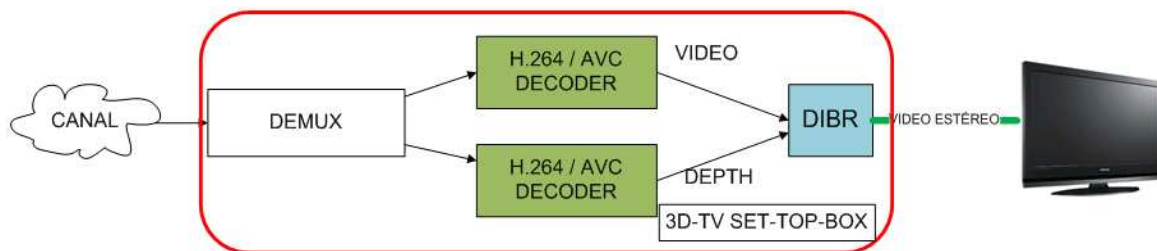


Figura 3. 3 Receptor IRD MPEG-C Parte 3.

6.2.4.2 Recepción para Video auxiliar

En este caso el receptor debe estar actualizado para entender que el video usa H.264 y también utiliza una señal de video auxiliar, estas dos señales (V+D) deben procesarse (DIBR) y transmitirse para que la pantalla pueda representarlos bajo cualquier técnica conocida.

6.2.5 RECEPCIÓN DE LA SEÑAL PARA FORMATOS AVANZADOS (MVC, MVD, LDV, DES)

Considerando que estos formatos utilizan el perfil multivista de H.264 el IRD debe estar normalizado a este estándar y ser capaz de transmitir a la pantalla flujos de datos que contengan las posibles capas y múltiples vistas que el formato exige. El uso de múltiples vistas y múltiples informaciones provoca que se aplique la técnica DIBR antes descrita.

6.3 DESCRIPCIÓN DEL SET-TOP-BOX PARA 3DTV.

A pesar de que para algunos formatos, los set top box existentes son aplicables, es necesario mencionar que el avance de los formatos y la calidad y cantidad de información involucrada, provocan una tendencia a diseñar nuevos chips decodificadores, set top box con más funcionalidades, y en el caso de que el sintonizador esté incluido en el televisor, este debe contar con todas las prestaciones para una recepción 3D. Hoy en día se pretende que el desarrollo de la tecnología apunte a crear IRD's que soporten muchos formatos de producción 3D, así mismo que los chips para Set Top Box puedan ser capaces de procesar al menos varios de los formatos compatibles con el cuadro HD.

La 3DTV al brindar más opciones de interactividad, crea la necesidad de que se desarrollen aplicaciones (GINGA) que permitan manejar este tipo de contenidos y permitirle al usuario escoger la información que desea ver, elegir el formato, y elegir de que manera desea reproducir el video.

También es importante señalar que todo el proceso del IRD debe ser transmitido a la pantalla por lo que la interface utilizada debe soportar los datos y formatos utilizados. Los Set-Top-Box también deben evolucionar con la tecnología 3D usada, es así que se pueden identificar las siguientes etapas:

- Formatos que funcionen sobre los Set-Top-Box actuales reprogramados.
- Formatos que funcionen sobre Set-Top-Box actualizados (chip decoder) y reprogramados para 3DTV.

En el primer caso podemos citar aquellos formatos como OCA o Simulcast, sabiendo que en este último el IRD debe ser capaz de identificar y enviar a la pantalla dos flujos de video.

En el segundo caso ya se incluyen herramientas como DIBR, varios procesadores para la posible existencia de múltiples vistas, chips que soporten algunos tipos de empaquetamiento de cuadro, etc.

Hoy en día los chips que se ofertan en el mercado son capaces de decodificar algunos formatos de 3DTV como empaquetamiento de cuadros, V+D con DIBR, y video con varias vistas, sin embargo aún no se comercializan chips desarrollados para soportar todos los formatos, estos chips ya se encuentran diseñados y en proceso de investigación. Un Set-Top-Box para 3DTV debe cumplir algunos requerimientos debido al gran procesamiento necesario⁴⁸:

- Capacidad de alto procesamiento para decodificar la vista real y sintetizar las vistas virtuales.

⁴⁸ Pei-KueiTsung ; Ping-ChihLin ; Kuan-YuChen ;Tzu-Der Chuang ; Hsin-Jung Yang ; Shao-YiChien ;Li-Fu Ding ; Wei-Yin Chen ; Chih-Chi Cheng ; Tung-ChienChen ; Liang-GeeChen,NationalTaiwanUniversity, Taipei."A 216fps 4096×2160p 3DTV Set-Top Box SoC for Free-Viewpoint 3DTV Applications", Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2011 IEEE International, p124-126

- Para soportar FVV (Free Viewpoint View) rotación 3D y traslación 3D (6D), la matriz de pandeo para reconstruir las vistas debe existir para cada pixel.
- La memoria del sistema debe funcionar a una tasa de 31.5 Gbps.

Dichos requerimientos no se tienen disponibles hoy en día en el mercado, sin embargo ya están siendo investigados. La Figura 3.4 presenta el proceso de sintetización de vistas en un Set-Top-Box para 3DTV.

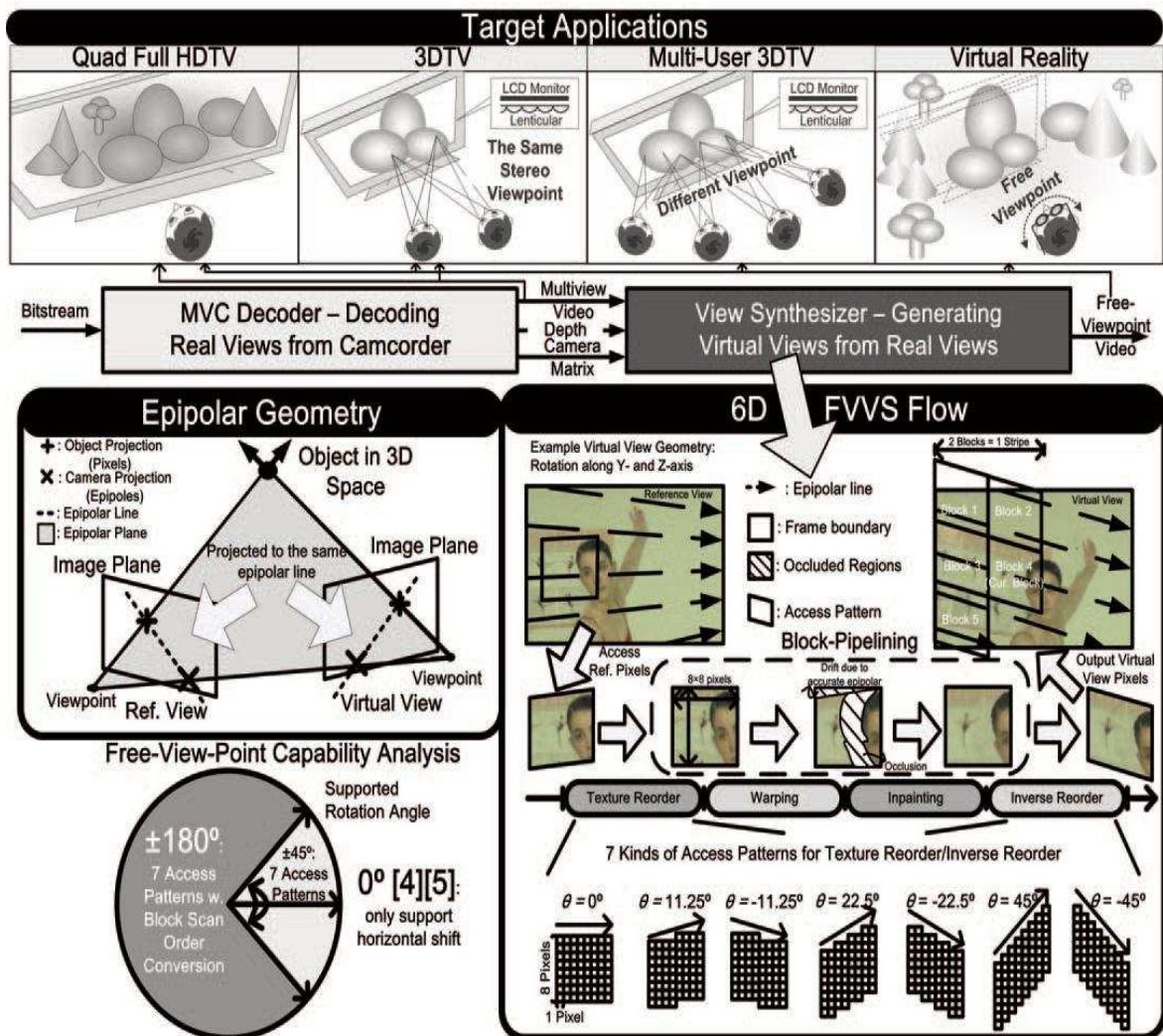


Figura 3. 4 Sintetización de vistas virtuales en un Set-Top-Box 3D.⁴⁹

⁴⁹ Pei-KueiTsung ; Ping-Chih Lin ; Kuan-Yu Chen ;Tzu-Der Chuang ; Hsin-Jung Yang ; Shao-Yi Chien ;Li-Fu Ding ; Wei-Yin Chen ; Chih-Chi Cheng ; Tung-Chien Chen ; Liang-Gee Chen, National Taiwan University, Taipei. "A 216fps 4096x2160p 3DTV Set-Top Box SoC for Free-Viewpoint 3DTV Applications", Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2011 IEEE International, p124-126

La arquitectura del IRD se muestra en la Figura 3.5.

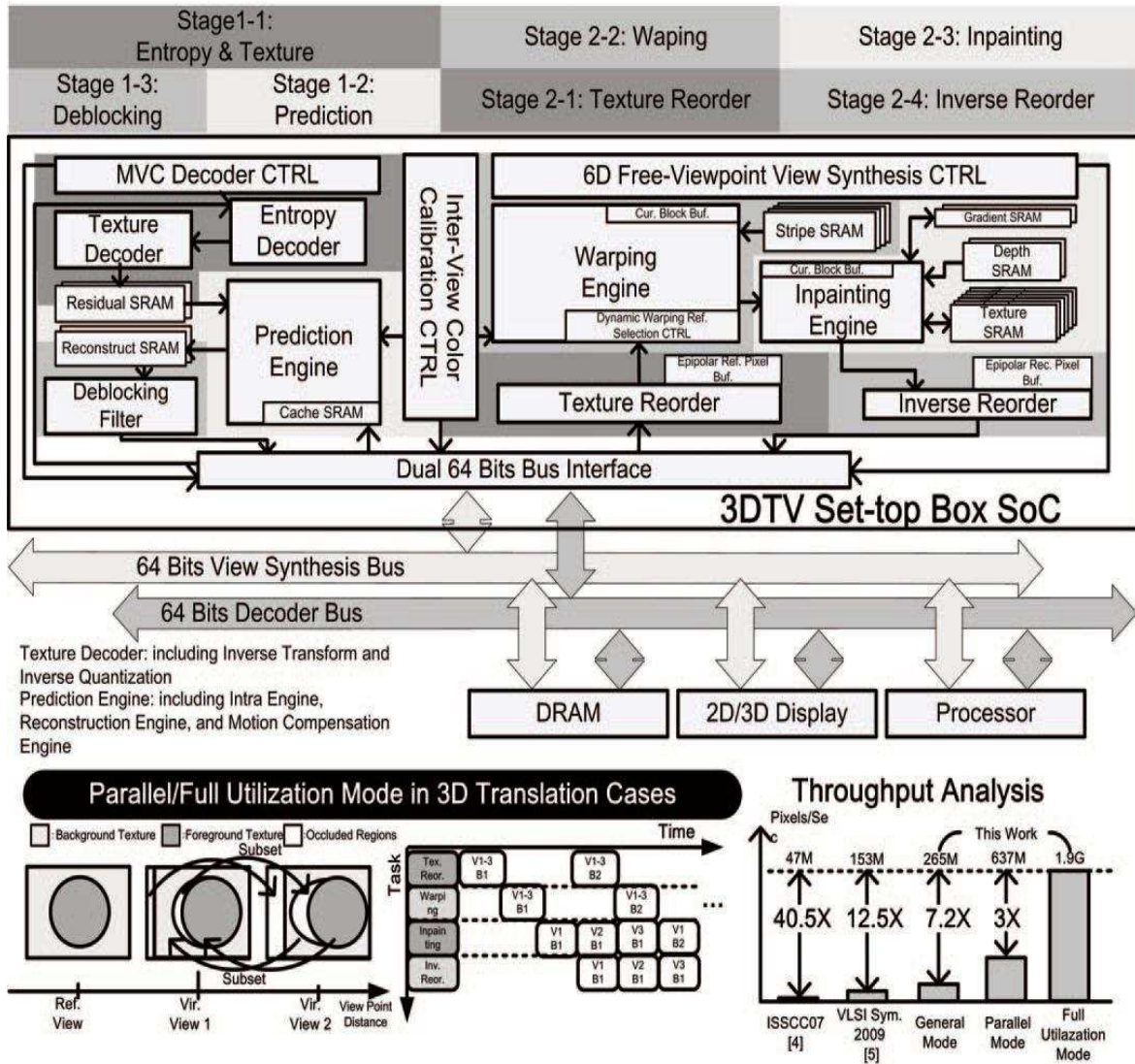


Figura 3. 5 Estructura de un Set-Top-Box 3D.⁵⁰

6.3.1 COMUNICACIÓN E INTERFACES ENTRE EL SET-TOP-BOX Y EL TELEVISOR.

En el caso en el que se utilice un Set-Top-Box es importante considerar que este debe enviar información que la pantalla pueda interpretar y reproducir, por lo cual ambos deben soportar el estándar H.264 en la versión que el formato en uso

⁵⁰ Pei-KueiTsung ; Ping-Chih Lin ; Kuan-Yu Chen ;Tzu-Der Chuang ; Hsin-Jung Yang ; Shao-Yi Chien ;Li-Fu Ding ; Wei-Yin Chen ; Chih-Chi Cheng ; Tung-Chien Chen ; Liang-Gee Chen, National Taiwan University, Taipei. "A 216fps 4096×2160p 3DTV Set-Top Box SoC for Free-Viewpoint 3DTV Applications", Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2011 IEEE International, p124-126

exija; además de ello se plantea que la comunicación involucre gran cantidad de información. Posterior al procesamiento del IRD, se debe considerar bajo qué tipo de información 3D trabaja la pantalla y que puertos de entrada posee para poder utilizar un protocolo de comunicación y una interfaz correcta.

Para el formato compatible con el cuadro HD el estándar aceptado es HDMI, por lo tanto se utiliza el interface del mismo nombre, es posible utilizar para otros formatos otras interfaces típicas en un televisor 3D (BNC, RJ-45, HDMI, USB, IEEE 1394, RS 232). A pesar de los ya existentes, los fabricantes de pantallas poseen interfaces propietarios en caso de que exista algún formato que necesite ser adaptado.

Dado que el 3D se puede lograr con 2 o más vistas, el interface y la comunicación varía según el caso. Para 2 vistas, HDMI 1.4 es el adecuado debido a que sabe cómo manejar el caso de distribuir y transmitir las vistas en varias posibilidades (ver Figura 3.6).



Figura 3. 6 Interfaz para formatos de $n=2$ vistas.

En el caso de $n>2$, las aplicaciones apuntan hacia autoestereoscopia y FVV, en ese caso aún no se comercializan displays capaces de transmitirlo y por lo tanto la interfaz adecuada no está determinada (ver Figura 3.7).

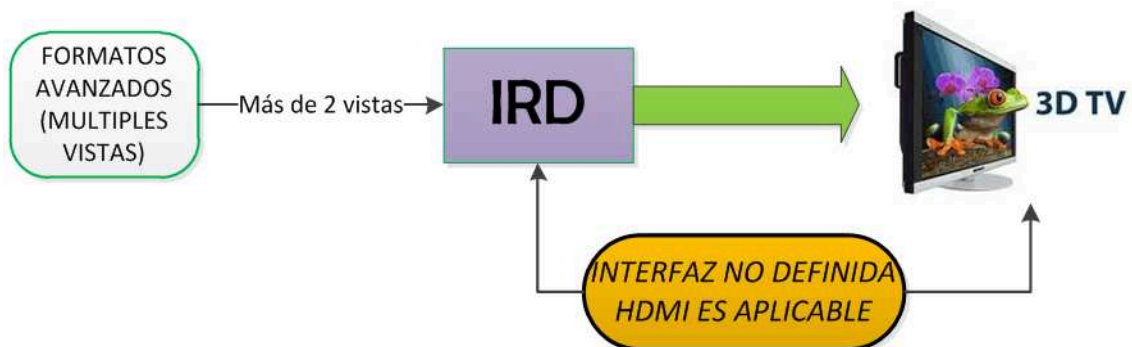


Figura 3. 7 Interfaz para formatos de $n>2$ vistas.

HDMI se vuelve muy atractivo por su capacidad y calidad, y considerando que hoy Blu-ray, que utiliza MVC con $n=2$, utiliza HDMI, es notorio que el estándar de compresión y esta interfaz tienden a resultar un buen complemento.

El desarrollo de la tecnología hace que los televisores con capacidad para 3D sigan aumentando en variedad, sin embargo la falta de estandarización y normativa hace que estos aún siendo de última tecnología requieran en algunos casos un Set-Top-Box. No obstante la industria sigue trabajando en pro de producir televisores digitales que incluyan tanto el receptor digital y el soporte para 3D, para ello es fundamental que cada estándar de TDT especifique como brindará el servicio de 3DTV, tal como DVB ya lo hizo.

6.4 REPRODUCCIÓN DEL VIDEO ESTEREOSCÓPICO.

Esta etapa comprende la cúspide de todo el sistema debido a que aquí es donde efectivamente el usuario deberá percibir el grado de calidad y realismo que una escena en 3D le puede proporcionar. Así como formatos, las técnicas de reproducción sobre una pantalla son de lo más variadas; obviamente el reproducir las imágenes dependerá del tipo de información que al televisor llegue, es de ahí que van surgiendo tecnologías, donde dependiendo del número de vistas, información exacta de profundidad, etc., la pantalla será capaz de mostrar al usuario las escenas estereoscópicas.

En el capítulo 1 se describieron las técnicas básicas para la presentación de los cuadros 3D, estos principios se mantienen y son usados en la actualidad, también de estos principios se desprende la clasificación de las pantallas estereoscópicas, la cual además de basarse en diferentes técnicas, marca el desarrollo de la tecnología para lo que en el futuro se debe desarrollar en pantallas.

A pesar de que la clasificación para las pantallas aún no es del todo clara se pueden establecer los siguientes tipos:

- Pantallas estereoscópicas.
- Pantallas autoestereoscópicas.
- Pantallas Volumétricas.

- Pantallas Holográficas.

Si bien aún no está definido si la autoestereoscopia, holografía y volumetrismo estén considerados como 3D en TV, no cabe duda que así como los formatos avanzan, estas pantallas representan la forma de mostrar al usuario múltiples vistas y calidad que con los formatos avanzados de 3D se pretende lograr. Por otro lado algunos autores mencionan a las pantallas holográficas y volumétricas como parte de las autoestereoscópicas. Aún es variada la forma de definirlo pero dentro de cada tipo existen algunas técnicas y además algunas clasificaciones.

6.4.1 PANTALLAS ESTEREOSCÓPICAS.

Estas pantallas parten del principio de la estereopsis, es decir de dos imágenes diferentes a cada ojo, para lograrlo se recurren a técnicas en las que el usuario debe utilizar filtros, gafas para aislar las imágenes correspondientes a cada ojo.

La percepción de profundidad se logra debido a que si las imágenes son diferentes, se produce la disparidad binocular horizontal que el cerebro humano representa como profundidad. Las técnicas de aislamiento de las vistas pueden realizarse simultáneamente o mediante una multiplexación temporal; para el primer caso se recurre a los fenómenos físicos de la luz que el ojo detecta, para darle características diferentes para cada ojo. En el segundo caso la idea es secuenciar la presentación de las vistas de manera muy rápida en diferentes instantes de tiempo a cada ojo, en este caso no se altera a la luz de proyección de la imagen como tal.

6.4.1.1 Pantallas Estereoscópicas Anaglíficas

En este caso el filtro a ser usado por el espectador es un selector de longitud de onda, la pantalla en sí no es especial a las tecnologías ya conocidas (TRC, PDP, LCD, LED-LCD). En el capítulo 1 y 2 se abordó tanto como formato como modo de presentación este tema, cabe mencionar que aún es utilizado aunque genere fatiga ocular en muchos casos y no promete ser un esquema que perdure mucho más, tomando en cuenta las tecnologías existentes y el prometedor avance de las mismas.

6.4.1.2 Pantallas Estereoscópicas Polarizadas

El tema de la polarización fue ya analizado en el capítulo 1, en este caso la pantalla debe ser capaz de generar una polarización de la luz para la imagen izquierda y otra para la derecha, y estas dos juntas presentarlas en el mismo cuadro al espectador. Se pueden aplicar los principios de polarización lineal o circular. La técnica micro polarizada (uPol) se inventó en 1990 por Sedeg Faris y se involucra la adición de un filtro especial polarizado. La forma más sencilla de implementar un display polarizado es utilizar un material en frente de la pantalla convencional (PDP, LCD, LED), que polarice la luz de las líneas sucesivas de los píxeles que contengan la información de imagen para cada ojo, comúnmente se utiliza polarización circular derecha e izquierda para cada ojo (ver Figura 3.8).

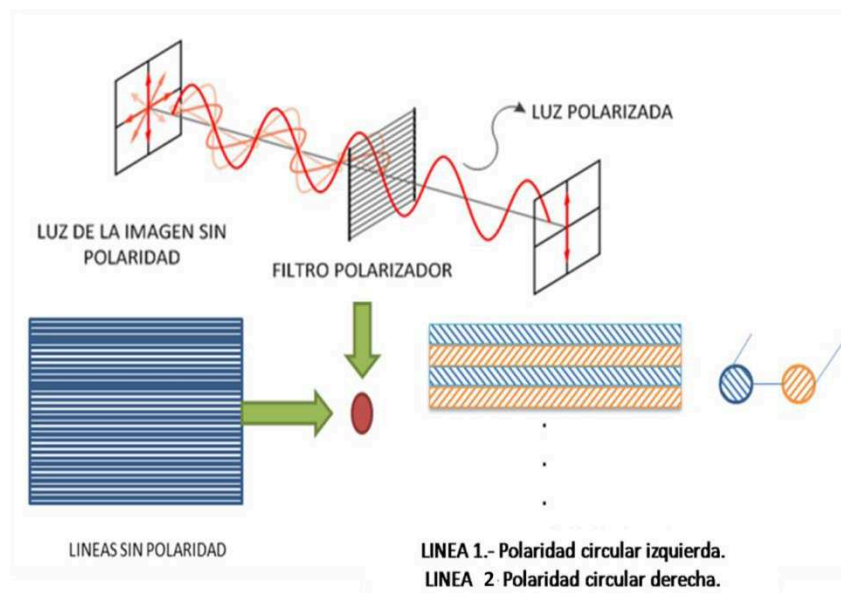


Figura 3. 8 Efecto de la pantalla polarizadora.

Obviamente se presenta un problema de reducción de resolución de la imagen original (Polarización unidimensional).

Para solucionar el problema de la resolución y de la pérdida de líneas se usa una polarización en 2 dimensiones, también llamada multiplexación espacial (ver Figura 3.9).

Cada imagen es modulada usando un operador para crear un patrón del tipo tablero de ajedrez siendo los bloques negros para imagen izquierda y el resto para la derecha.

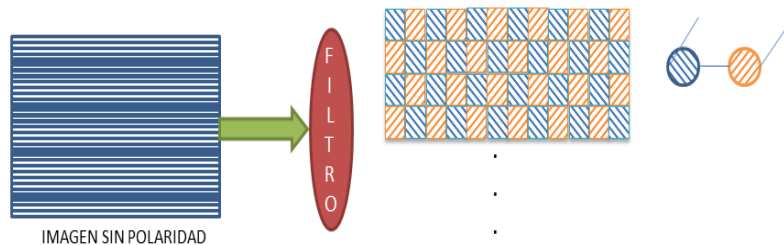


Figura 3. 9 Polarización en 2 dimensiones.

Se nota aún algún tipo de pérdida de resolución en la imagen aunque la imagen mejora en calidad por ser tratada más uniformemente.

Actualmente algunos productos usan un uPol (micro polarizador) llamado X-pol, que incluye una máscara negra entre las columnas de los píxeles para incrementar el rango vertical de la zona de vista y disminuir el Crosstalk. La técnica polarizada es mayormente usada en LCD y ha evolucionado a tecnologías innovadoras, como la del panel dual LCD, también conocida como display de ángulo de polarización variable (Figura 3.10). Convencionalmente cada sub-píxel del LCD trabaja como un obturador que controla el paso de la luz, en este caso la función óptica es diferente, dado que la disposición de los elementos ópticos varía según la Figura 3.10 lo detalla.

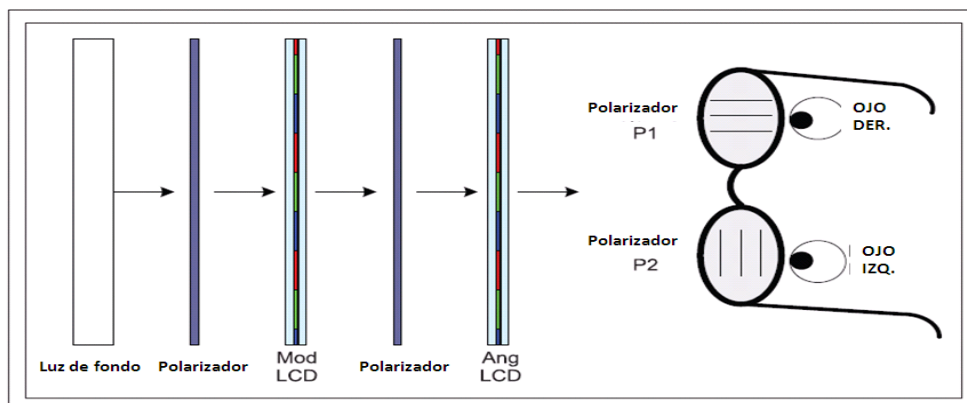


Figura 3. 10 Pantalla de ángulo de polarización variable.

El primer panel (Mod LCD), modula el brillo de la luz de cada pixel tal como lo convencional, excepto que la imagen enviada al primer panel es una mezcla de imágenes izquierda y derecha. El segundo panel (Ang LCD), actúa para controlar el ángulo de polarización de salida para cada pixel (usando como rotador de polarización a la celda de cristal líquido).

6.4.1.3 Pantallas Estereoscópicas Basadas en la Multiplexación Temporal.

Estas pantallas inicialmente fueron inviables para LCD debido a que no tenía el tiempo de respuesta necesario para pasar dos cuadros donde normalmente se mostraría uno, por ello para el PDP resultó más fácil la implementación. Estos displays son propensos a Crosstalk debido las propiedades de desintegración por sí misma, de la tecnología de la pantalla. Aquí se muestran alternadamente las imágenes izquierda y derecha (resultaría a 120 Hz.) (ver Figura 3.11). Estos cuadros se observan mediante obturadores de cristal Líquido (LCS) o gafas LCS (Liquid Crystal Shutter); que permiten la visión del ojo izquierdo o derecho sincronizadamente con la que se está presentando en la pantalla. La velocidad de conmutación de estas LCS es alrededor de 2 microsegundos; a la pantalla pueden llegar dos vistas en cualquier formato ya conocido, con su respectiva señalización para que estas puedan ser separadas y presentadas.

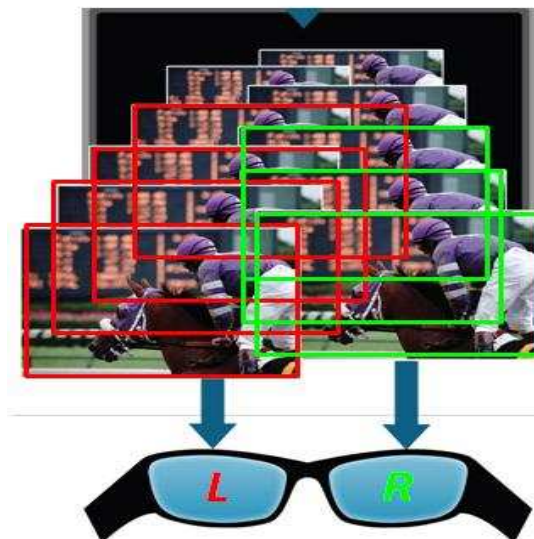


Figura 3. 11 Pantalla de multiplexación temporal.

De esta tecnología se pueden mencionar las desventajas de que las gafas activas resultan más costosas y con una infraestructura más compleja debido a que se debe tener especial cuidado con la sincronización.

Obviamente en estas pantallas se representan las imágenes sin afectar la resolución, porque estas se presentan íntegras al ojo sin disminuir el número de líneas que la compone; no obstante si en transmisión se ha utilizado algún formato como checkboard que ha disminuido la resolución, se emplea el método de Wobulation para recuperar la calidad presentando los cuadros a 120 Hz.

6.4.1.3.1 Funcionamiento y estructura de los LCS (Liquid Crystal Shutter)

En estas pantallas la única tecnología requerida es un procesador de video capaz de trabajar a una tasa de 100, 120 o 144 cuadros por segundo. Donde realmente se desarrolla la técnica es en las gafas LCS, estas gafas contienen para cada ojo un vidrio que contiene un cristal líquido de material nemático, el cual tiene la propiedad de obscurecerse si se aplica un voltaje, y de otro modo ser transparente.

Los cristales son controlados remotamente desde la pantalla por Infrarrojo, RF, DLP Link o Bluetooth; el transmisor envía una señal de sincronismo que permite que los cristales se obscurezcan alternadamente sobre cada ojo de acuerdo a la tasa de la pantalla.

La comunicación entre las gafas y la pantalla es de dos estados discretos 1L ó 0L definiendo uno para cada vista, ya sea por RF con ASK (Amplitude-shift keying), infrarrojo con pulsos si codificar, etc. Cada fabricante utiliza esta técnica a su manera y la forma de enviar los pulsos es variable, los protocolos de transferencia del sincronismo son variados, y por esta variedad es muy difícil una estandarización en este aspecto, la duración de la apertura para cada ojo depende del fabricante y del método de transmisión que este usa⁵¹.

La publicación “A Survey of 3D Sync IR Protocols” hace una buena reseña de cómo funcionan los protocolos en cada marca y los tiempos que obtura el

⁵¹Andrew Woods and Jesse Helliwell, Curtin University, “A Survey of 3D Sync IR Protocols”, Perth, western Australia, Marzo 31, 2011

sincronismo para las gafas activas, un ejemplo de ello es DLP Link que utiliza pulsos de luz visible, según lo explica la Figura 3.12.

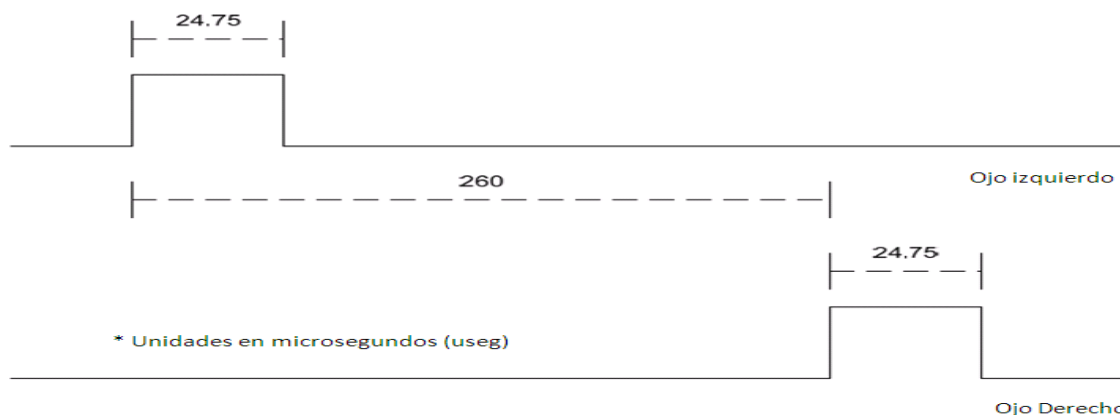


Figura 3. 12 Protocolo LCS – DLP Link⁵².

El presentado es un protocolo de los más sencillos y menos sofisticados, otras marcas, utilizan más pulsos por unidad de tiempo con el fin de combatir posibles espurias de señal IR, sin embargo utilizar pulsos más largos y en configuraciones más complejas aumentan la probabilidad de interferencia debido al aumento de período.

6.4.2 PANTALLAS AUTO-ESTEREOCÓPICAS⁵³.

Esta pantalla es capaz de transmitir una imagen a cada ojo sin necesidad de que el espectador utilice filtros o gafas de ningún tipo. La pantalla tiene una tecnología tal que limita el paso de la luz correspondiente a cada ojo, haciendo que a la posición de la retina izquierda o derecha solamente llegue su respectiva imagen. El problema principal es la precisión de la tecnología para direccionar las imágenes izquierda y derecha, por ello los fabricantes han desarrollado algunas técnicas para localizar los ojos del espectador y aumentar el número de espectadores que puedan ver la imagen.

En principio, si consideramos un espectador, el número de vistas necesario para representar el 3D es 2, sin embargo el objetivo de no usar gafas es que el sistema

⁵²Andrew Woods and Jesse Helliwell, Curtin University, "A Survey of 3D Sync IR Protocols", Perth, western Australia, Marzo 31, 2011

⁵³Dodgson N. "Autoestereoscopic 3D Displays", Computer Laboratory, University of Cambridge, IEEE Computer Society, Agosto 2005.

se haga cómodo, no solo para uno si no para muchos televidentes. De ahí que existe una relación directa entre los formatos de producción 3D y las pantallas autoestereoscópicas, ya que a mayor número de vistas adquiridas, más características se tiene de un objeto; y en una pantalla autoestereoscópica se pretende emular el volumen de un objeto, por lo tanto se vuelve necesario involucrar formatos multivista.

Si se lo implementa muy exactamente el efecto que el usuario percibe es el de estar viendo un holograma, aunque solo sea un conjunto de imágenes en el plano de la pantalla.

Las pantallas autoestereoscópicas combinan 2 efectos, el paralaje estéreo y el paralaje en movimiento. Cuando un observador mira una escena de la vida real, él observa una imagen diferente para cada ojo y diferentes imágenes si éste mueve la cabeza, es decir se tienen infinito numero de imágenes de una escena.

Para las pantallas autoestereoscópicas se pretende dividir el campo visual del usuario en slots o espacios para que en cada uno, solo se situé una imagen, y que el usuario pueda, según el modelo, tener espacios discretos de vistas de la escena, con ello cada espacio representa una vista y cada vista debe ser proyectada en la pantalla (ver Figura 3.13).

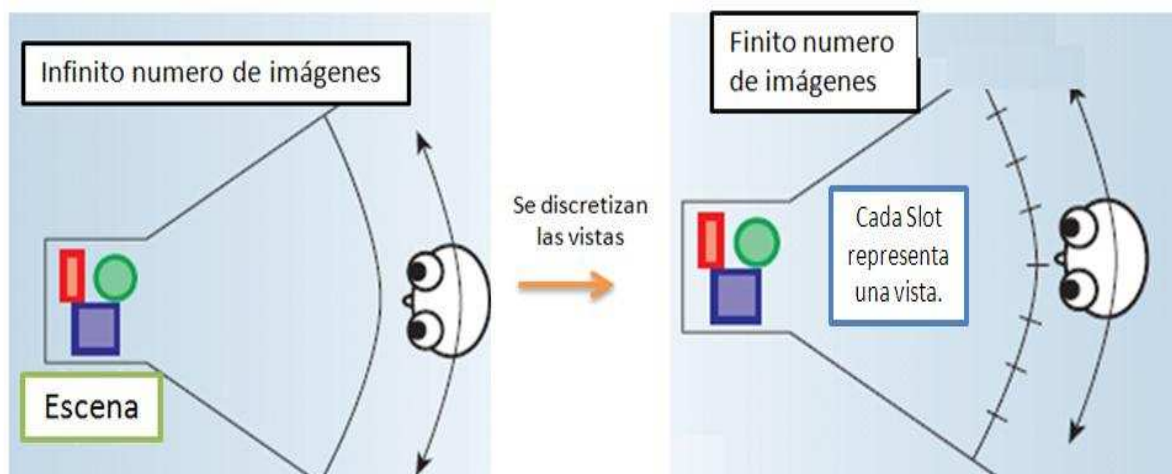


Figura 3. 13 Discretización del espacio visual

De acuerdo a su construcción y número de vistas, las pantallas autoestereoscópicas pueden ser:

- Pantallas de dos vistas.
- Pantallas con seguimiento de la cabeza (comúnmente 2 vistas).
- Pantallas multivista ($n > 2$).

6.4.2.1 Pantallas Auto-Estereoscópicas de dos vistas.

El objetivo es claro, determinar un método para filtrar las imágenes correspondientes a cada ojo, para ello los investigadores^{54, 55} han usado 2 técnicas, expuestas en la Figura 3.14, la barrera de paralaje, y la hoja (capa) lenticular.

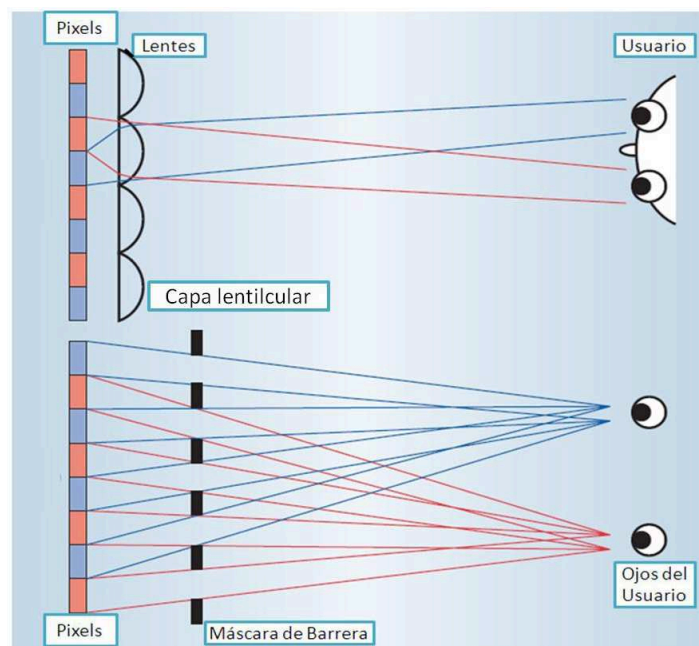


Figura 3. 14 Métodos para direccionar las imágenes a cada ojo.

La Figura 3.14 muestra, de dos formas, cómo se pueden implementar pantallas autoestereoscópicas con multiplexación espacial para 2 vistas. Estas pantallas dividen la resolución horizontal de la pantalla (normalmente LCD) en dos partes, los píxeles de las columnas pares corresponden a la imagen destinada al ojo

⁵⁴S.A. Benton, ed., "Selected Papers on Three-Dimensional Displays", SPIE—Int'l Soc. for Optical Eng., 2001.

⁵⁵ L. Lipton, "Foundations of the Stereoscopic Cinema", Van Nostrand Reinhold, 1982

izquierdo y las columnas impares contienen la información del otro ojo. Tanto la barrera como los lentes hacen que solamente los haces que corresponden a cada ojo se proyecten; se observa la gran desventaja de que si el usuario se mueve no se percibiría el efecto. En el primer caso se tiene, un arreglo de pequeños lentes (hoja lenticular) en frente de la imagen dividida en sus pixeles, direccionando la luz de las columnas de los pixeles adyacentes hacia diferentes espacios de visión, si el usuario se ubica a la distancia ideal, cada ojo solo mirará las columnas de pixeles necesarias de manera muy exacta. Esta hoja lenticular contiene una serie de lentes cilíndricas moldeadas en un substrato plástico.

En el caso de la barrera, esta se hace normalmente de un material opaco con fracturas verticales transparentes y finas con un espaciamiento regular. Cada espaciamento transparente actúa como ventana a un trozo vertical de la imagen puesta detrás de ella. Sin embargo con la tecnología, hoy se puede utilizar el principio de los obturadores LCD de las LCS que se diseñan, de tal manera que se aplica un voltaje en el LCD de las columnas que se desea opacar. En este caso de dos vistas existen solo determinadas posiciones en las cuales el espectador recibe correctamente la imagen (ver Figura 3.15), si el espectador se ubica a la distancia precisa, este percibirá una imagen estereoscópica, no obstante los problemas son variados; existe un 50% de probabilidad de que el espectador se ubique mal ante la pantalla y no perciba el efecto (efecto pseudoscópico), y es realmente incómodo para el espectador limitarse a una posición y que esta se pierda fácilmente ante cualquier movimiento.

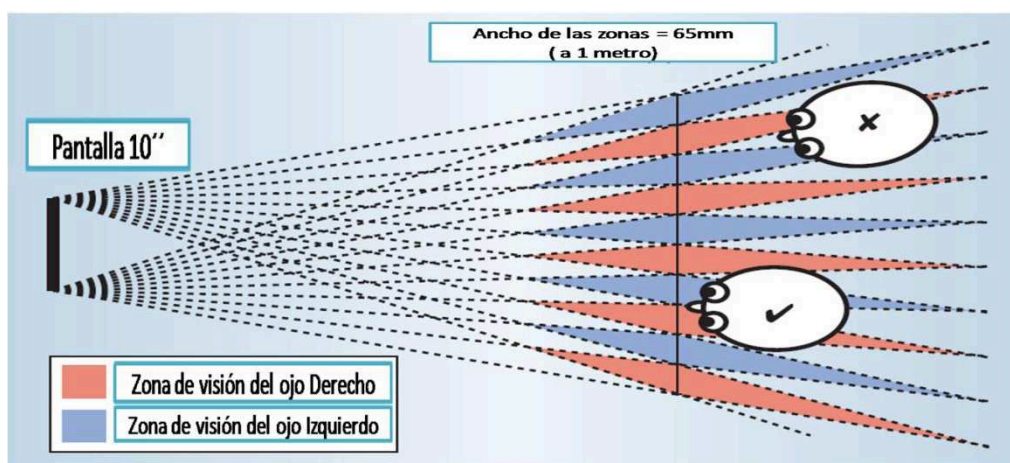


Figura 3. 15 Zonas de visión para n=2 vistas.

Tanto la barrera como el arreglo lenticular, provocan zonas para que el espectador ubique sus ojos, una pequeña variación de esta posición provoca una falla en la percepción tal como se muestra en la Figura 3.15. Asimismo el alejarse de la distancia óptima de visión provoca un desajuste de la zona, el tamaño de esta zona termina siendo del tamaño de la distancia inter-pupilar aproximadamente, considerando una pantalla de 10'', ubicándose el espectador a 1 metro.

6.4.2.2 Pantallas Auto-Estereoscópicas con seguimiento de la cabeza.

Estas pantallas tienen la capacidad de adaptar las zonas de visión de las imágenes izquierda y derecha la posición del espectador, es decir la pantalla conoce la posición del usuario. Para mover las zonas hacia el espectador existen dos opciones fundamentales (ver Figura 3.16), la primera desplazar todas estas zonas coherentemente con el movimiento del espectador, trabajando evidentemente para un solo espectador. Otra opción es crear solamente dos zonas de visión (L, R) y estas ponerlas en la posición perfecta para el observador.

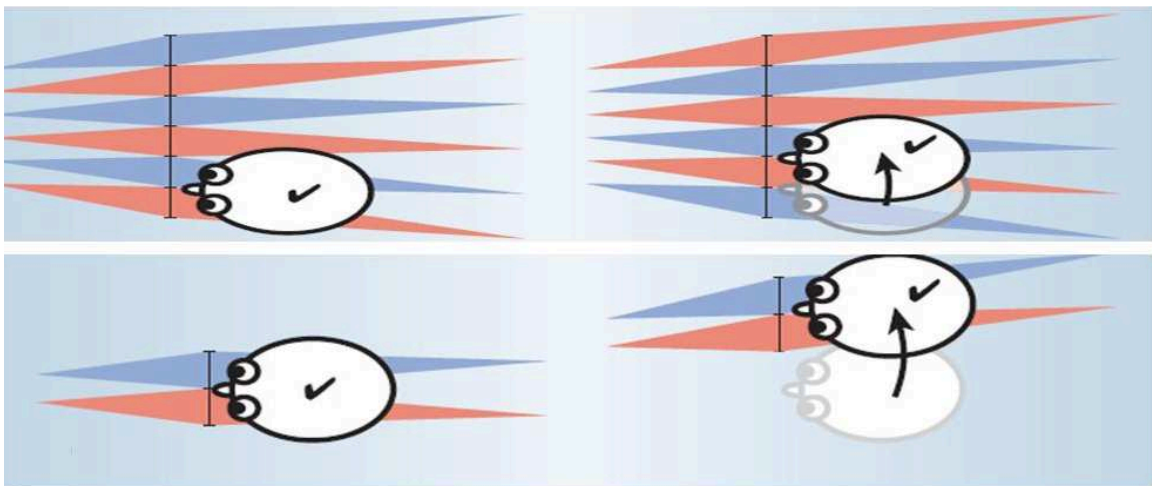


Figura 3. 16 Opciones de movimiento de la zona de visión.

El problema surge cuando la distancia de separación de los ojos de algún usuario es muy diferente a lo que la pantalla puede calcular.

Hoy en día se propone una nueva tecnología completamente diferente donde se muestran solo dos zonas permitiendo que estas sean movidas físicamente. La

primera alternativa, desarrollada por Xenotech, propone dos métodos de seguimiento: o los proyectores de la pantalla se mueven para desplazar las zonas de visión, o toda la pantalla en sí rota para seguir el movimiento de la cabeza del espectador. Por otro lado Sharp Electronics, propone un sistema en el cual la fuente de luz debe ser desplazada para mover las zonas de visión (ver Figura 3.17).

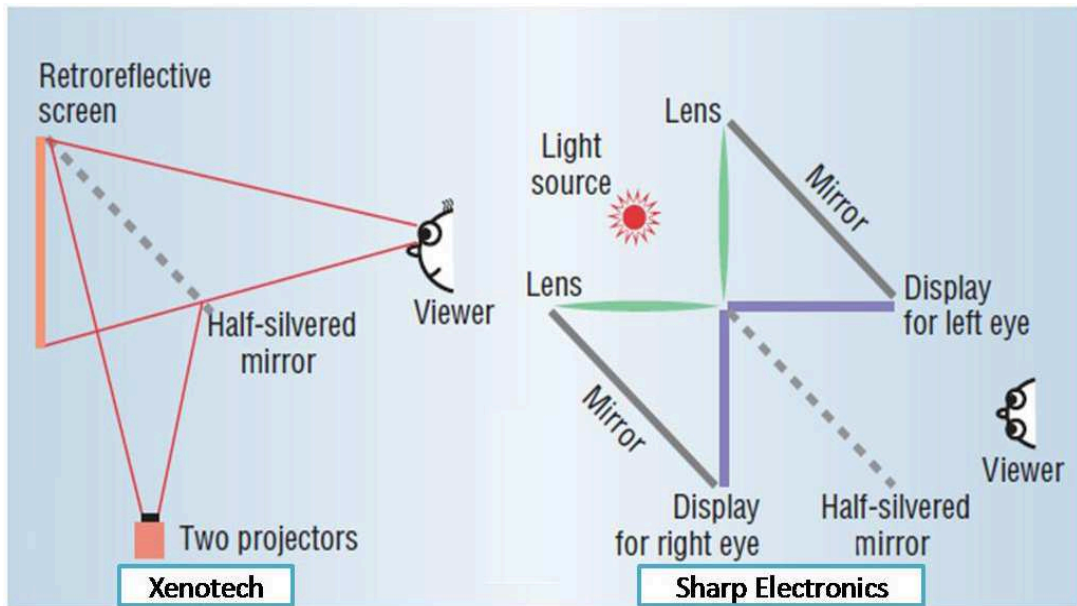


Figura 3. 17 Métodos de movimiento de la zona de visión Xenotech y Sharp.

Xenotech establece dos proyecciones de las imágenes L+R, estas se reflejan en el espejo semi-plateado hacia la pantalla retro-reflectiva, la cual devuelve los haces luminosos de la imagen hacia el usuario en la misma dirección.

Sharp propone en cambio dos pantallas, las cuales se iluminan por la fuente de luz y en el momento de que los espejos son ligeramente desplazados, se generan dos zonas de visión haciendo que cada ojo solo observe una de las dos pantallas.

Estos sistemas deben superar algunas dificultades para ser viables; se debe impedir que para el seguimiento, el espectador utilice equipos adicionales. Por otro lado algunos sistemas utilizan movimientos mecánicos para el seguimiento, este debe ser preciso, exacto y muy eficiente; y por ultimo estos sistemas deben ser multiusuario ya que establecer toda una estructura de seguimiento y una pantalla de tanta tecnología para un usuario es subutilizar el sistema.

6.4.2.3 Pantallas Auto-Estereoscópicas multivista.

Estas pantallas tienen como información de la escena más de 2 vistas, y por lo tanto crean la ventaja de que los espectadores perciban una imagen 3D si sus dos ojos están dentro de una zona de visión (ver Figura 3.18). Estas pantallas permiten el acceso de múltiples espectadores a la imagen, teniendo cada uno de ellos una perspectiva diferente de la escena; el tener tantas vistas provoca que no sea necesario un mecanismo de seguimiento, ya que no importa el lugar de los espectadores, estos estarán en alguna zona donde una de las vistas de la imagen ocurre y se da con una percepción 3D. Las dificultades que se presentan son; la dificultad de generar muchas vistas, y la gran complejidad de la construcción de una pantalla que genere tantas zonas de visión y más aún, si tanta complejidad va a ser aprovechada al máximo por n usuarios y no va a ser subutilizada.

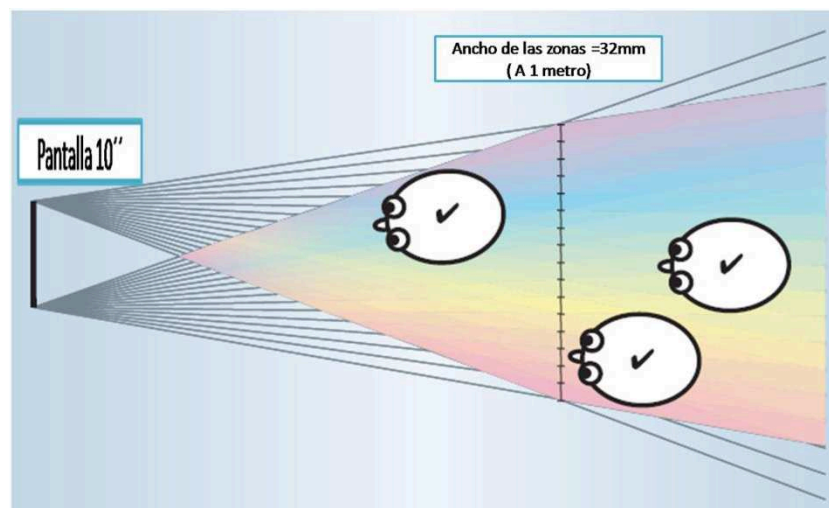


Figura 3. 18 Zonas de visión en pantallas autoestereoscópicas multivista.

6.4.2.4 Tecnologías para la elaboración de pantallas Auto-Estereoscópicas.

Una vez conocidos los tipos de pantallas existentes es necesario considerar las tecnologías que se aplican para la fabricación de las pantallas autoestereoscópicas; la investigación ha determinado tres tipos de tecnologías:

- Multiplexación espacial.
- Tecnología Multi-proyector.
- Tecnología secuencial en tiempo.

3.2.4.2.1 *Tecnología de multiplexación espacial.*

En esta tecnología se divide la resolución en función del número de vistas existentes; lo usado hasta ahora para hacerlo han sido: barreras de paralaje, la iluminación del paralaje, y hojas lenticulares. Las pantallas deben tener un Pixel Pitch⁵⁶ fijo para poder alinear la barrera, un CRT con tal precisión es imposible por lo cual se recurre al uso de PDP y LCD. El tamaño del pixel y la resolución del cristal líquido y del plasma han limitado la aplicabilidad a 4 vistas.

Además, el uso de las barreras de paralaje provocan una disminución de la luz conforme se aumentan el número de vistas, y de igual forma se convierte muy compleja la construcción de múltiples vistas.

Al implementar las pantallas con hojas de lentes, se provoca un oscurecimiento de algunas zonas entre pixeles, para combatir este efecto, los investigadores y las compañías que piensan implementar estas pantallas, modifican la estructura y la disposición de las lentes y la hoja lenticular para disminuir la incomodidad visual. Entre las técnicas probadas están, multiplexar horizontal y verticalmente, colocar la hoja lenticular ligeramente de forma diagonal, y usar elementos ópticos holográficos.

3.2.4.2.2 *Tecnología Multiproyector.*

Se expuso el principio multi-proyector para dos vistas en la sección 3.4.2.2. No obstante se ha propuesto otro método válido que usa un proyector por cada vista existente, estos proyectan las imágenes sobre una pantalla especial transmisiva o reflexiva, tal como una doble hoja lenticular (ver Figura 3.19).

Los inconvenientes de estas tecnologías son: el costo de tener tantos proyectores como vistas en un número elevado de las mismas, y que estas pantallas proyecten las imágenes correcta y precisamente alineadas unas con otras.

Sin embargo del costo, experimentalmente se han logrado sistemas con más de 100 vistas.

⁵⁶ Pixel Pitch: Espacio entre dos fósforos coloreados de un píxel de una pantalla.

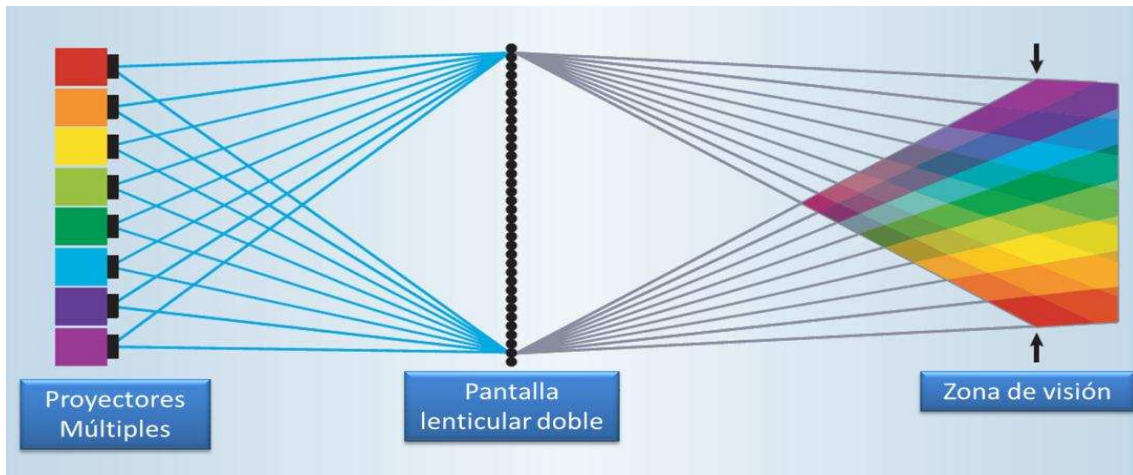


Figura 3. 19 Funcionamiento de la tecnología multi-proyector.

3.2.4.2.3 Tecnología secuencial en tiempo.

Esta tecnología utiliza un dispositivo de pantalla simple a una tasa de cuadro alta, posteriormente se utiliza un elemento óptico secundario para direccionar las imágenes a los lugares correspondientes (ver Figura 3.20). La implementación se logra mediante barras de iluminación que actúan sobre los lentes de tal manera que crean la zona de visión precisa, esta iluminación es a una alta tasa para que el ojo humano no perciba el cambio. Se ha implementado prácticamente con CRT, sin embargo esa tecnología esta obsoleta; hoy en día se utilizan LCD, LED y unidades de procesamiento de imágenes GPU más eficientes que logran la multiplexación a altas tasas para que el espectador perciba el efecto 3D.

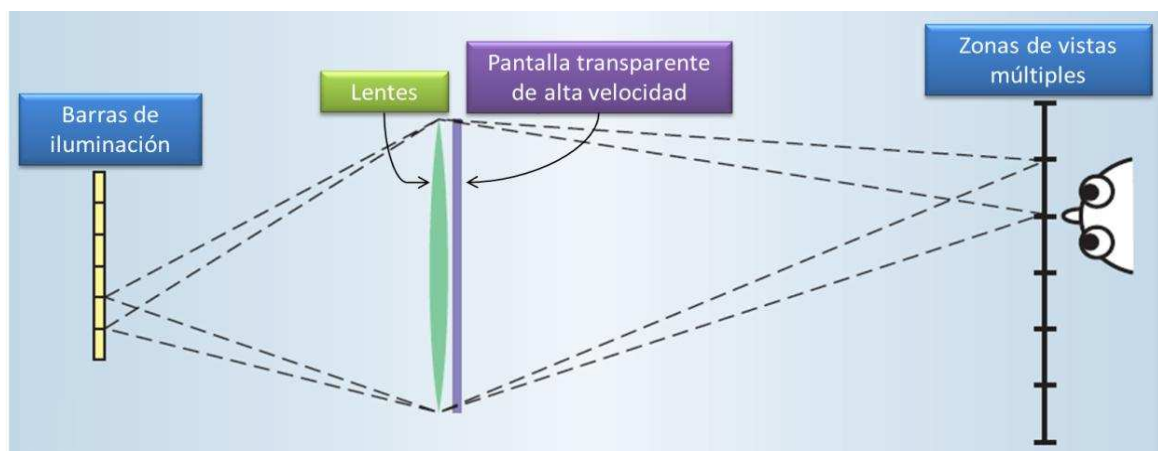


Figura 3. 20 Funcionamiento de la tecnología secuencial en tiempo.

6.4.3 PANTALLAS VOLUMÉTRICAS Y HOLOGRÁFICAS.

Es importante mencionar que para muchos estudios, las pantallas holográficas son una clase de pantallas volumétricas y que su estudio está íntimamente relacionado. Es decir se considera una pantalla volumétrica, aquella que físicamente plasma un volumen correspondiente a un objeto o escena, utilizando para este fin diversos mecanismos ópticos y eléctricos.

Tanta complejidad de pantalla, sin duda merece información de alta calidad, por lo cual es necesario recurrir a los formatos avanzados o cualquier metodología que permita captar la “onda del objeto”.

La luz que atraviesa el espacio hacia nuestros ojos no es una fuente puntual de luz, sino más bien la suma de todas las señales luminosas procedentes de todos los puntos y todos los ángulos. Esta totalidad de los rayos de luz, que pasarían a través de una imagen vacía si la sostenemos entre nosotros y los objetos, se denomina “onda del objeto”. El registro óptico de esta “onda del objeto” producirá una reproducción sin fatiga de la imagen real para que podamos verla.

La “onda del objeto”, como todas las ondas, tiene magnitud, longitud de onda y fase. Lamentablemente, en la actualidad no se cuenta con los medios necesarios para registrar en un sensor algo más que la amplitud de la onda. Si se puede encontrar un sensor o un método práctico para grabar la fase de la onda del objeto al igual que la amplitud, podrá desarrollarse una solución para la televisión 3D real. Hoy en día la onda del objeto se crea por la renderización⁵⁷ y procesamiento en software de la información que se obtienen en la adquisición.

6.4.3.1 Pantallas Volumétricas

Las pantallas volumétricas iluminan directamente puntos del espacio dentro del volumen mismo de la pantalla, a través de iluminar u opacar un espacio de la imagen volumétrica, con eso se crea la apariencia de un volumen transparente en el espacio. El elemento común para las diferentes pantallas volumétricas, es que el volumen o región que ocupa las tres dimensiones tiene componentes que

⁵⁷ Renderización: Proceso de interpretar la escena en tres dimensiones en una imagen bidimensional que contenga la información del volumen de los objetos de la escena.

pueden ser posicionados representados y/o manipulados en algunos casos. Muchas de las tecnologías de estas pantallas no presentan restricción en lo correspondiente al ángulo de visión, haciendo que el espectador no tenga que buscar una posición ideal frente a la pantalla. Algunos tipos de pantallas volumétricas son:

- *Dispositivos de estado sólido*, que generando pulsos de luz con una fuente externa, despliegan información de los voxels (pixel con volumen) dentro de un substrato transparente. Ejemplo de ello se da al usar dos haces de láser infrarrojo con diferente longitud de onda para excitar los electrones hacia un nivel de energía más alto y poder emitir luz visible.
- *Pantallas Volumétricas multi-planares* que construyen el objeto 3D a partir de un conjunto de imágenes 2D multiplexadas en tiempo, mediante movimientos rápidos o elementos rotativos en la pantalla.
- *Pantallas de espejo varifocal* que aplican espejos flexibles para barrer una imagen de una pantalla CRT, a través de diferentes planos de profundidad del volumen de una imagen.

6.4.3.2 Pantallas Holográficas

Estas pantallas tienen un enfoque diferente a las anteriores, las pantallas holográficas reconstruyen la luz emitida por un objeto 3D a partir de franjas de interferencia generadas a través de un proceso de grabación holográfica. Las franjas de interferencia, si son correctamente iluminadas, funcionan como una rejilla difractiva compleja que reconstruye tanto la dirección como la intensidad de la luz reflejada del objeto.

La capacidad de almacenamiento de una placa fotográfica utilizada para holografía es enorme, y va mucho más allá de los actuales medios electrónicos de distribución. Por ello, hoy existen varios métodos experimentales tales como, paralaje horizontal, compresión del ancho de banda holográfica y hardware digital más rápido, que representan opciones para lograr un holograma aceptable y real para el video Streaming.

Ya sean volumétricas u holográficas los prototipos existentes son muy variados y en ellos se aplican principios, técnicas y algunos conceptos muy precisos que a continuación se abordan.

6.4.3.3 Infraestructura de las pantallas volumétricas.

Para representar el volumen de una escena, las técnicas que se han usado son variadas, entre ellas se encuentran:

- *Barrido de Volumen*, este principio es el fundamental de las pantallas volumétricas, consiste en llenar el volumen de la imagen por medio de reflejar o transmitir luz de una superficie plana rotativa u oscilante dentro del volumen deseado. Dado que la superficie de proyección se barre a través del volumen en cuestión, esta refleja o emite luz sincronizada con su posición. Si el volumen es frecuentemente repetido (20 Hz) se percibirá el efecto 3D.
- *Volumen estático*, estas pantallas pueden generar la imagen 3D a partir de la persuasión de volumen por luz emitida, en la cual las propiedades macroscópicas quedan intactas; se logra con cristales dopados con iones de tierras raras, que emiten puntos de luz que se excitan por dos haces de luz láser duales.
- *Hologramas y estereogramas holográficos*, si bien aún no se acepta totalmente que un holograma cree una imagen volumétrica, este fenómeno ya fue explicado en el punto 3.4.3.2
- *Pantallas 3D altamente multivista*, con la tecnología estudiada se pueden lograr de 200 a 300 vistas y el espectador no tiene restricción en las zonas de visión.

La mayoría de estas pantallas volumétricas tienen la propiedad de tener imágenes con paralaje total, permitiendo múltiples espectadores con puntos de vista tanto verticales como horizontales, dándole una gran ventaja ante las autoestereoscópicas. La resolución y la frecuencia de cuadro dependen del motor de proyección triple, estos tienen una resolución 1024x768 pixeles y una frecuencia de cuadro de entre 5 y 10 KHz.

El ancho de banda óptico es muy importante considerar, dada la cantidad de datos involucrados, 8000 cuadros por segundo en imágenes monocromáticas de profundidad, teniendo valores de 6.3 Giga pixeles por segundo.

6.4.3.4 Algunos prototipos de pantallas volumétricas y holográficas existentes.

Entre otras, las pantallas con capacidad de generar hologramas se enumeran en la Tabla 3.1

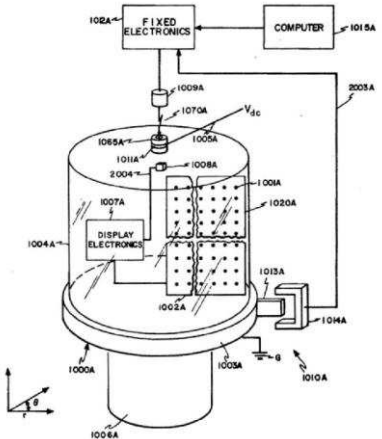
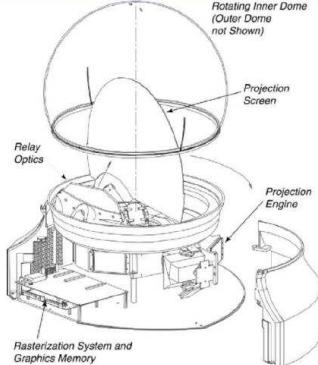
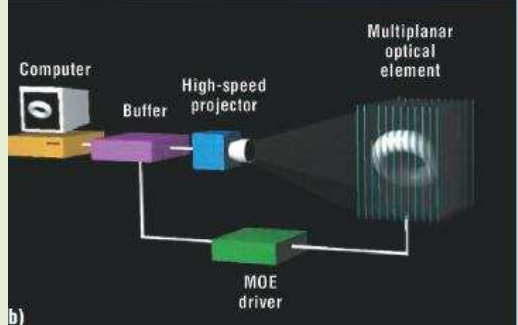
Pantalla.	Descripción.
	<p>Se barre el volumen, usando una superficie cubierta con emisores de alta velocidad como LEDS.</p>
	<p>Sistema con esfera de 10 cm de diámetro, con llenado completo de imágenes, en 360°.</p> <p>Para generar una imagen 3D, se proyectan 198 patrones 2D, llamados slices, que rotan a 900 rpm. La pantalla barre el volumen entero 2 veces teniendo una frecuencia de refresco de 30 Hz.</p>
	<p>Pantalla Depth Cube, genera el 3D con una secuencia de planos 2D, el buffer es dual para que el proyector actúe sobre cada plano con su respectiva proyección.</p>

Tabla 3. 1 Prototipos de pantallas holográficas existentes.

Aún estas tecnologías no han sido investigadas a fondo para aplicaciones de Televisión, sin embargo el avance existente en medicina, microscopía, imágenes médicas, etc. ha hecho que exista la posibilidad de que se den cambios aplicables a la TDT.

CAPÍTULO 4

SITUACIÓN ACTUAL DE LA TV-3D Y TEMAS RELACIONADOS

4.1 INVESTIGACIÓN ACTUAL Y PROYECTOS VIGENTES.

El servicio de televisión estereoscópica consta, según lo estudiado, de un gran número de procesos y elementos que deben ser desarrollados y considerados; de ahí que tantas compañías tecnológicas, organismos de estandarización, universidades técnicas, etc. proponen, diseñan y estructuran alternativas en todos los campos, para que el servicio de TV-3D llegue a su futura implementación. Los estudios y proyectos vigentes comprenden una gama extensa de contenidos, desde el estudio y perfeccionamiento de los principios físicos de la estereoscopía, hasta soluciones técnicas para la reproducción, procesamiento y transmisión de las imágenes tridimensionales. Las empresas productoras de tecnología cumplen aquí un rol muy importante, marcas reconocidas en el mundo de la televisión buscan hoy introducir sus productos, tanto para el hogar como para las estaciones tele-difusoras y así lograr ganar terreno en una aplicación de TV aún no explorada.

Los organismos de estandarización también cumplen un papel fundamental, si bien es cierto estos no pueden influir en el mercado de manera directa, si pueden y deben establecer los lineamientos para que se pueda crear tecnología que garantice compatibilidad y aplicabilidad.

En los capítulos anteriores, se mencionan algunos de los proyectos que han existido y que han aportado al desarrollo de la 3D-TV, no obstante esta sección está dedicada completamente al análisis de algunas de las propuestas, soluciones, y por supuesto ventajas y desventajas de los factores vigentes. Los documentos, exposiciones y convenciones son muy numerosos, por lo tanto es imposible detallar el análisis de cada uno de ellos; se procurará dar una idea clara de los más relevantes y sus aportaciones más significativas.

4.1.1 ORGANIZACIÓN “SD&A”

La SD&A (Stereoscopic Displays and Applications) comprende una serie de conferencias y con mundiales acerca de la innovación 3D en todos los aspectos, anualmente (desde 1995 hasta la actualidad) se realizan convenciones que generan cientos de conferencias y miles de publicaciones sobre la innovación de pantallas, modelos de cámaras, contenidos de alta calidad, interfaces de usuario, etc. Además aquí se presentan productos que las compañías más reconocidas lanzan al mercado, empresas como NVIDIA, JVC, Fujifilm, 3D-TV Solutions, Real 3D, participan para dar a conocer sus avances. En estas convenciones también tienen lugar prototipos creados por Universidades como The Osaka City University y The University of Tsukuba.

Esta organización se ha convertido en la expresión más representativa y la que reúne a más miembros de la comunidad tecnológica, que desarrolla la técnica estereoscópica.

Las memorias de cada una de sus convenciones se exponen en su página web <http://www.stereoscopic.org>. Algunos de los prototipos y temas presentados en los últimos años en dicha convención se detallan a continuación:

- Tema: Tecnología estereoscópica para la realización de la película 3D “Beauty and the Beast”, año 2010 por Tara Turner de Disney Feature Animation.
- Foro de discusión "3D Video Standards - At Last!!" y "The Business of 3D: How to Build Successful Business Models". Año 2010 por Mark Stockfisch (Consumer Electronics Association and Quantum Data, Inc.), Steve Venuti (HDMI LLC), David Broberg (Society of Cable Telecommunications Engineers and Cable Labs), Pete Ludé (SMPTE and Sony Electronics), y Mike Weissman (True Vision Systems).
- Demostración de LCS (gafas activas) NVIDIA 3D con 4 de los nuevos monitores LCD Acer GD235HZ full-HD 24" 3D LCD, se incluye 3D Blu-ray, 3D Fuji 3D camera. – 2010 por NVIDIA (Dave Cook and Michael McSorley).

- Demostración del algoritmo de conversión automática 2D a 3D en una pantalla 3D HDTV. – 2010 por Philips (Rene Klein Gunnewiek)
- Demostración de un sistema de display volumétrico, (ver Figura 4.1) usando un arreglo de rejilla de espejos superiores (RMGA) y un motor de proyección DLP. El RMGA consiste de un arreglo bi-dimensional de reflectores esquineros diedrales y se forma una imagen en una posición simétrica al plano por Osaka City University (Daisuke Miyazaki)



Figura 4. 1 Prototipo de sistema Volumétrico.

- Foro de discusión “Factores que determinan el tamaño de la Pantalla” por Bernard Mendiburu (Chair), David Broberg (Cable Labs), John Merrit (The Merritt Group), Jukka Häkkinen (Aalto University) y Frédéric Devernay (INRIA).
- Demostración por Volfonide de gafas activas que pueden trabajar con cualquier técnica pasiva de polarización circular para un display polarizado.

Estos son algunos de los más significativos temas tratados en las convenciones del 2010 y 2011, cabe señalar que desde 1995 se tienen miles de exposiciones y temas tratados.

Cabe destacar que las convenciones no se han dedicado plenamente al estudio de las tecnologías y de los equipos existentes, sino también a evaluar factores de viabilidad y factibilidad de la implantación de la tecnología 3D más formalmente.

4.1.2 PROYECTO “3D VIVANT”

El proyecto 3D Vivant⁵⁸ está apoyado por la European Comission a través del programa de las Tecnologías de Información y Comunicación FP7 (Frame Programm 7), que consta de nueve miembros, cada uno con su función específica (ver Tabla 4.1). El proyecto se enmarca en el desarrollo de un sistema completo, desde la adquisición hasta la presentación de video Holoscópico 3D, es decir crear un holograma o volumen de color que se represente en el espacio independientemente de en qué posición el espectador lo vea. Dentro de los objetivos de este proyecto se encuentran:

- El diseño y construcción de una cámara capaz de adquirir el holograma en ultra alta definición a partir de un solo lente, con ello poder desarrollar un algoritmo para calcular la profundidad de las escenas 3D.
- Usar displays autoestereoscópicos existentes para poder desplegar las imágenes.
- Añadir a la experiencia de video 3D, técnicas de audio espacial para generar un sistema completo que brinde al espectador sensación total de volumen en audio y video.
- Proveer de un códec de video escalable basado en códecs de código abierto existentes como H.264 / AVC.
- Por último se busca integrar completamente un sistema end-to-end desde la captura y producción, procesamiento y visualización en dos escenarios, el Broadcasting y el Streaming por IP.

MIEMBRO	FUNCIONES
Universidad de Brunel - West London,	Coordinador del proyecto y contribuye a la investigación de imágenes Holoscópicas 3D, llevando a la simulación de diseño de la óptica y las pruebas, la segmentación en 3D Holoscópico del objeto de vídeo, la mezcla real y virtual del vídeo Holoscópico 3D y la demostración del contenido 3D en un LCD Holoscópico superpuesto con una matriz de microlentes.
CERTH (Centro de Investigación y Tecnología Hellas, Informática e Instituto de Telemática.)	El Centre for Research and Technology Hellas–Informatics and Telematics Institute, contribuye al etiquetado de contenidos, búsqueda y recuperación necesarias en el contenido 3D Holoscópico.

⁵⁸Website del proyecto: <http://dea.brunel.ac.uk/3dvivant/consortium.html>

Institut für Rundfunktechnik GmbH (Instituto de Radio Tecnología)	IRT (Alemania) se centra en las soluciones técnicas para la captura de audio espacial y la reproducción. IRT también trabaja en el desarrollo de un modelo de atención visual en 3D y contribuye a las interfaces 3D de usuario.
Holografika (Hungría)	Holografika contribuirá a la tecnología de visualización 3D, el desarrollo de algoritmos capaces de representar el contenido Holoscópico en tiempo real en las pantallas de campo de luz y el desarrollo de una interfaz gráfica de usuario en 3D
RAI Research Centre	RAI contribuye a la definición de los casos de uso de radiodifusión y de los escenarios de usuarios de VIVANT 3D para usuarios aficionados y profesionales, así como la percepción del usuario y las pruebas de campo con los productores y consumidores de servicios
Rundfunk Berlin-Brandenburg (Radiodifusora Berlin-Brandenburg)	RBB es un socio de contenido / servicio en VIVANT 3D. Sus principales contribuciones son el desarrollo de escenarios de utilidades y servicios de referencia que se utilizará para el desarrollo de VIVANT 3D, así como para fines de demostración. En el RBB se establecen los requisitos para los socios de tecnología en términos de contenido y creación de servicios y validar los resultados de VIVANT 3D y hacer recomendaciones de los puntos de vista del sector de la emisión y producción.
Instituto de Telecomunicações (Instituto de Telecomunicaciones)	El IT (Portugal), contribuirá a la creación de Códecs 3D necesarios para la entrega de los contenidos en 3D Holoscópico haciendo posible la transmisión de vídeo 3D Holoscópico en un ambiente propenso a errores, también se encargan de lograr alcanzar una calidad de vídeo adecuada en el lado del decodificador.
European Broadcast Union (Union Europea de Radiodifusion UER)	La EBU se centrará en el proyecto de captura y validación de los requerimientos del usuario en el entorno de la radiodifusión en desarrollo, y utilizar su influencia en la difusión, capacitación y normalización. Además, encargarse de la distribución y evaluación subjetiva, la UER contribuirá a la evaluación de la calidad en 3D, el estudio y las cuestiones de los sistemas generales.
Arnold & Richter Cine Technik (München – Alemania)	ARRI lleva las especificaciones y los paquetes de la innovación técnica de trabajo para el sistema de cámaras VIVANT 3D, con base en los resultados de la encuesta de necesidades del usuario, y también proporcionará el hardware para realizar pruebas. La compañía participará en la integración y pruebas de los sistemas desarrollados en el proyecto, así como en las pruebas de campo.

Tabla 4. 1 Miembros y funciones del 3D Vivant.⁵⁹

⁵⁹ Fuente: <http://dea.brunel.ac.uk/3dvivant/consortium.html> (Web Site del Proyecto)

Este proyecto conforma un claro ejemplo de la tendencia hacia lograr el desarrollo total del sistema en su conjunto, y no solamente a limitarse a cámaras y pantallas.

4.1.3 PROYECTO “3D4YOU”

Proyecto formado en el año 2008, cuya característica principal es enfocar la aplicación 3D a la televisión, 3D4YOU tiene el objetivo de establecer un sistema eficiente totalmente independiente a las tecnologías de las pantallas y compatibles con los servicios monoscópicos, se promueve también la búsqueda de un sistema total de producción que sea práctico para la televisión tanto técnica como comercialmente. 3D4YOU se fundó bajo el programa ICT (Information and Communication Technologies) Work Programme 2007-2008 junto con el FP7. Los miembros de este proyecto son: Philips (Holanda, Coordinator), la BBC (Inglaterra) de Londres, el Instituto Alemán Fraunhofer Heinrich Hertz Institut, Technicolor Thomson R&D de Francia, la empresa France Télécom, la Universidad alemana Christian Albrechts Universität y la Productora de cine alemana KUK Film produktion.

El concepto básico de este proyecto es realizar un sistema de TV3D totalmente práctico, y para ello la clave es tener contenido 3D para difundir. El proyecto define el formato en el proceso de adquisición (formatos estudiados en el capítulo 2) con el fin de crear un sistema con interoperabilidad y compatibilidad.

El proyecto 3D4YOU ha identificado 3 problemas técnicos principales para un sistema de TV3D.

- Es muy complejo capturar video 3D con la tecnología actual de cámaras. Al menos se requieren de dos cámaras que operen simultáneamente, con una geometría ajustable y conocida. El desplazamiento de las cámaras estéreo, debe ser ajustable para poder capturar la profundidad del objeto cuando este está cerca y lejos del punto de referencia.
- El video 3D que es adquirido con dos cámaras, no brinda suficientes elementos para crear un sistema que no utilice gafas, y pantallas autoestereoscópicas multivista.

- Los métodos de post-producción del video y los estándares son incipientes y poco desarrollados.

Este proyecto 3D4YOU pretende atacar estos problemas, utilizando de manera inteligente un arreglo complejo de 2 o 3 cámaras de alta resolución, más una o dos cámaras de baja resolución con sensores de rango de profundidad. Estos elementos, crearían video 3D de alta calidad sin necesidad de adquirir e implementar equipos muy costosos y de última tecnología.

3D4YOU propone el sistema genérico de adquisición, y la mayor ventaja es el costo que representa frente a sistemas multivista de $n=100$ vistas, sin embargo el arreglo y la combinación de las cámaras en la adquisición debe ser muy ingenioso y de un diseño muy elaborado.

4.1.4 PROYECTO “MOBILE 3DTV”

Este proyecto culminó y presentó sus resultados el 8 de Junio del 2011, los colaboradores del proyecto fueron: Tampere University of Technology, Technische Universität Ilmenau, Middle East Technical University, Fraunhofer Heinrich-Hertz-Institute, Multimedia Solutions Ltd. y Tamlink Innovation-Research-Development Ltd.

El proyecto se fundamenta sobre DVB-H, el proyecto, de duración 3 años, fue fundado por la FP7 de la Unión Europea. El principal objetivo es demostrar la viabilidad de la implementación de TV3D móvil sobre DVB, y con ello que los mecanismos de creación del video (formatos) codificación y transmisión sean apropiados.

Este proyecto también se enmarca el sistema end-to-end, abordando desde la adquisición del video hasta la experiencia visual del usuario, incluyendo los equipos móviles que serían parte del sistema.

4.1.5 PROYECTO “TV3D 2.0”

Proyecto planeado para trabajar entre 2010 y 2013, cuyos colaboradores son: Alcatel-Lucent Bell, Androme, Grin, Philips, Innovative Applications, Grass Valley

Nederland B.V. y los grupos de investigación: IBBT-MICT-UGent , IBBT-MMLab-UGent , IBBT-EDM-UHasselt , IBBT-SMIT-VUB , IBBT-IPI-UGent.

El objetivo del proyecto es preparar la siguiente generación de aplicaciones 3DTV, a través de la mejora en la tecnología de adquisición y displays, y del estudio de la experiencia del usuario en su ambiente natural a través de experimentos de laboratorio.

Este proyecto está mucho más enfocado a la experiencia visual y la mejora de contenidos, ya que se enfoca en las cámaras y pantallas, asimismo estudia y procura entender las necesidades del espectador para que más allá de lo técnico, se genere un sistema cómodo para el usuario y con acogida.

4.1.6 AGENTES INVESTIGADORES Y DESARROLLADORES PARA LA TV3D

La cantidad de organizaciones, universidades, corporaciones, etc. que investigan y desarrollan el tema es muy grande, en las secciones anteriores se abordaron algunos proyectos de relevancia, en los cuales ya se mencionaron las organizaciones que se han apropiado del tema, y han generado los conocimientos necesarios para crear equipos y hacer viable al sistema.

A continuación se exponen algunos agentes investigadores en diferentes áreas, que permiten que el servicio de televisión estereoscópica se desarrolle y pueda alcanzar una completa evolución.

4.1.6.1 Organismos de estandarización.

Sin duda estos organismos deben estar involucrados, ya que serán los que en último término crearán las normas para que los fabricantes puedan desarrollar el mercado de la TV3D; la UIT tiene dentro de sus diferentes grupos de trabajo, cuestiones y recomendaciones alusivas a la TV3D en todos los niveles. La Comisión de Estudio 6 del Sector de Radiocomunicaciones de la UIT (UIT-R) ha publicado un nuevo informe en el que se establece una hoja de ruta para la introducción de la futura TV3D, en el que se prevén unos sistemas tan inteligentes que son capaces de reproducir fielmente la manera en que los ojos y cerebros

perciben el mundo visual, este informe es el estudiado en el capítulo 2 y el que permitió establecer los niveles y generaciones del servicio de la TV3D.

Otros organismos de estandarización también están presentes, tales como; el MPEG (Moving Picture Experts Group) que es un grupo de trabajo del ISO/IEC y que ha permitido el desarrollo de los estándares de codificación vitales en la televisión estereoscópica, la ISO-IEC también muy importante a la hora de proponer soluciones basándose en lo estándares de MPEG.

El estándar HDMI también ha participado generando formatos que pueden funcionar sobre la interfaz del mismo nombre.

El IEEE (Institute of Electrical and Electronics Engineers) también ha contribuido, ya que sus miembros generan constantemente publicaciones basadas en experiencias y estudios que permiten un desarrollo sostenido de la TV3D. Parte importante del presente estudio se ha basado en las publicaciones realizadas por miembros IEEE.

4.1.6.2 Investigadores académicos y centros de investigación.

La comunidad académica y científica como tal tiene un espacio importante, ya que generan conocimientos científicos puros y sin influencia del mercado ni factores externos, toman mayoritariamente en cuenta el aspecto técnico.

Por otro lado, los centros de investigación de las corporaciones electrónicas más importantes en el mundo generan también información valiosa en el desarrollo. La Tabla 4.2 detalla algunos de estos agentes investigadores y algunas de sus aportaciones científicas más significativas.

En la Tabla 4.2 se observa claramente investigaciones en todos los aspectos, y en muchos casos temas que intentan abordar en la totalidad el sistema 3DTV, existen centros de investigación característicos que han generado más de una investigación y otros especializados en ciertos temas. Lo importante en todo caso, es la convergencia hacia el tema y el gran interés que despierta, ya que los elementos mostrados en la Tabla 4.2 apenas representan un pequeño porcentaje del total de agentes investigadores alrededor del mundo.

AGENTES INVESTIGADORES	TEMAS TRATADOS
Tzu-Der Chuang, Pei-Kuei Tsung, Pin-Chih Lin, Lo-Mei Chang, Tsung-Chuan Ma, Yi-Hau Chen, Yu-Han Chen, Chuan-Yung Tsai, and Liang-Gee Chen DSP/IC Design Lab, Graduate Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan	<ul style="list-style-type: none"> • Chip escalable decoder para 3d HDTV y aplicaciones de Streaming de 59,5 mW • Chip codificador MVC para 212 Mpixeles para 3DHDTV
Disney Research, Zurich - Dr.-Ing. Aljoscha Smolic	3D Video and Free Viewpoint Video
Pei-Kuei Tsung, Ping-Chih Lin, Kuan-Yu Chen, Tzu-Der Chuang, Hsin-Jung Yang, Shao-Yi Chien, Li-Fu Ding, Wei-Yin Chen, Chih-Chi Cheng, Tung-Chien Chen, Liang-Gee Chen	A 216fps 4096×2160p 3DTV Set-Top Box SoC for Free-Viewpoint 3DTV Applications
Fraunhofer Institute for Telecommunications - Heinrich-Hertz-Institut	Investigaciones varias 3DTV, formatos, MVD, etc.
Thomson Research Digest - Technicolor	Encodings contents for 3DTV, pruebas y resultados experimentales
Mitsubishi electric research laboratories	Investigación de los formatos y propuesta de estándar para TV autoestereoscópica.
Nokia Research Center	Colaboración para la edición y lanzamiento del anexo MVC de H.264
Jung-Young Son, Seokwon Yeom, Dong-Su Lee, Kwang-Hoon Lee, and Min-Chul Park como parte del Mid-career Researcher Program through National Research Foundation of Korea (NRF)	A Stereoscopic Camera Model of Focal Plane Detector Array
Qualcomm MM R&D team	A 3D camera solution for mobile platforms
Australian Communications and Media Authority (ACMA)	Temporary trials of 3D TV and other emerging technologies
Levent ONURAL - Dept. of Electrical and Electronics - Bilkent University - Ankara, Turkey	An assessment of 3DTV technologies
Thomas SIKORA- Institut für Telekommunikationssysteme - Technische Universität Berlin - Berlin, Germany	
Jörn OSTERMANN Institut für Informationsverarbeitung - University of Hannover - Hannover, Germany	
Aljoscha SMOLIC Fraunhofer-Inst. für Nachrichtentechnik - Heinrich-Hertz-Institut - Berlin, Germany	
M. Reha CİVANLAR - Dept. of Computer Eng - Koç University - Istanbul, Turkey.	
John WATSON Dept. of Engineering University of Aberdeen, - King's College - Aberdeen, Scotland	

Tabla 4. 2 Agentes Investigadores y aportaciones significativas

4.2 SITUACIÓN TÉCNICA ACTUAL DE LA TELEVISIÓN ESTEREOSCÓPICA.

La situación actual, vista desde el aspecto técnico, está íntimamente relacionada a factores como: desarrollo experimental y pruebas de campo realizadas, equipos disponibles en el mercado y experiencia de uso del sistema en otras áreas diferentes a la televisión. Estos aspectos brindarán un panorama claro al respecto de establecer el “estado del arte” del servicio de televisión estereoscópica principalmente para el estándar ISDB-T internacional.

4.2.1 DESARROLLO EXPERIMENTAL Y PRUEBAS DE CAMPO REALIZADAS.

El desarrollo experimental se ha dividido en tantas partes como procesos involucra el sistema de TV3D, es así que, desde el desarrollo y la investigación de cámaras, algoritmos de codificación, pantallas entre otros, se configuran en elementos suficientes para que las empresas pertinentes puedan producir la tecnología acorde a las necesidades.

4.2.1.1 Pruebas y experimentos en la etapa de adquisición del video 3D.

Para adquirir el video, los formatos estudiados permiten obtener muchas experiencias, desde las más sencillas con cámaras estéreo, hasta usar decenas de cámaras para poder captar muchas vistas de un objeto.

Las cámaras estéreo de dos lentes ya se comercializan actualmente, por lo tanto ya no siguen siendo objeto de investigación profunda, en este caso las imágenes se obtienen de acuerdo al formato o de manera independiente.

Para sistemas multivista los experimentos son interesantes como el publicado en la revista de procesamiento de señal del IEEE del 17-12-2010. Se utilizaron arreglos de 100 cámaras de alta definición en diferentes configuraciones (ver Figura 4.2).

Con tal número de cámaras las escenas realmente adquieren un gran realismo en las pantallas que usualmente son volumétricas u holográficas, el mismo experimento estableció su propio sistema de reproducción.

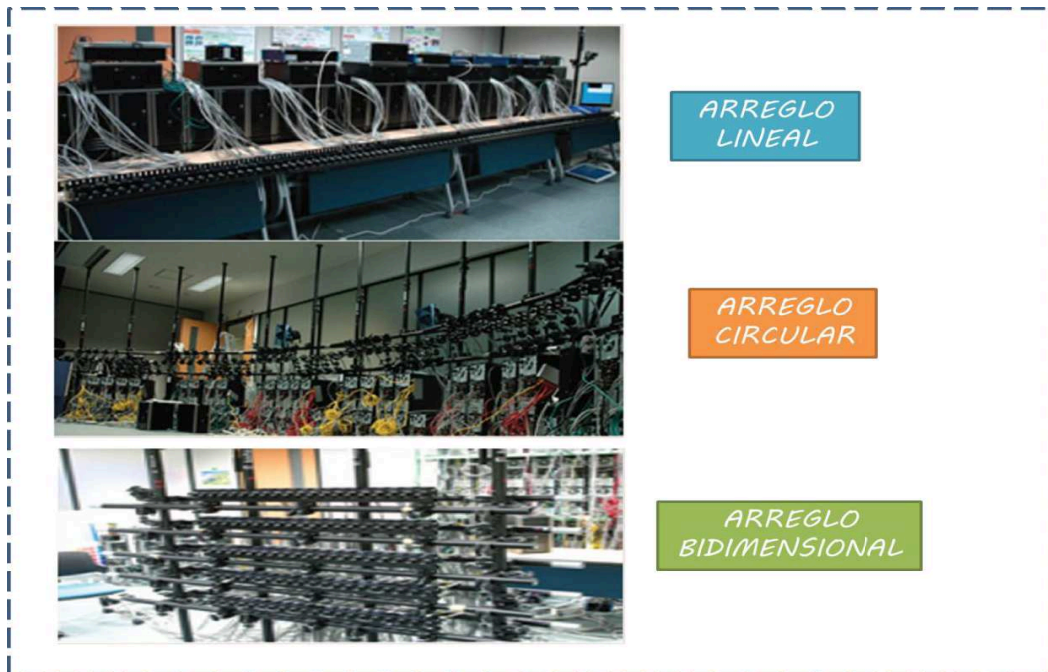


Figura 4. 2 Arreglo de cámaras (100) en sistema Multivista

Las cámaras usualmente a nivel experimental generan imágenes de baja resolución para poder aplicar mayores procesos, sin embargo resoluciones de SDTV y HDTV también han sido probadas a tasas de 30 y 60 fps. Los detectores CCD de dichas cámaras utilizan 8 bits por muestra. En el caso de las cámaras de profundidad, estas son de baja resolución monocromáticas de 8 bits.

El análisis geométrico de la ubicación es analizado mediante pruebas y resultados de las imágenes, de todas formas los posibles errores en la geometría de la adquisición pueden ser solventados en el pre procesamiento, mediante los algoritmos de interpolación y creación de vistas.

En general el sistema 3DTV requiere de un sistema de n cámaras ($n \geq 2$) en los cuales intervienen aspectos como calibración, estimación de la relación estéreo y estimación del mapa de profundidad en caso de existirlo, los experimentos desarrollados han permitido establecer sobre todo la parte de la calibración para que esta sea usada en la producción de cámaras, para el mercado el desarrollo

experimental de las pruebas que permiten la calibración es extenso y puede ser realizado de varias maneras y depende del tipo y número de cámaras, las pruebas y métodos de calibración se presentan en el Anexo 3.

4.2.1.2 Pruebas y experimentos en la etapa de procesamiento y codificación.

Cuando ya se cuenta con la información del video, las pruebas se enfocan principalmente a establecer la eficiencia de los codificadores H.264, las pruebas son tantas como formatos existentes, es decir, existen resultados experimentales para sistemas Simulcast, para los formatos compatibles con el cuadro HD, y para los videos codificados con las actualizaciones SVC y MVC de H.264; muchos de estos datos fueron usados en el presente estudio para determinar el uso de los segmentos OFDM para un canal de TV3D. Las condiciones de estudio son varias, experimentos en LD, SD, HD, utilizando secuencias variadas de video. Por ejemplo un estudio realizado por Michael Dröse, Carsten Clemens, y Thomas Sikora de la Universidad de Berlín, donde aplicando 8 cámaras y codificación MVC, se obtienen los resultados experimentales mostrados en la Figura 4.3; el Anexo 4 muestra algunos otros trabajos con resultados experimentales no solo en codificación si no también en transmisión en algunos casos.

View	JSVM		MVC	
	Rate	PSNR	Rate	PSNR
0	227.5	33.0	200.0	33.9
1	217.3	33.4	200.0	33.8
2	218.4	33.1	200.0	34.3
3	212.4	33.5	200.0	33.5
4	179.6	34.4	200.0	34.3
5	177.2	34.6	200.0	34.3
6	230.0	32.8	200.0	33.5
7	214.4	33.3	200.0	34.1
Avg.	209.6	33.5	200.0	34.0

Table 1. Sequence Race1 320 frames encoded using JSVM and the presented MVC (PSNR [dB], Rate[kbit/s]).

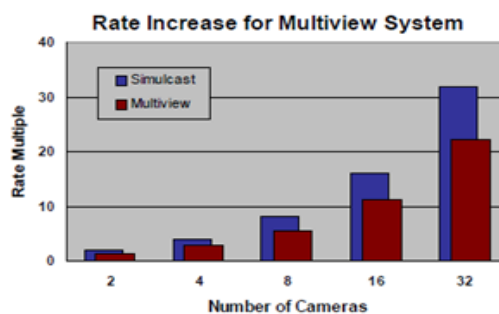


Figura 4. 3 Resultados experimentales de codificación MVC.⁶⁰

Como se observa en la Figura 4.3 los experimentos se realizan para algunas secuencias, y así mismo en condiciones diferentes de calidad (PSNR), evaluando

⁶⁰ "Extending single-view scalable video coding to multi-view based on H.264/AVC" Michael Dröse, Carsten Clemens, and Thomas Sikora; Department of Communication Systems Technical University of Berlin, 10587 Berlin, Germany; 2006.

algunos criterios clave para la posterior transmisión, como es la tasa de bits para cierta relación SNR o para cierto formato de las vistas.

4.2.1.3 Pruebas y experimentos en la etapa de Transport Stream, Multiplexación y Transmisión y recepción de la señal.⁶¹

Si se comprueba que la codificación ha sido eficiente con tasas de bits adecuadas, el siguiente paso es evaluar los anchos de banda y el flujo que se va a enviar, si bien es cierto ISDB-T aún no tiene definido claramente cómo repartir sus segmentos en 3D otros sistemas ya han realizado transmisiones con algunas características particulares. La Figura 4.4 muestra las condiciones de diferentes procesos experimentales de broadcasting donde ya se experimentan resoluciones propias de alta definición y dependiendo del estándar varían las técnicas utilizadas.

	Terrestrial	Cable	Satellite
Broadcasting Standard	ATSC	OpenCable	DVB-S2
Frequency	CH 66	CH 97	12.59 GHz
Bandwidth	6MHz	6MHz	36MHz
Modulation	8 VSB	256 QAM	8 PSK
FEC (code rate)	TCM (2/3)	TCM (38/40)	LDPC (3/4)
Effective data rate	19.39 Mbps	38 Mbps	35 Mbps
Bit-rate allocation for 3D video (video codec)	Left : 12 Mbps (MPEG-2) Right: 6 Mbps (H.264)	Left : 17.5 Mbps (H.264) Right: 17.5 Mbps (H.264)	Left : 17.5 Mbps (H.264) Right: 17.5 Mbps (H.264)
Image Resolution	1920x1080 @ 60i	1920x1080 @ 60i	1920x1080 @ 60i
Service area	Local (Seoul)	Local (Seoul)	Nationwide



Figura 4. 4 Condiciones de Prueba de la TV3D en algunos sistemas⁶¹.

⁶¹ Fuente: "3DTV Broadcasting and Distribution Systems" Namho Hur, Member, IEEE, Hyun Lee, Member, IEEE, Gwang Soon Lee, Sang Jin Lee, Member, IEEE, Atanas Gotchev, Member, IEEE, and Sang-II Park, Member, IEEE -IEEE TRANSACTIONS ON BROADCASTING, VOL. 57, NO. 2, JUNE 2011. (Las Figuras de esta sección se refieren a la misma fuente)

Para verificar la conveniencia de adaptar los servicios 3DTV HD a un método dual-stream, los métodos experimentales han existido desde el inicio del 2010; en primer lugar, a las pruebas se han enfocado en la capacidad de los IRD's y los STB de trabajar con flujos dobles de datos, y cuántos de estos equipos disponibles en el mercado pueden hacerlo. También se debe probar si es que la gran calidad de las imágenes se mantiene bajo estas condiciones.

Como resultado de algunas investigaciones en algunos equipos receptores existentes, se determinó que algunos de estos equipos fallaron al decodificar las señales 3DTV usando el método dual Streaming; pero que para la mayoría no existieron problemas para decodificar y reproducir las señales estereoscópicas.

Para evaluar la calidad de las imágenes, los cuadros izquierdo y derecho fueron codificados en algunas posibilidades de bitrate, por ejemplo 12 Mbps para la imagen izquierda y 6 Mbps para la imagen derecha, 13 Mbps para la izquierda y 5 Mbps para la derecha y así sucesivamente. Para estas pruebas tanto el hardware y el software de los codificadores 3DTV fueron también desarrollados, también se desarrolló el STB 3D sobre una tarjeta RF PCI (Peripheral Component Interconnect) y su respectivo software. También se determinó la interfaz HDMI 1.4a para la conexión hacia el televisor, el laboratorio de recepción de pruebas se muestra en la Figura 4.5.



Figura 4. 5 Laboratorio de Pruebas de la TV3D.

En Octubre de 2010 se llevó a cabo exitosamente una transmisión de 3DTV sobre tres tipos de redes de transmisión, terrestre, cable y satélite. Para el servicio se configuraron líneas terrestres dedicadas, transmitiéndose la información sobre el

canal 66 (785.14 Mhz), alrededor de 100 STB's y 100 Televisores se instalaron en diversas ubicaciones tanto públicas como privadas. Con ello se pudieron evaluar tanto la experiencia del usuario como la calidad subjetiva de las imágenes.

4.2.1.4 Pruebas “Codificación y transmisión de 1080p 3D en 6 segmentos ISDB-T internacional.”⁶²

Esta es una de las pocas pruebas realizadas sobre el estándar ISDB-T internacional, esta prueba consistió en la transmisión de señal estereoscópica de alta definición, utilizando equipos existentes y basándose en CPU's para el procesamiento del video.

La prueba completa se describe en el Anexo 4. El diagrama del sistema de prueba se observa en la Figura 4.6



Figura 4. 6 Sistema de Pruebas de TV3D sobre ISDB-T

Esta prueba demuestra la adaptabilidad del sistema ISDB-T internacional a la tecnología 3D, ya que se concluye, en esta prueba, que las modificaciones que son hechas por el CPU para procesar y poder transmitir el cuadro estereoscópico, pueden ser reemplazadas por un chip en el Set-Top-Box y por el eficiente y correcto manejo de los equipos de transmisión.

⁶² Fuente: “Nuevos Encoders y normas para TDT y Argentina Conectada”, Dr. Ing. Mario Mastriani, Universidad Nacional de Tres de Febrero, Laboratorio de Imágenes y Señales, Mayo 2011

4.2.1.5 Pruebas y experimentos en la etapa de reproducción de la señal.

Para formatos de dos vistas la reproducción no representa grandes obstáculos, es por ello que las pruebas se han enfocado a experimentar la representación de perfiles multivista, con pantallas especializadas y con métodos de reproducción complejos. Estas pruebas⁶³ no han considerado efectos del canal ni de transmisión, solo han especificado la calidad de reproducción de una escena adquirida con n vistas. Con las 100 vistas descritas en las pruebas de adquisición se utilizó un sistema “space-ray” descrito en la figura 4.7.

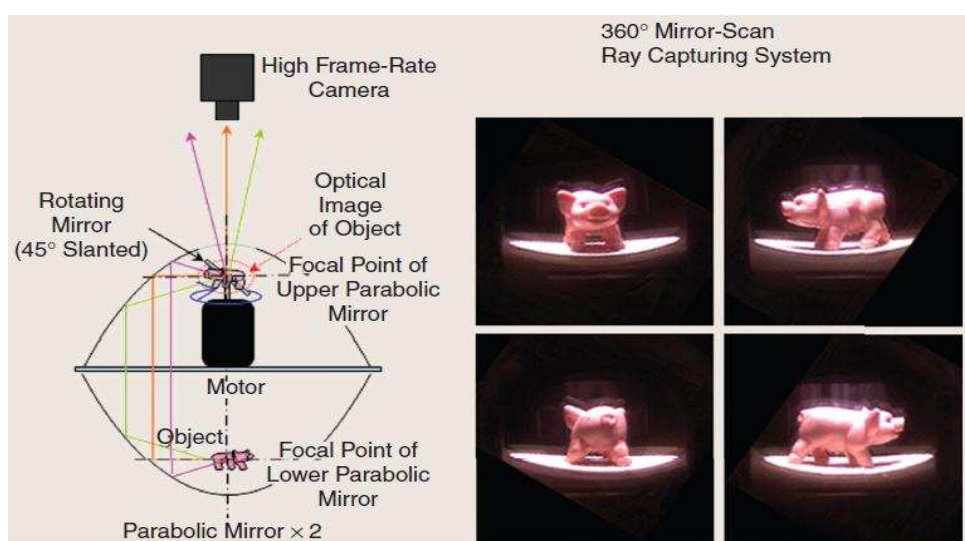


Figura 4. 7 Representación “Space-Ray” de un sistema de 100 vistas

La Figura 4.7 también muestra la calidad de la reconstrucción de la imagen, además de tener prácticamente 360° de vistas donde el usuario no debe cambiar su posición y que recibirá información de alta calidad desde cualquier posición.

4.2.1.6 Pruebas de TV3D móvil.⁶⁴

El estándar chino de televisión Digital DMB, estableció algunos parámetros base sobre los cuales se realizaron pruebas de campo para percibir la calidad y para

⁶³ Fuente: “Free-Viewpoint TV”, Tanimoto M. Mehrdad T., Toshiaki F. y Tomohiro Y.” IEEE SIGNAL PROCESSING MAGAZINE, enero 2011

⁶⁴Fuente: “3DTV Broadcasting and Distribution Systems” NamhoHur, Member, IEEE, Hyun Lee, Member, IEEE, Gwang Soon Lee, Sang Jin Lee, Member, IEEE, AtanasGotchev, Member, IEEE, and Sang-Il Park, Member, IEEE -IEEE TRANSACTIONS ON BROADCASTING, VOL. 57, NO. 2, JUNE 2011 (Las Figuras de esta sección se refieren a la misma fuente)

mejorar el servicio a futuro, el sistema 3DTV se orienta en aquel estándar de acuerdo lo muestra la figura 4.8



Figura 4. 8 Descripción del sistema 3DTV móvil en DMB

Las pruebas de transmisión tomaron en cuenta las consideraciones expuestas en la Figura 4.9

TABLE I
DETAILS OF BROADCASTING TRIALS

	S-DMB (TU Media)
Pilot service	Oct. 2009~Dec. 2009
Trial service	Jan.1 st , 2010~Apr.30 th , 2010
Channel resource	2CDM(256kbps/1CDM*2=512kbps), Channel number:24
Running time per day	8 hours(16:00~24:00)
System structure	Same as 2D system
3D service scenario	3D video service: 320 x 240 @ 15 fps with side-by-side format 3D data service plan: 3D Advertisements during channel switching time
Terminal	Cellular phone type made by Samsung Model name: SCH-W960 LCD : 3.2" WQVGA(240 x 400) AMOLED 3D LCD (parallax barrier)

Figura 4. 9 Parámetros de Prueba de 3DTV sobre DMB ⁶⁴

Estas pruebas fueron muy importantes ya que brindaron la posibilidad para que otros estándares también busquen incorporar el servicio móvil 3D, DVB también incursionó en algunas pruebas teniendo resultados alentadores, como por ejemplo obtener calidad aceptable en DVB-H con 600 Kbps y un FEC de 3/4.

4.2.1.7 Pruebas de percepción de Imagen y experiencia del usuario.

Las técnicas de evaluación subjetiva de imágenes fueron descritas en el Capítulo 1, por lo tanto en esta sección es importante señalar qué tipo de pruebas se han realizado a los usuarios de la TV3D, si bien es cierto las pruebas arrojan gran cantidad de resultados que no se analizarán en este estudio es conveniente señalar algunas características y conclusiones importantes de las experiencias que se tienen de aceptación al usuario; el cual no considera elementos técnicos sino solamente el efecto de comodidad y de calidad percibida.

El estudio “Visual Experience of 3DTV”⁶⁵ tuvo como objetivo comprender, medir y eventualmente modelar y predecir la experiencia Visual en 3D, esta tesis investiga que los modelos de calidad de imagen monoscópica, no son suficientemente adecuados para medir la calidad en 3D, dado que típicamente las distorsiones estereoscópicas y la medida de la profundidad es inexistente para imágenes 2D; parte importante del estudio es encontrar criterios apropiados de evaluación para sistema tridimensionales, estableciendo nuevos criterios de evaluación propios de la TV3D.

La tesis describe un determinado número de experimentos que evalúan diferentes parámetros, por ejemplo en el experimento 1 se evalúa la calidad de imagen, la profundidad, la naturalidad, la presencia y experiencia de visualización.

El estudio también, presenta los trabajos empíricos sobre estos criterios de evaluación para imágenes estereoscópicas, con el uso de diferentes algoritmos de conversión de 2D a 3D.

El Anexo 5 muestra un extracto de esta tesis donde se encuentra la metodología y procedimiento de todos los experimentos realizados.

⁶⁵ “Visual Experience of 3DTV” / by Petrus Johannes Hendrikus Seuntiens - Eindhoven: Technische Universiteit Eindhoven, 2006. - Proefschrift.

4.2.2 EQUIPOS EXISTENTES PARA CONFORMAR EL SISTEMA DE TV3D.

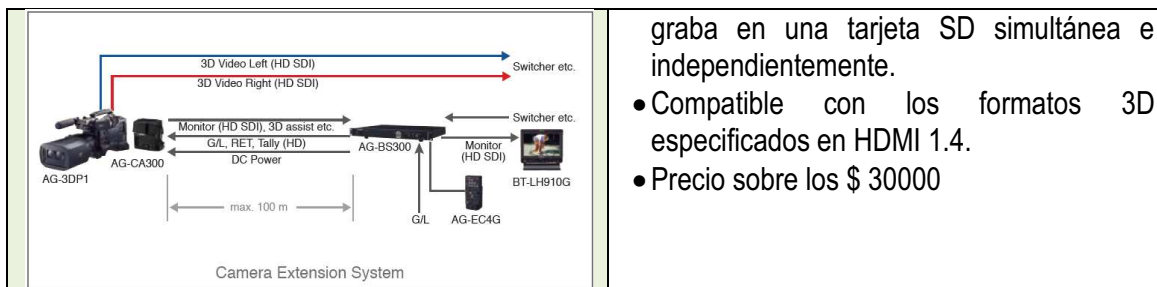
Parte de la situación actual, consiste en describir los últimos productos desarrollados para realizar los distintos procesos que se requieran. Es así que manteniendo el orden del sistema, se detallarán algunos productos resaltando las características más importantes de cada uno.

4.2.2.1 Equipos para la adquisición.

4.2.2.1.1 Cámaras estereoscópicas

Corporaciones tradicionales en la producción de cámaras de video participan muy activamente en el desarrollo de cámaras estereoscópicas como las que se presentan en la Tabla 4.3. La documentación con las especificaciones completas de algunos de los equipos se detalla en el Anexo 6 (Equipos en el mercado).

CÁMARA	CARACTERÍSTICAS
	<ul style="list-style-type: none"> • Codificación H.264 Dual –Stream • Soporte estándar SBS con salida de video vía HDMI. • Procesador para soportar MVC, 1080p30/720p60 • Para trabajar con HDMI 1.4 • Soporta calibración para 3D y alineación del sensor dual. • Flujo de video saliente: Líneas entrelazadas o concatenadas para obtener un único flujo de salida estéreo.
	<ul style="list-style-type: none"> • Marca SONY modelo HXR-NX3D1U • Formato de grabación 3D: MVC, 1080/60i, 50i, 24p. • Soporta HD con MPEG-4 H.264 AVC • Tasa de grabación para 3D: 28 Mbps. • Trabaja con HDMI. • 96 GB de memoria Interna • Precio \$ 3000
	<ul style="list-style-type: none"> • Marca PANASONIC modelo AG-3DP1 • Integra dos sensores tipo 1/3 de 2.2 megapíxeles capaces de grabar simultáneamente dos señales 1920 x 1080 a 10bit 4:2:2. • Emplea MPEG-4 H.264/AVC. • Trabaja con Flujos Independientes de Video (Izq, Der), cada flujo es HD y se



- graba en una tarjeta SD simultánea e independientemente.
- Compatible con los formatos 3D especificados en HDMI 1.4.
 - Precio sobre los \$ 30000

Tabla 4. 3 Modelos de cámaras estereoscópicas disponibles en el mercado.

4.2.2.1.2 Cámaras de profundidad Z-Cam – TOF

Al igual que las cámaras estereoscópicas, las Z-Cam también están disponibles en el mercado y algunas se detallan en la Tabla 4.4 (Información ampliada en el Anexo 6 de Equipos disponibles en el mercado)

CÁMARA	CARACTERÍSTICAS
	<ul style="list-style-type: none"> • PMD[vision]® CamCube 3.0 <ul style="list-style-type: none"> • 200 x 200 pixels • Tasa de Cuadro : 40 fps @ 200x200 pixels, m60 fps @ 176x144 pixels, 80 fps @ 160x120 pixels • Longitud de onda de iluminación: 870 nm
	<ul style="list-style-type: none"> • Frecuencia de Modulación: 29/30/31 Hz. <ul style="list-style-type: none"> • Longitud de Onda 850 nm. • Pixel Pitch 40um. • Precio \$4295
	<ul style="list-style-type: none"> • FOTONIC-B70 by Fotonic

Tabla 4. 4 Modelos de cámaras de profundidad Z-CAM disponibles en el mercado.

4.2.2.1.3 Cámaras para sistema multivista.

El mercado aún no ha explotado comercialmente al máximo estos sistemas, sin embargo lo que hoy se tiene son arreglos de varios lentes y con un posterior uso

de herramientas informáticas para el manejo de tal cantidad información, en todo caso el uso de un codificador de última tecnología puede incorporar el procesamiento requerido por un sistema multivista.

4.2.2.1.4 Equipos para el procesamiento, codificación, transporte y transmisión de la Señal.

La etapa de transmisión comprende varios procesos, en la Tabla 4.5 se describen algunos los que deben ser especiales para TV3D ya que dentro de todo el proceso algunos equipos tradicionales si pueden ser utilizados.



EQUIPO	CARACTERÍSTICAS
Imagen no Disponible	<ul style="list-style-type: none"> • AL8000 MVC/SVC broadcast encoder <ul style="list-style-type: none"> • AVC high profile at level 4.2 • SVC scalable high profile at level 4.2 • MVC stereo high profile at level 4.1 <ul style="list-style-type: none"> • Adaptable a half 3D y a full 3D
	<ul style="list-style-type: none"> • Codificador H.264 / MPEG-4 AVC, Baseline/Main/High, en todos los niveles. • Valido para formatos de cuadro compatible y para algunos V+D.
	<ul style="list-style-type: none"> • Multiplexer NDS3105A marca DSDVB • 12 Flujos Entrantes de hasta 214Mbps por flujo <ul style="list-style-type: none"> • Modos 2K 4K y 8 K para ISDB-T • Dos grupos de salida con codificación RS. <ul style="list-style-type: none"> • Parámetros de Modulación: • Intervalo de Guarda : 1/4 1/8 1/16 1/32 <ul style="list-style-type: none"> • Tasa de Codificación: Coding 1/4 2/3 3/4 5/6 7/8 • Soporta: DQPSK QPSK 16QAM 64QAM <ul style="list-style-type: none"> • Anchos de Banda de 6 7 y 8 Mhz

Tabla 4. 5 Equipos de transmisión disponibles en el mercado.

4.2.2.1.5 Equipos para la recepción de la señal.

A pesar de que en investigación se han desarrollado Set-Top-Box y chips completos para los formatos 3D más exigentes en el mercado, aún no se los observa. En la Tabla 4.6 se detallan chips para Set-Top-Box y Televisores para recibir la señal de video estereoscópica. A pesar de que fabricantes como Sony y

Toshiba ya han presentado prototipos de televisores Autoestereoscópicos⁶⁶ aún el mercado no permite que estos sean comercializados masivamente.

EQUIPO	CARACTERÍSTICAS
	<ul style="list-style-type: none"> • Sigma SMP8910 <ul style="list-style-type: none"> • Adaptable a Blue-Ray, IPTV, Set-Top-Boxes • Soporta formatos compatibles con el Cuadro HD, V+D, Dual Streaming. • MPEG -4.10 (H.264) BP@L3, MP@L4.2, HP@L4.2, MVC <ul style="list-style-type: none"> • Capaz de realizar conversiones 3D en tiempo real <ul style="list-style-type: none"> • Varios formatos de salida de video.
<p style="font-size: 24pt; color: #0070C0; text-align: center;">BCM7424</p>	<ul style="list-style-type: none"> • Set-Top-Box para 3DTV satelital <ul style="list-style-type: none"> • Soporta AVC, SVC, MVC
	<ul style="list-style-type: none"> • Semisky (Shenzhen) Electronics Co., Ltd -ISDB-T HD TV Set Top Box • Accord with ISDB-T Standard, in suport of MPEG2, MPEG4 and H.264 Decoder. • In support of H.264 Video Decoder, the video output formats include CVBS. YPbPr, and HDMI. Para formatos compatibles con el cuadro HD
	<ul style="list-style-type: none"> • Visor personal 3D SONY <ul style="list-style-type: none"> • Precio \$800 • Video Proveniente de algún dispositivo HDMI • Tecnología OLED de pantallas de alta resolución
	<ul style="list-style-type: none"> • LG 42LW5700 Cinema 3D • Tecnología 3D: Soportada, Cinema 3D con TruMotion 120 Hz <ul style="list-style-type: none"> • Soportada, norma ISDB-T • FPR (Film-type Patterned Retarder), polarizado
	<ul style="list-style-type: none"> • Sony BRAVIA® HX729 Series LED HDTV • Sistema de Visualización con gafas activas • Requiere de contenido 3D a través de HDMI
	<ul style="list-style-type: none"> • 5 gafas pasivas (polarizadas) marca LG <ul style="list-style-type: none"> • Precio por las 5 gafas \$39,53
	<ul style="list-style-type: none"> • PANASONIC TY-EW3D10E Gafas 3D Gafas Activas <ul style="list-style-type: none"> • Precio por unidad: \$125,17

Tabla 4. 6 Equipos de recepción 3DTV disponibles en el mercado.

⁶⁶Televisores presentados en el Consumer Electronics Show (CES), Las Vegas USA, Enero 2011

4.2.3 PERSPECTIVAS DE NORMALIZACIÓN, EVOLUCIÓN DE LOS ELEMENTOS TÉCNICOS DEL SISTEMA 3DTV.⁶⁷

Se pueden identificar claramente las fases del servicio de televisión estereoscópica considerando como fundamentales aquellos niveles que pueden funcionar sobre las plataformas actuales. Con todos los elementos hasta ahora estudiados se puede concluir que la evolución de la tecnología permite identificar tres fases de la TV3D (ver Figura 4.10).



Figura 4. 10 Generaciones del servicio 3DTV

Si hoy en día nos situamos en la 1ª generación, y con los elementos técnicos y equipos disponibles, entonces la solución óptima y disponible para el servicio de mejor calidad posible viene dada por los formatos compatibles con el cuadro HD en cualquiera de sus configuraciones SbS, TaB, etc. Aunque hoy por hoy sea solo posible el HALF-3DTV, en el futuro se podrá duplicar el tamaño del cuadro y llegar al FULL-3DTV.

Con los equipos disponibles y las técnicas utilizadas convencionalmente, se hace imprescindible desarrollar nuevo Hardware para las futuras generaciones, para que así se produzca la evolución del sistema (Ver Figura 4.11).

⁶⁷Fuente: "Radiodifusión de la 2ª Generación de 3DTV" Jorge Rodríguez López, Grupo Hispasat, Julio 2011.

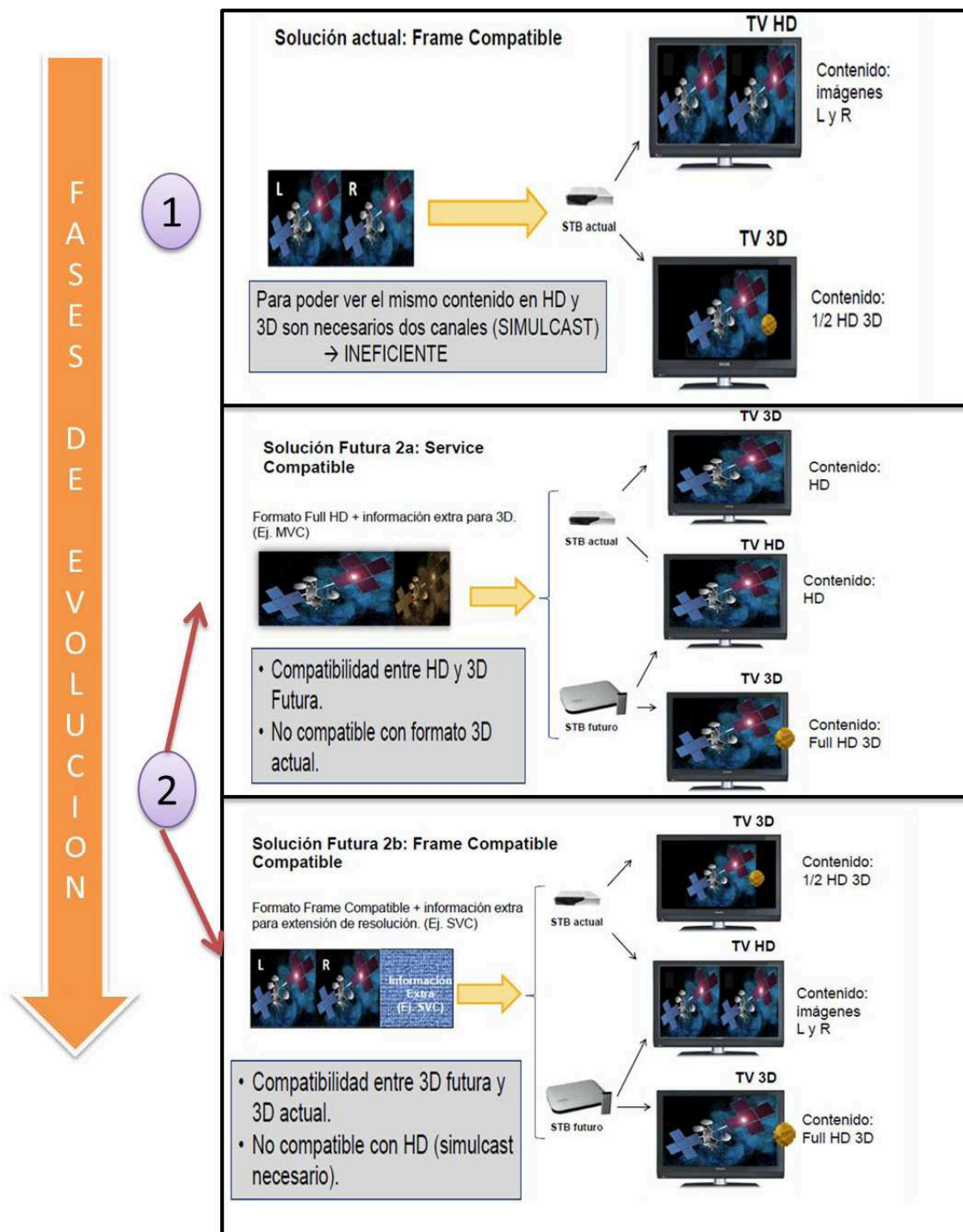


Figura 4. 11 Fase de evolución de la 3DTV

En cuanto a las herramientas de codificación, estas también han seguido una evolución, partiendo de la adaptación de las actualmente conocidas hacia las desarrolladas recientemente, las cuales brindan mayores prestaciones y mejor calidad. Comenzando por H.264/AVC, se puede decir que si bien representa el primer paso en la codificación 3D, esta herramienta es, de las aplicables y

disponibles en el mercado, la más reciente, después debemos mencionar a las extensiones SVC y MVC de H.264 que representan la evolución hacia un estándar propio de codificación para video estereoscópico. Tanto SVC como MVC tienen sus fortalezas y debilidades, por ejemplo se ha demostrado que SVC permite mejorar las tasas sin reducir drásticamente la calidad y que MVC representa la herramienta óptima para explotar la relación entre las vistas. La ITU junto con el grupo MPEG estudian la posibilidad de evolucionar la codificación y llevarla al siguiente nivel, H.265, con el que se pretende tener un ahorro del 50% de ancho de banda que el de MPEG-4 obteniendo la misma calidad. La fase de desarrollo actual plantea un previsto lanzamiento al mercado en 2014. H.265 pretende ser la base para la codificación futura del video 3D.

Analizando la situación técnica actual, se puede establecer las fortalezas y debilidades de cada esquema de codificación existente (Ver Figura 4.12)

	Codificación para formatos compatibles HD – H.264/AVC	Codificación H.264/SVC	Codificación H.264/MVC
CALIDAD DE LA IMAGEN	3D 1/2 HD 	3D Full HD 	3D Full HD 
SERVICIOS COMPATIBLES		 	
SOPORTA MULTIVISTA			
DISPONIBLE COMERCIALMENTE			

Figura 4. 12 Comparación de los esquemas de codificación H.264

Otro elemento técnico importante en la evolución del servicio es el número de vistas, ya sean estas adquiridas, interpoladas o creadas, se supone que a mayor número de vistas, y mayor número de profundidades la imagen en reproducción será de mejor calidad es por ello al crecer este número de vistas también crece el

nivel de servicio brindado y factible de implementar, tal y como lo muestra la Figura 4.13

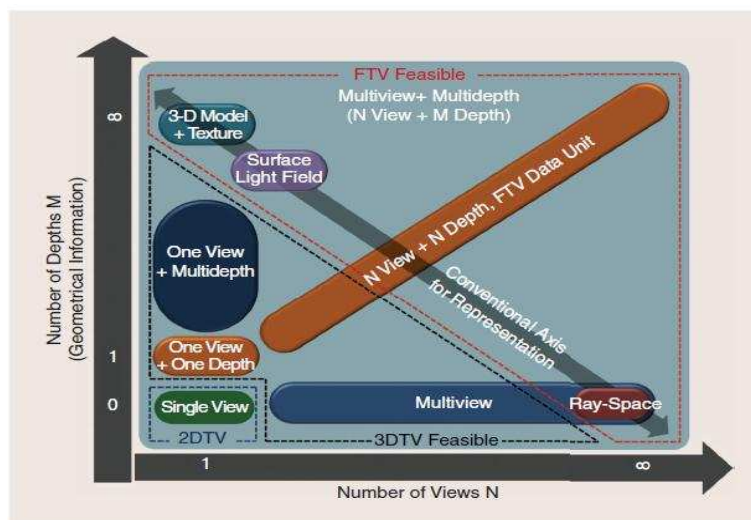


Figura 4. 13 Relación entre el número de vistas y el servicio brindado.

En la Figura 4.13 se diferencian claramente las etapas existentes 2DTV, 3DTV, y FTV (FreeViewPoint TV) determinando en cada una el número de vistas que hacen a cada servicio factible. Como último punto cabe mencionar que el estándar HDMI 1.4a o posterior, es el que se debe seguir utilizando en los equipos ya que con un estándar anterior no existe el soporte de servicios 3D hoy compatibles. Por lo cual se concluye que es parte del éxito del sistema el avance de esta interfaz capaz de soportar la exigencia del servicio de TV3D.

4.3 SITUACIÓN DE MERCADO DEL SISTEMA Y SUS TENDENCIAS⁶⁸.

En esta sección corresponde estudiar los elementos que sin ser meramente técnicos facilitan o dificultan el desarrollo del servicio del TV3D. Para ello las experiencias de operadoras de televisión digital Europea, pueden ser de gran utilidad para evaluar estos aspectos y poder establecer algunas tendencias futuras.

⁶⁸ Datos para la situación de mercado obtenido de : “HD y 3D Situación y tendencias” ASTRA, José Luis Garate, Madrid, Octubre 2010.

El servicio de 3D, junto con la alta definición, pretenden ser la siguiente evolución en la televisión, y hoy en día algunas plataformas brindan estos servicios audiovisuales avanzados por televisión.

Se pueden mencionar dos factores que impiden un desarrollo rápido del servicio:

- Altos precios en el equipamiento de recepción.
- La falta de contenidos.

El desarrollo tecnológico no escasea, es por ello que los pocos operadores existentes pueden escoger la técnica; hoy en día existe un estándar de “facto” para la TV3D, y es el compatible con cuadro HD, el cual es el más viable actualmente.

Los llamados a superar los problemas antes mencionados son las empresas operadoras, y es mandatorio que lo hagan, ya que se ha demostrado que el servicio de 3D promete ser el más atractivo al público, muestra de ello es que las películas 3D pueden en algunos casos duplicar las ganancias de lo que se obtiene por 2D. En el año 2009 la película Bolt recaudó más del doble en taquilla en la versión 3D que en la versión 2D. Otra variable que permite asegurar que la TV3D tendrá acogida, es el número de películas 3D producidas en los últimos años, mientras en 2005 se produjeron 2 películas 3D y de animación, en el 2011 se produjeron alrededor de 50 películas en total, siendo casi 30 las de acción en vivo grabadas con cámaras estereoscópicas y 20 las de animación creadas por computadora; en la actualidad el número de películas superan las 70, mayoritariamente las de acción en vivo. Los contenidos más apetecibles para producir 3D, según las plataformas que están en el mercado, son los eventos deportivos, el 80% de los contenidos 3D son deportes mientras que el 10% o menos está destinado a películas⁶⁹.

Hoy en día la Televisión por satélite lidera la comercialización de canales 3D entre ellos: ESPN, Sky, Comcast, Directv, Digital+ etc. en países como E.E.U.U, Inglaterra, España, Japón, y próximamente en Brasil, y Australia.

⁶⁹ Fuente: “HD y 3D Situación y tendencias” ASTRA, José Luis Garate, Madrid, Octubre 2010.

Por otro lado en una realidad más cercana, Chile ya ha realizado pruebas de emisiones 3D en el estándar ISDB-T con su canal TVN junto con Direc-TV en el año 2010, mientras que China ya dispone de un canal público que emite 4,5 horas de programas 3D cada día, que son repetidos dos veces, en total son 13 horas y media de emisión al día, desde las 10:30 de la mañana a la medianoche. Todos estos sucesos apuntan a una evolución sostenida del servicio.

En cuanto a la situación del mercado y del comportamiento del consumidor hasta hoy, los fabricantes han puesto a disposición del cliente algunos productos que permiten analizar la reacción de consumo del mismo (ver Tabla 4.6).







	 LG	 Panasonic	 SAMSUNG	 SONY	 TOSHIBA	 PHILIPS
Tipo de Pantalla	LCD	Plasma	•Plasma •LCD	LCD	LCD	LCD
Tecnología de gafas	*G. Polarizadas *G. Activas (LCS)	G. Activas (LCS)	G. Activas (LCS)	G. Activas (LCS)	G. Activas (LCS)	G. Activas (LCS)
3DTV's 2010 (numero de modelos)	12	9	43	12	3	8
3D BR Players	1	5	1	3	1	---
Tamaños de la pantalla	32''- 55''	50''- 65''	40''- 67''	40''- 60''	55''- 65''	37''- 58''
Resolución Max. 3D	1080 i/p	1080p	1080p	1080p	1080p	1080p
Conversión 2 D a 3D	No	No	Si	Si	Si	?
Precio aprox.	\$2500(47'')	\$3500(50'')	\$3200(46'')	\$3500(46'')	---	\$3500(56'')

Tabla 4. 7 Oferta de Televisores 3D en el mercado⁷⁰

En Europa, en los primeros 4 meses de 2010 se vendieron 100.000 televisores 3D, lo que habla de gran aceptación, la evolución de dicho avance se muestra en la figura 4.14

⁷⁰ Fuente: "HD y 3D Situación y tendencias" ASTRA, José Luis Garate, Madrid, Octubre 2010.

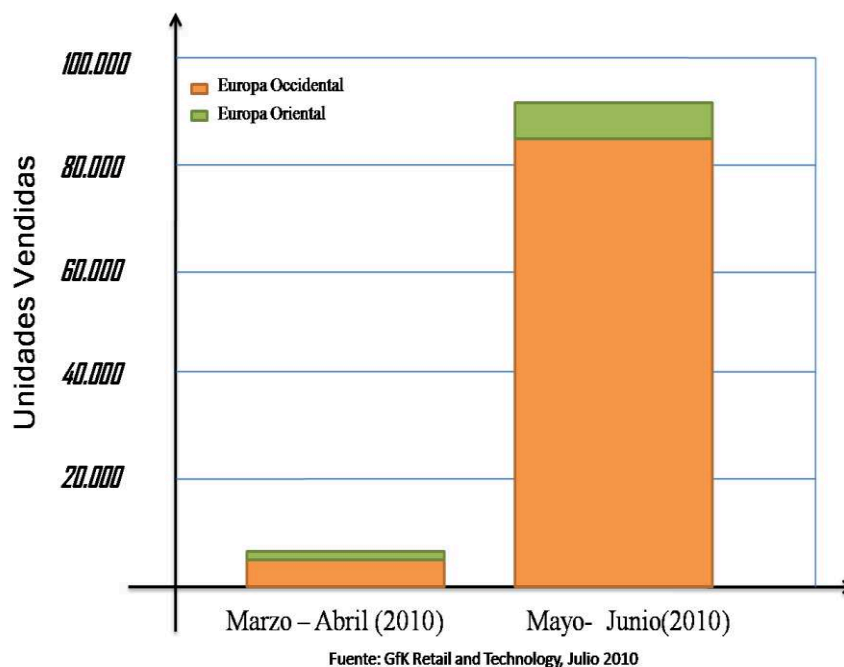


Figura 4. 14 Unidades de Televisores 3D vendidas en 2010

Dentro de otros datos interesantes del mismo estudio, se afirma que la penetración de la 3DTV en el hogar se dará de manera fuerte entre 2010 y 2015.

Mientras en 2010 la penetración era del 1% aproximadamente, para el 2014 se espera un crecimiento al 25% y que seguirá al alza, de tal forma que el 50% de los televisores vendidos serán aptos para 3D (en Europa), presentándose el 3D como la gran innovación en el mercado.

Por otro lado se prevé también que para 2015 habrá 25 millones de hogares con TV3D de los cuales 7 millones utilizarán de manera activa y frecuentemente el servicio.

La tendencia creciente del servicio, que se piensa ocurre de 2010 a 2015, se muestra en la Figura 4.15; para años posteriores la intención es que la tendencia a la alza sea más brusca y más rápida.

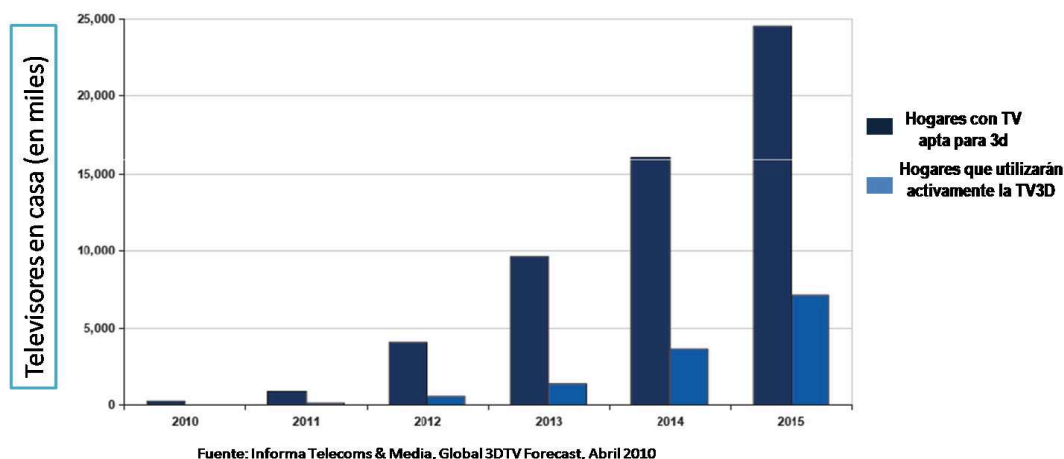


Figura 4. 15 Tendencia de consumo de televisores 3D.

Si bien es cierto esta información se enmarca en un contexto europeo, la tendencia siempre será creciente en el mundo entero; la tecnología no se detiene y es uno de los aspectos que avanza a pasos agigantados en cualquier parte del planeta, ya sea en Europa o América.

4.4 RELACIÓN CON OTRAS TECNOLOGÍAS Y PERSPECTIVAS FUTURAS.

La tecnología de video estereoscópico no avanza sola, involucra una serie de elementos complementarios que también tienen un desarrollo sostenido, para esto cabe señalar la evolución del sistema de televisión en cuanto a los servicios brindados a lo largo de la historia (ver Figura 4.16).

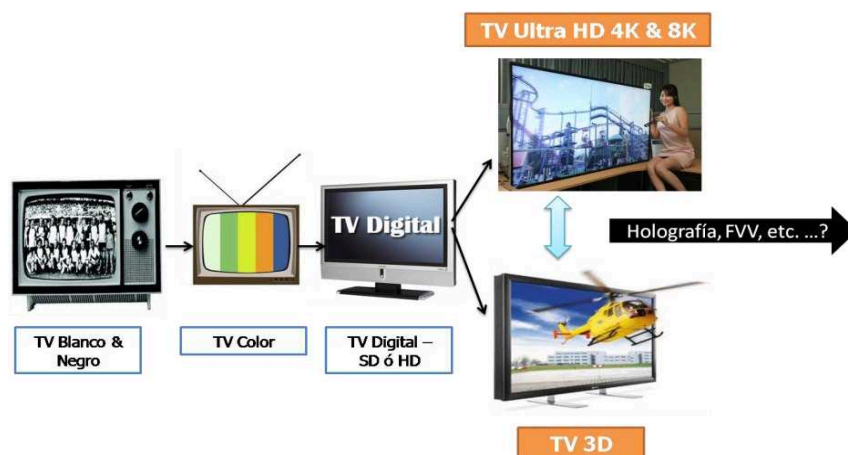


Figura 4. 16 Servicios brindados por la TV a lo largo de la Historia

La Figura 4.16 muestra la clara relación entre el HD y el 3D, dado que el HD brinda la primera posibilidad de hacer 3D sobre uno de sus cuadros y así mismo, en el caso de que HD llegase a la evolución UHD, se tendrán beneficios también para señales estereoscópicas, por otro lado existen otros servicios que también se desarrollaran afines a la TV3D, entre ellos se encuentran:

- FVV - Televisión multivista.
- Televisión Holográfica
- Tecnología de almacenamiento de información
- 3D sobre IP, etc.

En el primer caso la FVV, de llegar a un sistema ya maduro, sería de mucha acogida para el usuario que perciba cientos de vistas de la escena desde cualquier ángulo o punto de vista; la evolución sigue y sin duda el poder hacer hologramas para el público también promete ser muy atractivo.

Con la cantidad de información que genera el video de alta calidad en UHD, el video 3D y el video con múltiples vistas, la tecnología de almacenamiento, discos duros y medios de transferencia de información sin compresión requieren gran desarrollo. Además dado que el Internet no deja de ser muy utilizado y popular, la TV3D puede estimular el uso del video 3D sobre IP, haciendo que se complemente la televisión y el Internet.

A futuro se prevé que se vayan escogiendo las técnicas y formatos de acuerdo al costo y facilidad de implementación, no obstante el desarrollo en investigación, avance y mejore, se debe ir de acuerdo al mercado y a la acogida del usuario.

Hoy en día existen las herramientas para hacer que el usuario pueda sentir la realidad visual en la TV tal y como es la vida diaria, sin embargo todavía cuesta que el 3D de primera generación penetre totalmente en el mercado y en los hogares; se espera que así como todos los aspectos tecnológicos, la TV3D avance y sea la tecnología tan esperada, prometedora y atractiva como hoy se cree y piensa.

CAPÍTULO 5

CONCLUSIONES Y RECOMENDACIONES

5.1 CONCLUSIONES

- La estereoscopía ha tenido un desarrollo sostenido y extenso a lo largo de la historia, por lo cual es muy necesario que se aplique acorde a la tecnología existente en todos los campos, especialmente en el servicio de Televisión.
- La Televisión Digital es un gran avance tecnológico, sin embargo este servicio debe competir con tecnologías como IP-TV e Internet; por lo cual el brindar nuevas prestaciones (3DTV) al usuario, harán que la TV sea atractiva y marque la línea de desarrollo para otras tecnologías.
- El hecho de adoptar un estándar de TDT en nuestro país debe suponer una evolución total y en todos los aspectos del servicio, por lo tanto no bastará con alta calidad y más programación, será muy necesario la inclusión de servicios muy novedosos e interactivos para el usuario.
- Todas las técnicas de codificación actuales para imágenes estereoscópicas se fundamentan en la norma H.264 / MPEG-4, dado que este esquema de codificación está vigente para ISDB-T, se tiene un gran punto a favor para poder implementar el servicio.
- Las técnicas de adquisición proponen variadas soluciones para múltiples vistas, en el corto plazo lo más aplicable son los métodos basados en 2 vistas (L+R), por lo tanto estos métodos son los más apropiados para iniciar pruebas en etapas de adquisición y de transmisión.
- Las técnicas que usan el mapa de profundidad como información importante, presentan un ahorro en tasas de bits; sin embargo merecen una infraestructura de adquisición de imágenes exclusiva y propia de los formatos V+D, que son más adaptables a los estándares de codificación

menos sofisticados, por lo tanto representan una opción conveniente en la etapa de transmisión.

- La calidad en la adquisición depende proporcionalmente del número de vistas adquiridas, es por ello que para un sistema multivista la calidad en reproducción será máxima; no obstante las dificultades en los requerimientos del canal y el costo de los equipos aumenta significativamente. La demanda de recursos técnicos en estos casos hace que lo más recomendable sea utilizar interpolación o creación de vistas no adquiridas y para evitar el uso excesivo del canal, esta creación se debería hacer en recepción, lo cual por ahora supone IRD's bastante complejos.
- El estándar H.264 es muy robusto en cuanto al tipo de información que puede procesar, los anexos SVC y MVC presentan una gran herramienta para el desarrollo de sistemas con gran cantidad de información; la 3DTV para ISDB-T resulta completamente aplicable desde este parámetro técnico y versátil al poder utilizar diferentes anexos de la norma dependiendo del formato de producción 3D elegido.
- La interfaz HDMI 1.4a establece la mejor solución para la comunicación de las señales 3D en recepción, ya que se encuentra estandarizada y muestra grandes prestaciones, además de contar con la aceptación del mercado.
- Es necesario reglamentar el uso de los 13 segmentos, para que así se establezca un proceso de aplicación de señales estereoscópicas a uno o más canales y que estos convivan tanto con señales HD, SD y one-seg.
- Dado que una innovación importante de ISDB-T es el one-seg, se deben utilizar las herramientas para aplicar 3DTV móvil, ya que para imágenes de baja resolución la demanda de recursos en 3D no es tan crítica como en resoluciones mayores.
- Con el ahorro de espectro radio eléctrico que implica la TDT, no es crítico que servicios como el 3D puedan aplicarse con técnicas como Simulcast o

Dual Streaming, ya que a pesar de duplicar los recursos que se utilizarían normalmente, la capacidad disponible es suficiente.

- El desarrollo de la HDTV va íntimamente ligado al desarrollo de la 3DTV, ya que la primera generación del servicio se fundamenta en la utilización del cuadro de alta definición.
- La decisión sobre qué contenidos deben transmitirse en canales 3D, debe estar ligada a la aceptación de los consumidores, por lo tanto es necesario tomar en cuenta los estudio previos que determinan la clase de eventos que los espectadores desean y esperan ver por televisión 3D.
- El problema de la incomodidad visual debe ser abordado en todas las etapas: en la adquisición, calibrando mejor las cámaras y utilizando mejores métodos de captación; en la etapa de pre-procesamiento ajustando el paralaje y la disparidad horizontal, en la codificación evitando pérdida de información de profundidad; en recepción efectuando una correcta reconstrucción de la señal visual y por último al utilizar pantallas de alta calidad con métodos polarizados o multiplexados en tiempo que generen la menor fatiga visual posible.
- Dentro de la reproducción sobre pantallas estereoscópicas, autoestereoscópicas, volumétricas u holográficas, se requieren amplios estudios y experimentos de evaluación subjetiva para determinar la técnica de mayor aceptación.
- Por el momento la solución económicamente más conveniente en la reproducción de la señal, es utilizar gafas polarizadas o LCS. Las gafas 3D activas (LCS) tienen mejores resultados en la resolución, ya que muestran imágenes de alta resolución en calidad Full-HD y menor distorsión de la imagen; por otro lado las polarizadas permiten obtener mayor luminosidad al no existir un parpadeo de la imagen a cada ojo; sin embargo en ambos casos la solución mucho más barata implica utilizar gafas polarizadas (pasivas).

- Aplicar el servicio es técnicamente posible dentro de ISDB-T, adoptándose en etapas, comenzando con los formatos de primera generación y avanzar a futuro hasta lograr sistemas multivista y autoestereoscópicos al usuario.
- Los proyectos analizados muestran que es necesario encontrar un sistema end-to-end aplicable e implementable, asimismo se concluye que es necesario encontrar técnicas comunes en los procesos de adquisición y recepción, y también desarrollar técnicas cómodas a la experiencia visual del usuario
- El presente estudio ha abordado en mayor o menor escala todos los aspectos que comprenden un sistema de TV3D, concluyendo que el sistema es viable técnicamente, y que la transición a la TV digital debe incluir este servicio como parte de las innovaciones de la digitalización.

5.2 RECOMENDACIONES

- Del estudio realizado se desprenden campos donde es necesario desarrollar investigación y generar aplicaciones propias del estándar, por lo que se recomienda poner énfasis en el desarrollo de los temas relacionados a la TV3D.
- Dadas las posibilidades de formatos se recomienda profundizar el estudio en particular de cada uno de ellos, V+D, formatos compatibles con el cuadro HD, SVC, etc.
- Se recomienda efectuar pruebas de campo y de calidad subjetiva para determinar el formato óptimo para anexar a la norma ISDB-T y establecer una línea base de la evolución del servicio, tomando en cuenta las generaciones estudiadas.
- Es necesario analizar métodos de empaquetamiento y/o transmisión para dar una idea de una medida de estandarización.

- Se recomienda realizar exhaustivamente, pruebas de mercado, costo de equipos, aceptación del público para poder emprender en el desarrollo experimental de la alternativa más conveniente y más económica.
- Dado que la TV satelital ha profundizado en mayor medida el mayor desarrollo de la aplicación TV3D, se recomienda estudiar los proyectos que se han lanzado, y recoger las técnicas y resultados favorables para adaptarlos a las normas de televisión digital terrestre.
- Dado que las pruebas de recepción se han realizado sobre Set-top-Box modificados, se recomienda analizar e investigar métodos para actualizar estos receptores para formatos 3D como DIBR, TaB, SbS, etc.
- Se recomienda desarrollar aplicaciones y metodologías para el middleware Ginga, con el objetivo de que el usuario interactúe con la señal de video 3D, pueda manipularla y elegir el formato de acuerdo a la pantalla o a su conveniencia.
- Es necesario establecer el grado de aceptación del servicio por parte de los consumidores en la sociedad, por lo tanto se recomienda realizar estudios de mercado considerando aspectos sociales y económicos de la realidad latinoamericana.
- En general para la TDT, es muy importante reglamentar el uso de los 13 segmentos del ISDB-T, por lo tanto se recomienda considerar en la reglamentación, el futuro uso de canales 3D; y de señales que eventualmente, requieran más segmentos de lo habitual, para generar servicios más exigentes.

REFERENCIAS BIBLIOGRÁFICAS

ARTÍCULOS Y PUBLICACIONES CIENTÍFICAS

- AM Ariyaeinia; “Analysis and design of stereoscopic television systems”; University of Hertfordshire, College Lane, Hatfield, Hertfordshire AL10 9AB, UK; 1998.
- BAL Can, “Three-dimensional video coding on mobile platforms”, a thesis submitted to the department of electrical and electronics engineering and the institute of engineering and sciences of Bilkent university in partial fulfillment of the requirements for the degree of Master of science, Septiembre 2009.
- BALAŠKO Hrvoje, “Comparison of Compression Algorithms for High Definition and Super High Definition Video Signals”; Audio Video Consulting Ltd., Karlovačka 36b, 10020 Zagreb, Croatia; 2010.
- BOISSON Guillaume, KERBIRIOU Paul y LOPEZ Patrick; THOMSON Research Digest; “Encoding contents for 3DTV: binocular stereo, 2D+Z and LDV”; IEEE International Conference on Image Processing (ICIP); 2009.
- BONANSEA Lucas; “3D Hand gesture recognition using a ZCam and an SVM-SMO classifier”; Tesis Iowa State University Ames; Iowa; 2009.
- BOURGE Arnaud, GOBERT Jean y BRULS Fons, “MPEG-C Part 3: Enabling the introduction of video plus depth contents”, Philips Applied Technologie (Suresnes, France), Philips Research (Eindhoven, The Netherlands), 2006.
- C. FEHN y P. KAUFF (Heinrich-Hertz-Institut, Germany), M. Op de BEECK y F. ERNST (Philips Research, The Netherlands), W. IJSSELSTEIJN (Eindhoven University of Technology, The Netherlands), M. POLLEFEYS y L. VAN GOOL (Katholieke Universiteit Leuven, Belgium), E. OFEK (3DV Systems, Israel) e I. SEXTO (De Montfort University, United Kingdom); “AN EVOLUTIONARY AND OPTIMISED APPROACH ON 3D-TV”, 2006
- CHUANG Tzu-Der, TSUNG Pei-Kuei, LIN Pin-Chih, CHANG Lo-Mei, MA Tsung-Chuan, CHEN Yi-Hau, CHEN Yu-Han, TSAI Chuan-Yung, y CHEN Liang-Gee; “A 59.5mW Scalable/Multi-view Video Decoder Chip for

Quad/3D Full HDTV and Video Streaming Applications “;DSP/IC Design Lab, Graduate Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan; 2010.

- Digital Video Broadcasting (DVB); “Frame Compatible Plano-Stereoscopic 3DTV (DVB-3DTV)” ; DVB Document A154; Febrero 2011.
- DONGHYUN Kim, DONGBO Min, KWANGHOON Sohn, “A Stereoscopic Video Generation Method Using Stereoscopic Display Characterization and Motion Analysis”, IEEE transactions on broadcasting, Vol 54, N°2, Junio 2008.
- FEHN Christoph; “Depth-Image-Based Rendering (DIBR), Compression and Transmission for a New Approach on 3D-TV” Fraunhofer-Institut für Nachrichtentechnik, Heinrich-Hertz-Institut (HHI); Einsteinufer 37, 10587 Berlin, Germany; 2004.
- FEHN Cristopher; “A 3D-TV System Based On Video Plus Depth Information”; Fraunhofer-Institut für Nachrichtentechnik, Heinrich-Hertz Institut; 2009.
- GARATE Jose Luis, “HD y 3D Situación y Tendencias”, Grupo ASTRA, Madrid, Octubre 2010
- HILTON Adrian, GUILLEMAUT Jean-Yves, KILNER Joe, GRAU Oliver, y GRAHAM Thomas; “3D-TV Production From Conventional Cameras for Sports Broadcast”; IEEE Transactions on broadcasting, vol. 57, no. 2, Junio 2011.
- HITACHI, Ltd., PANASONIC Corporation., PHILIPS Consumer Electronics, INTERNATIONAL B.V., SILICON Image, Inc., SONY Corporation, TECHNICOLOR, S.A., TOSHIBA Corporation; “High-Definition Multimedia Interface Specification Version 1.4a Extraction of 3D Signaling Portion”; 4 Marzo; 2010.
- KAWAKITA M., KURITA T., KIKUCHI H., INOUE S.; “HDTV AXI-VISION Camera”; NHK Science & Technical Research Laboratories, Japan; 2002.
- KERBIRIOU Paul, BOISSON Guillaume, SIDIBÉ Korian, HUYNH-THU Quan, “Depth-based representations: which coding format for 3D Video broadcast applications?; Technicolor, 3D4YOU Project, 17 Marzo 2011.

- LAKIS Christodoulou, LIAM M. Mayron, HARI Kalva, OGE Marques, y BORKO Furht; “3D TV Using MPEG-2 and H.264 View Coding and Autostereoscopic Displays”; Dept. of Computer Science and Engineering, Florida Atlantic University, Boca Raton, FL 33431, Octubre 2006.
- LEE Jin Young, WEY Hochen, y PARK Du-Sik; “A Fast and Efficient Multi-View Depth Image Coding Method Based on Temporal and Inter-View Correlations of Texture Images”; IEEE; 2011.
- LI Ping, FARIN Dirk, GUNNEWIEK Klein Rene, DE WITH Peter H. N; “On Creating Depth Maps from Monoscopic Video using Structure From Motion”; Eindhoven University of Technology, Philips Research Eindhoven, Eindhoven University of Technology and Logica CMG Netherlands B.V.; 2006.
- LUTHRA Ajay, “MPEG-4 MVC/H.264 Digital Video Compression Standard Digital”, Adv. Tech., Motorola Inc. – ITU-T, ITU-T Workshop and IMTC Forum, Mayo 2006.
- MEESTERS Lydia M. J., IJSSELSTEIJN Wijnand A., y SEUNTIËNS Pieter J. H.; “A Survey of Perceptual Evaluations and Requirements of Three-Dimensional TV”; IEEE transactions on circuits and systems for video technology, VOL. 14, NO. 3, Marzo 2004.
- MÜLLER Karsten, “Representation and Coding Formats, for 3D Video”; 3DTV Conference IEEE, IHH; Alemania; 2010.
- NAMHO Hur, HYUN Lee, GWANG Soon Lee, SANG Jin Lee, ATANAS Gotchev, y SANG-II Park; “3DTV Broadcasting and Distribution Systems”; IEEE transactions on broadcasting, VOL. 57, NO. 2, Junio 2011.
- OHM Jens-Rainer; “MPEG Developments in Multi-view Video Coding and 3D Video”; RWTH Aachen University Lehrstuhl und Institut für Nachrichtentechnik; Alemania 2009.
- PEI-KUEI Tsung, LI-FU Ding, WEI-YIN Chen, TZU-DER Chuang, YU-HAN Chen, PAI-HENG Hsiao, SHAO-YI Chien, y LIANG-GEE Chen; “Video Encoder Design for High-Definition 3D Video Communication Systems”; IEEE Communications Magazine ; Abril 2010.
- Puri, R.V. Kollarits, B.G. Haskell; “Basics of stereoscopic video, new compression results with MPEG-2 and a proposal for MPEG-4”; AT&T

- Labs - Research, IO1 Crawfords Corner Road, Holmdel, NJ 07733, USA; 1997.
- ROHDE & SCHWARZ; “3D TV test signals in line with the HDMI 1.4^a interface standard”; Broadcasting/Test & Measurement Application Brochure; München Alemania; 2011.
 - RUBÉN Fernández, ROMÁN Antonio, “Adquisición y visualización de vídeo 3D” Tesis Ing. De Telecomunicaciones, Universidad Politécnica de Cataluña, Noviembre 2010.
 - SATU Jumisko-Pyykkö y TIMO Utriainen; “A Hybrid Method for Quality Evaluation in the Context of Use for Mobile (3D) Television”; Human-Centered Technology, Tampere University of Technology, Septiembre 2010.
 - SMOLIĆ Aljoscha Dr.; “Compression for 3DTV - with special focus on MPEG standards”; Fraunhofer HHI Image Processing Department; Alemania, 2007.
 - SMOLIC Aljoscha, MUELLER Karsten, MERKLE Philipp, VETRO Anthony; “Development Of A New MPEG Standards For Advanced 3D Video Applications” Mitsubishi electric research laboratories (<http://www.merl.com>); Noviembre 2009.
 - SMOLIC Aljoscha; “3D Video and Freem Viewpoint Video”; Multimedia Communications, Disney research; Zurich; 2005.
 - TAM Wa James, SPERANZA Filippo, YANO Sumio, SHIMONO Koichi, ONO y Hiroshi; “Stereoscopic 3D-TV: Visual Comfort”; IEEE Transactions on broadcasting, VOL. 57, NO. 2, Junio 2011.
 - TANIMOTO Masayuki, TEHRANI Mehrdad PANAHPOUR, FUJII Toshiaki, y YENDO Tomohiro; “ Free Viewpoint TV”; IEEE Signal processing magazine N°67, Enero 2011.
 - TSUNG Pei-Kuei, LIN Ping-Chih, CHEN Kuan-Yu, CHUANG Tzu-Der, YANG Hsin-Jung, CHIEN Shao-Yi, DING Li-Fu, CHEN Wei-Yin, CHENG Chih-Chi, CHEN Tung-Chien, CHEN Liang-Gee; “A 216fps 4096×2160p 3DTV Set-Top Box SoC for Free-Viewpoint 3DTV Applications”; National Taiwan University; IEEE International Solid-State Circuits Conference; Taipei, Taiwán; 2011

- VETRO Anthony, TOURAPIS Alexis M., MÜLLER Karsten, y CHEN Tao “3D-TV Content Storage and Transmission”, IEEE Transactions on Broadcasting -- Special Issue on 3D-TV Horizon: Contents, Systems and Visual Perception, Octubre 2010.
- VETRO Anthony, YEA Sehoon, SMOLIC Aljoscha, “Towards a 3D Video Format for Auto-Stereoscopic Displays”, Mitsubishi Electric Research Laboratories, Septiembre 2008.
- WOJCIECH Matusik, HANSPETER Pfister, “3D TV: A Scalable System for Real-Time Acquisition, Transmission, and Autostereoscopic Display of Dynamic Scenes”, Mitsubishi Electric Research Laboratories, Cambridge, MA, 2004.
- WOODS Andrew, HELLIWELL Jesse; “White Paper: A Survey of 3D Sync IR Protocols”; Centre for marine Science and Technology Curtin University Perth, Western Australia, 31 Marzo 2011.
- ZHI-WEI Gao, WEN-KUO Lin, YU-SHIAN Shen, CHIA-YEN Lin, y WEN-CHUNG Kao; “Design of Signal Processing Pipeline for Stereoscopic Cameras”; IEEE Transactions on Consumer Electronics, Vol. 56, No. 2, Mayo 2010.

LIBROS

- **Lipton**, Lenny; “Foundations of the Stereoscopic Cinema”. Van Nostrand Reinhold Company, Nueva York, 1982.
- **Minoli**, Daniel; “3DTV Content Capture, Encoding and Transmission”. John Wiley & Sons; New Jersey; 2010.
- **Richardson**, Iain; “H.264 and MPEG-4 Video compression”. Wiley, Inglaterra, 2003.

REPORTES Y RECOMENDACIONES UIT

- INFORME UIT-R BT.2088, “Televisión estereoscópica”, 2006.
- ITU-T Recommendation H.264, “Advanced video coding for generic audiovisual Services” 2010.

- RECOMENDACIÓN UIT-R BT.601. Cuestión UIT-R 1/6, "Parámetros de codificación de televisión digital para estudios con formatos de imagen normal 4:3 y de pantalla ancha 16:9". 2007.
- RECOMENDACIÓN UIT-R BT.1198, "Televisión estereoscópica basada en dos señales de canal de ojo derecho y ojo izquierdo"; 1995.
- RECOMENDACIÓN UIT-R BT.1438, "EVALUACIÓN SUBJETIVA DE LAS IMÁGENES DE TELEVISIÓN ESTEREOSCÓPICA", 2000.
- Report ITU-R BT.2160-1; "Features of three dimensional television video systems for broadcasting BT Series", 2010.

PÁGINAS WEB

- <http://estereocosas.webs.com/>
- http://pro.sony-asia.com/product/resources/en_BA/images/Brochure/BCAP/PMW-TD300_2309.pdf
- http://pro-av.panasonic.net/en/sales_o/broch_pdf/ag-3dp1.pdf
- <http://web.media.mit.edu/~halazar/autostereo/autostereo.html>
- <http://www.21stcentury3d.com/>
- <http://www.3dcgi.com/cooltech/displays/displays.htm>
- <http://www.3dvivant.eu/>
- <http://www.actuality-systems.com/>
- <http://www.dibeg.org/techp/techp.htm>
- http://www.faqs.org/patents/imgfull/20110234769_05
- <http://www.felix3d.com/web/index.php?show=Publicaciones>
- <http://www.red3i.es/stereoweb/video3d.htm>
- <http://www.stereographics.com/>
- <http://www.stereoscopic.org/proc/index.htm>
- <http://www.tmbroadcast.es/index.php/television-estereoscopica/>
- www.3d.curtin.edu.au
- www.elsevier.com/locate/jvci
- www.sigmadesigns.com

GLOSARIO DE TÉRMINOS

3DTV	Three dimensional Television
AAC	Advanced Audio Coding
AVC	Advanced Video Coding
CABAC	Context-adaptive binary arithmetic coding
CAD	Computer Aided Design
CAVLC	Context-Adaptive Variable Length
CCD	Charged Couple Display
CFC	Conventional HD Frame Compatible
CRT	Cathode Ray Tube
CSV	Conventional Stereo Video
DES	Depth Enhanced Video
DIBR	Digital Image Based Rendering
DMB	Digital Multimedia Broadcasting
DQPSK	Diferential QPSK
DSCQS	Double-Stimulus Continuous Quality-Scale
DVB-T	Digital Video Broadcasting- Terrestrial
FHD3D	Full HD 3D
FVV	Free View-point Video
GOP	Group of Picture
HDTV	High Definition Television
HHD3D	Half HD 3D
ISDB-T	Integrate Service Digital Broadcasting – Terrestrial
L+R	Left + Right
LCD	Liquid Crystal Display
LCS	Liquid Crystal Shutter
LDV	Layered Depth Video
MVC	Multiview Video Coding
MVD	Multiview Video Coding + Depth
NAL	Network Abstraction Layer
OCA	Optimized Chromatic Anagliph
OFDM	Orthogonal Frecuency Division Multiplexing

PDP	Plasma Display Panel
PSNR	Peak Signal Noise Ratio
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
SbS	Side by Side
SDTV	Standar Definition Television
SFN	Single Frecuency Network
STB	Set Top Box
SVC	Scalable Video Coding
TaB	Top and Bottom
TDT	Televisión Digital Terrestre
TOF	Time of Flight
V+D	Video + Depth
VCL	Video Coding Layer

ANEXO 1

Extracto de la Especificación High-Definition Multimedia Interface
Versión 1.4a (HDMI 1.4a) (porción de señalización para video 3D –
Formato Digital)

ANEXO 2

Detalle de la codificación para los formatos LDV (Layered Depth Video).

ANEXO 3

Métodos de configuración y calibración de las cámaras para producción de Televisión estereoscópica (3DTV) (Formato Digital).

ANEXO 4

Extracto de algunos resultados de pruebas experimentales
relacionadas a la 3DTV.

(Pruebas completas y detalladas en formato digital)

ANEXO 5

Extracto del estudio de la experiencia visual del usuario de señales de televisión estereoscópica (3DTV) al usuario (Estudio completo - Formato Digital)

ANEXO 6

Características de algunos equipos para televisión estereoscópica
(3DTV) (Formato Digital)