

**JUAN CARLOS CUÉLLAR QUIÑONEZ**

**MODELO PARA LA MEDICIÓN DE CALIDAD DE LA  
EXPERIENCIA PARA EL SERVICIO DE IPTV**

**Tesis presentada a la Facultad de Ingeniería  
Electrónica y Telecomunicaciones de la  
Universidad del Cauca para la obtención del  
Título de**

**Doctor en:  
Ingeniería Telemática**

**Director:  
José Luis Arciniegas Herrera, PhD**

**Co-director:  
Jesús Hamilton Ortiz, PhD**

**Popayán  
Junio 14 de 2018**



**Página Aceptación**

**(Pendiente)**



A mi familia y amigos cercanos por todo el apoyo brindado.



## AGRADECIMIENTOS

A Dios, mi familia y amigos cercanos por todo el apoyo que me brindaron durante estos años de sacrificio para poder culminar este proyecto.

A Flor Magnolia Gómez, esa persona especial que me acompañó en 2016 y parte de 2017, su apoyo y compañía fue fundamental mientras hacía las pruebas de laboratorio y durante mi estancia en Madrid. Gracias, gracias, gracias.

A mis tutores José Luis Arciniegas y Jesús Hamilton Ortiz, por todo el apoyo y guía durante el proyecto. Sus consejos y recomendaciones fueron muy valiosas.

A José Luis, mi hermano, y al profesor Norbey Marín por las muchas explicaciones en la parte estadística involucradas en este proyecto.

A Álvaro Pachón, mi amigo y compañero de trabajo, por todo el apoyo y sugerencias durante el desarrollo del proyecto.

A Valentina Moreno, David Acosta y Miguel Ossa por el apoyo durante el desarrollo de las pruebas de laboratorio.

A Sara Ortiz, Elmer Galvez y Don Alejandro Ortiz por recibirme en Madrid y ayudarme durante toda la estancia realizada.

A Encarna Pastor, Luis Bellido y Carlos Lentisco, que me colaboraron en la Universidad Politécnica de Madrid durante mi estancia.

Al Comité de Doctorado y Vice-Rectoría de Investigaciones de la Universidad del Cauca por el apoyo para realizar la pasantía en Madrid.

A la Universidad Icesi y la Decanatura de la Facultad Ingeniería por el apoyo durante todo el proyecto y durante la pasantía realizada en Madrid.



## Resumen estructurado

**Antecedentes:** Con el crecimiento de la demanda de servicios de video en Internet y en especial el servicio de IPTV, se hace necesario que los proveedores de servicio implementen esquemas de calidad de servicio en sus redes de núcleo, con el fin de garantizar el funcionamiento adecuado del servicio de IPTV que ofrecen a sus usuarios. Y con el fin de verificar el funcionamiento del servicio de IPTV, la comunidad de investigadores ha propuesto diferentes tipos de modelos que permiten evaluar la calidad de experiencia del video transmitido. Generalmente el proceso seguido para obtener estos modelos se basa en pruebas subjetivas que utilizan un número limitado de parámetros de calidad de servicio [32-40]. A diferencia de otras propuestas, esta investigación propone un modelo para medir la calidad de la experiencia, basándose en parámetros de calidad de servicio y utilizando un modelo objetivo para la evaluación de la calidad del video.

**Objetivos:** Proponer un modelo que permita medir la calidad de la experiencia en vivo para el servicio de IPTV en un entorno residencial, con la finalidad de permitir a los proveedores de servicio monitorizar de forma continua la calidad de la experiencia del video y realizar ajustes en tiempo real en la red, para garantizar el buen funcionamiento del servicio al usuario final.

**Método:** Basándose en la revisión del estado del arte, se identificaron los parámetros de calidad de servicio, el método objetivo con el cual se generará el modelo mediante la metodología de diseño de experimentos. Se grabaron tres videos de prueba utilizando la maqueta de experimentación diseñada para el proyecto, se transmitieron los videos de prueba modificando los parámetros de calidad de servicio definidos. Después de transmitidos los videos y procesados en el destino, la calidad de la experiencia de los videos fue evaluada utilizando el método objetivo elegido. Con los valores de los parámetros de calidad de servicio y las medidas obtenidas con el método objetivo, después de procesar los datos y utilizando metodología de superficies de respuesta se generó el modelo propuesto.

**Resultados:** Con los datos obtenidos en la maqueta de experimentación, se genera un modelo para cada video de prueba. El modelo propuesto permite calcular la calidad de la experiencia asociada al video conociendo los valores de los parámetros de calidad de servicio (retardo, variación del retardo y pérdida de paquetes) del video analizado. Para verificar el desempeño de cada modelo, los resultados entregados por cada uno de ellos son comparados con pruebas subjetivas y dos métodos objetivos estandarizados. Los resultados obtenidos indican que existe una correlación positiva moderada entre los resultados que entregan los modelos propuestos y las pruebas subjetivas y los métodos objetivos estandarizados.

**Conclusiones:** El modelo propuesto tiene las siguientes ventajas con respecto a las propuestas analizadas en la revisión del estado del arte:

- Se puede utilizar para la medición de la calidad de la experiencia en línea, sin decodificar el video y sin recurrir a pruebas subjetivas
- La complejidad computacional que requiere el modelo es baja en comparación con la complejidad computacional alta que requieren los modelos basados en métodos objetivos, ya que solo requiere el valor de los parámetros de calidad de servicio para obtener la medición de la calidad de la experiencia del video analizado.

**Palabras clave:** Calidad de Servicio, Calidad de la Experiencia, IPTV.

## Structured abstract

**Background:** With the growth in demand of video services on the Internet and especially the IPTV service, it has become necessary for service providers to implement quality of service schemes in their core networks to guarantee the functionality of such services. To verify the functioning of the IPTV service, the research community has proposed different types of models that permit one to evaluate the quality of experience of the transmitted video. Nevertheless, the process for obtaining these models is generally based on subjective tests and the utilization of a limited number of quality of service parameters [32-40]. Therefore, this research proposes a model for measuring quality of experience based on quality of service parameters and utilizing an objective model for the evaluation of video quality.

**Aims:** To propose a model that allows one to measure the quality of experience for an online IPTV service, so that service providers can monitor this quality of experience and make real time adjustments to guarantee optimum functioning for the end user.

**Method:** Based on a systematic revision of the state of the art, the parameters of quality of service and the objective model were defined. Three test videos were recorded and these were transmitted in a testbed modifying the defined parameters of quality of service. The quality of experience of the videos was evaluated utilizing an objective model. The proposed model was generated with the data obtained from the quality of service parameters and the measurements obtained from the objective model.

**Results:** With the data obtained in the experimental model, a model is generated for each test video. The proposed model allows to calculate the quality of the experience associated with the video knowing the values of the quality of service parameters (delay, jitter and packet loss) of the video analyzed. To verify the performance of each model, the results delivered by each of them are compared with subjective tests and two standardized objective methods. The results obtained indicate that there is a moderate positive correlation between the results delivered by the proposed models and the subjective tests and the standardized objective methods.

**Conclusions:** The proposed model has the following advantages with respect to the proposals analyzed in the review of the state of the art:

- It can be used to measure the quality of the online experience, without decoding the video and without resorting to subjective tests.
- The computational complexity required by the model is low compared to the high computational complexity required by models based on objective methods. Since it only requires the value of the parameters of quality of service to obtain the measurement of the quality of the experience of the analyzed video.

**Keywords:** Quality of Service, Quality of Experience, IPTV.

# Contenido

	Pág.
Lista de Figuras.....	15
Lista de Tablas.....	17
Lista de Acrónimos.....	19
1 Introducción.....	21
1.1 Planteamiento del problema.....	21
1.2 Objetivos.....	22
1.3 Contribución de la tesis.....	22
1.4 Organización de la tesis.....	23
2 Estado del arte.....	25
2.1 Metodología.....	25
2.2 Antecedentes.....	26
2.2.1 Modelos Subjetivos.....	26
2.2.2 Modelos objetivos.....	27
2.2.3 Modelos Indirectos.....	29
2.3 Trabajos relacionados.....	30
2.4 Comentarios del Capítulo.....	34
3 Componentes para generar los datos y obtener el modelo para medir QoE para el servicio de IPTV.....	37
3.1 Parámetros de calidad de servicio seleccionados.....	38
3.1.1 Retardo ( <i>Delay</i> ).....	38
3.1.2 Variación del Retardo ( <i>Jitter</i> ).....	38
3.1.3 Pérdida de paquetes ( <i>Packet Loss</i> ).....	38
3.2 Modelo objetivo seleccionado para medir la calidad de la experiencia.....	39
3.3 Maqueta de Experimentación.....	44
3.4 Videos Fuente.....	45
3.5 Comentarios del Capítulo.....	50
4 Diseño del Experimento para obtener los datos para generar el modelo.....	51
4.1 Diseño del Experimento.....	52
4.1.1 Identificación de los factores y sus respectivos rangos.....	52

4.1.2	Selección de variable de respuesta .....	53
4.1.3	Elección del diseño experimental .....	53
4.1.4	Ejecución del experimento.....	61
4.1.5	Proceso para obtener los rangos de los factores del experimento .....	62
4.1.6	Análisis estadístico de los datos .....	64
4.2	Comentarios del capítulo.....	78
5	Análisis de Evaluación de Desempeño de los Modelos Propuestos .....	81
5.1	Ajuste de los datos para realizar los análisis comparativos.....	82
5.2	Comparación de desempeño entre modelos propuestos vs modelos objetivos	88
5.2.1	Modelo Video Movimiento Bajo para el video Entrevista .....	91
5.2.2	Modelo Video Movimiento Moderado - Video Cafetería .....	92
5.2.3	Modelo Video Movimiento Alto para el Video Baloncesto.....	93
5.3	Comparación de desempeño entre modelos propuestos vs pruebas subjetivas	95
5.3.1	Pruebas para comparar el desempeño de los modelos propuestos.....	95
5.3.2	Proceso de comparación de desempeño: modelos propuestos vs. pruebas subjetivas .....	96
5.4	Pruebas estáticas de cada modelo.....	103
5.5	Comentarios del capítulo.....	106
6	Conclusiones y recomendaciones .....	111
6.1	Conclusiones .....	111
6.2	Productos obtenidos durante el desarrollo del proyecto.....	113
6.2.1	Publicaciones.....	113
6.2.2	Base de Datos .....	114
6.3	Trabajos Futuros .....	115
6.4	Lecciones aprendidas.....	116
	Anexo 1: Configuración VLC Y NetEM.....	117
	Anexo 2: Pruebas Subjetivas .....	123
7	Bibliografía.....	131

## Lista de Figuras

	Pág.
Figura 1. Descripción del contenido del capítulo 1.....	21
Figura 2. Descripción del contenido del capítulo 2.....	25
Figura 3. Métodos Objetivos para evaluar la calidad de la experiencia del video. Tomado de [9]. .....	29
Figura 4. Descripción del contenido del capítulo 3.....	37
Figura 5. Encapsulamiento de H.264. ....	41
Figura 6. Maqueta de experimentación implementada para obtener los datos para generar el modelo. ....	44
Figura 7. Ejemplos de videos de prueba con sus características espacio-temporales. Tomado de [14]. .....	47
Figura 8. Distribución de los valores IE & IT de los videos originales utilizados en el experimento.....	50
Figura 9. Descripción del contenido del capítulo 4.....	51
Figura 10. Tabla de codificación porción factorial del DCC para tres factores y su representación geométrica.....	55
Figura 11. (a) Puntos Axiales y factoriales del DCC. (b) Valores asociados a la codificación del eje X, parámetro retardo. ....	55
Figura 12. Proceso para lo obtener los datos para la generación de los tres modelos .....	64
Figura 13. Diagrama de barras de la QoE obtenida al video Entrevista en escala VQM. ....	66
Figura 14. Gráfica de contorno QoE_Entrevista vs Pérdida de paquetes, Variación del Retardo. ....	68
Figura 15. Gráfica de contorno QoE_Entrevista vs Pérdida de Paquetes, Retardo. .	68
Figura 16. Gráfica de contorno QoE_Entrevista vs Retardo, Variación del Retardo.	69
Figura 17. Diagrama de barras de la QoE obtenida al video Cafetería en escala VQM. ....	71
Figura 18. Gráfica de contorno QoE_Cafetería vs Retardo, Perdida de Paquetes. ..	72
Figura 19. Gráfica de contorno QoE_Cafetería vs Retardo, Variación del Retardo. .	73
Figura 20 . Gráfica de contorno QoE_Cafetería vs Perdida de Paquetes, Variación del Retardo. ....	73
Figura 21. Diagrama de barras de la QoE obtenida para el video Entrevista en escala VQM. ....	75

Figura 22. Gráfica de contorno QoE_Baloncesto vs Pérdida de Paquetes, Retardo. ....	77
Figura 23. Gráfica de contorno de QoE_Baloncesto vs Pérdida de paquetes, Variación del retardo. ....	77
Figura 24. Gráfica de contorno de QoE_Baloncesto vs Retardo, Variación del retardo. ....	78
Figura 25. Descripción del contenido del capítulo 5. ....	81
Figura 26. Gráfica de la regresión no lineal correspondiente al MVMB - video Entrevista. ....	86
Figura 27. Gráfica de la regresión no lineal correspondiente al MVMM - video Cafetería. ....	86
Figura 28. Gráfica de la regresión no lineal correspondiente al MVMA - video Baloncesto. ....	87
Figura 29. Interpretación de los valores que entrega el coeficiente de correlación de Pearson. Adaptado de [86]. ....	91
Figura 30. (a) Medidas en escala MOS obtenidas al evaluar los parámetros de QoS del Modelo Entrevista. (b) Medidas obtenidas al realizar las pruebas subjetivas a los 16 tratamientos del video Entrevista. ....	98
Figura 31. (a) Medidas en escala MOS obtenidas al evaluar los parámetros de QoS del MVMM. (b) Medidas obtenidas al realizar las pruebas subjetivas a los 16 tratamientos para el video cafetería. ....	100
Figura 32. (a) Medidas en escala MOS obtenidas al evaluar los parámetros de QoS del MVMA. (b) Medidas obtenidas al realizar las pruebas subjetivas a los 16 tratamientos del video Baloncesto. ....	102
Figura 33. MOS vs Pérdida de paquetes y variación del retardo para el MVMB para el video entrevista. ....	104
Figura 34. MOS vs Pérdida de Paquetes y retardo, para el MVMM para el video cafetería. ....	105
Figura 35. MOS vs Pérdida de paquetes y variación del retardo para el MVMA para el video baloncesto. ....	105
Figura 36. Diagrama de Flujo del algoritmo para obtener el modelo para medir QoE para el servicio de IPTV. ....	109
Figura 37. Algoritmo para obtener el modelo para medir QoE para el servicio de IPTV. ....	110
Figura 38. Descripción del contenido del capítulo 6. ....	111

## Lista de Tablas

	Pág.
Tabla 1. Cuadro comparativo de trabajos relacionados que utilizan modelos indirectos en sus propuestas. ....	32
Tabla 2. Especificaciones técnicas de las máquinas utilizadas en la maqueta de experimentación. ....	45
Tabla 3. Características espacio temporales de los videos usados en la maqueta de experimentación. ....	48
Tabla 4. Rango de los parámetros de calidad de servicio configurados en el experimento . ....	52
Tabla 5. Tratamientos para el experimento con los valores codificados y valores reales asociados a cada factor. ....	57
Tabla 6. Datos para el modelo presentado en la Ecuación 4 . ....	58
Tabla 7. Medición de QoE obtenida para el tratamiento No.1 .....	61
Tabla 8. Datos obtenidos para generar la modelo del video entrevista. ....	65
Tabla 9. Análisis de Varianza para los datos obtenidos al video entrevista. ....	66
Tabla 10. Datos obtenidos para generar el modelo del video cafetería. ....	70
Tabla 11. Análisis de Varianza para los datos obtenidos al video Cafetería. ....	71
Tabla 12. Datos obtenidos para generar el modelo del video Baloncesto. ....	74
Tabla 13. Análisis de Varianza para los datos obtenidos al video Baloncesto. ....	75
Tabla 14. Coeficientes de los tres modelos obtenidos. ....	79
Tabla 15. Resultados obtenidos al evaluar cada modelo con los parámetros de QoS definidos para cada tratamiento. La QoE obtenida está en escala VQM. ....	82
Tabla 16. Valores de las pruebas subjetivas realizadas a los tres videos con su correspondiente valor en escala VQM obtenido por cada modelo propuesto. ....	85
Tabla 17. Valores de las constantes de a,b,c,d y valor del R_cuadrado de la regresión no lineal para cada modelo propuesto. ....	85
Tabla 18. Escalas para mapear MOS vs VQM para cada video. ....	87
Tabla 19. Medidas obtenidas de cada modelo en escala MOS. ....	88
Tabla 20. Resultados obtenidos al evaluar los 16 tratamientos para los tres videos con PSNR. ....	89
Tabla 21. Resultados obtenidos al evaluar los 16 tratamientos para los tres videos con SSIM. ....	90
Tabla 22. Mapeo de escala MOS a PSNR y SSIM. ....	90

Tabla 23. Datos obtenidos para evaluar el modelo video movimiento bajo para el video Entrevista. ....	91
Tabla 24. Resultados estadísticos para comparar el desempeño del modelo propuesto para video Entrevista.....	92
Tabla 25. Datos obtenidos para evaluar el modelo Cafetería.....	93
Tabla 26. Resultados estadísticos para comparar el desempeño del modelo video movimiento moderado para el video cafería.....	93
Tabla 27. Datos obtenidos para evaluar el modelo Baloncesto.....	94
Tabla 28. Resultados estadísticos para comparar el desempeño del modelo video movimiento alto para el video Baloncesto. ....	94
Tabla 29. Resultados en escala de MOS del modelo propuesto y las pruebas subjetivas realizadas a los 16 tratamientos del video Entrevista.....	97
Tabla 30. Datos estadísticos de los resultados del MVMB – video entrevista y los resultados obtenidos mediante las pruebas subjetivas. ....	97
Tabla 31. Resultados de pruebas de desempeño del MVMB. ....	98
Tabla 32. Resultados en escala de MOS del modelo propuesto y las pruebas subjetivas realizadas a los 16 tratamientos del video Cafetería. ....	99
Tabla 33. Datos estadísticos de los resultados del modelo video movimiento moderado para el video Cafetería y los resultados obtenidos mediante las pruebas subjetivas. ....	100
Tabla 34. Resultados de pruebas de desempeño del MVMM –video cafetería. ....	101
Tabla 35. Resultados en escala de MOS del modelo propuesto y las pruebas subjetivas realizadas a los 16 tratamientos del video Baloncesto.....	101
Tabla 36. Datos estadísticos de los resultados del MVMA y los resultados obtenidos mediante las pruebas subjetivas. ....	102
Tabla 37. Resultados de pruebas de desempeño del modelo Baloncesto.....	103
Tabla 38. Resumen de pruebas de desempeño de los tres modelos propuestos...	106

## Lista de Acrónimos

IPTV: Televisión sobre Protocolo IP

QoS: Quality of Service; Calidad de Servicio

QoE: Quality of Experience, Calidad de la Experiencia

MOS: Mean Opinion Score

PSNR: Peak-Signal-to-Noise-Ratio

SSIM : Structural SiMilarity

VQM: Video Quality Metric

NetEm: Network Emulator

GoP: Group of Picture

MSR: Metodología de Superficies de Respuesta

CCP: Coeficiente de Correlación de Pearson

CCS: Coeficiente de Correlación de Spearman

PVA: Porcentaje de Valores Atípicos

DCC: Diseño Central Compuesto

MVMB: Modelo Video Movimiento Bajo

MVMM: Modelo Video Movimiento Moderado

MVMA: Modelo Video Movimiento Alto



# 1 Introducción

En este capítulo, se presenta el planteamiento del problema, el objetivo general y los objetivos específicos del proyecto, la contribución de la tesis y al final, la organización del documento. En la Figura 1, se presenta la descripción general del capítulo y su contenido, identificando cada sección con el fin de que se pueda observar de manera general el contenido del capítulo y cada sección.

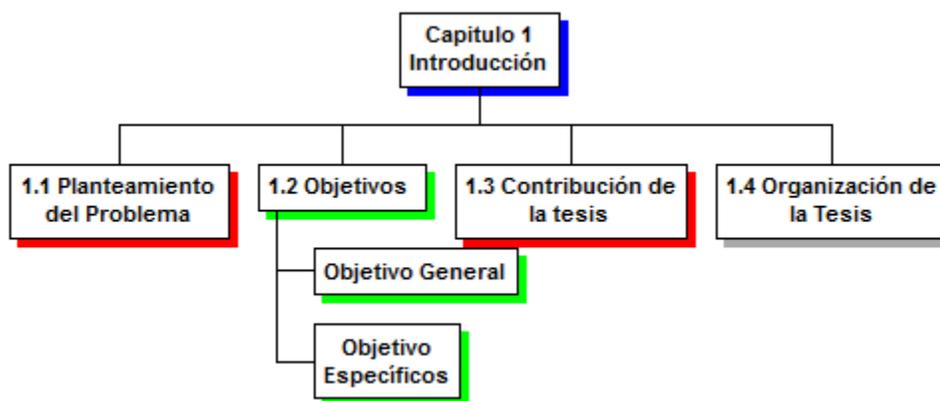


Figura 1. Descripción del contenido del capítulo 1.

## 1.1 Planteamiento del problema

El crecimiento de los servicios de video sobre IP, incluyendo el servicio de IPTV (Televisión sobre Protocolo IP) [1] hará que, para el 2020, éste corresponda con el 82% del tráfico global de Internet, frente al 70% en 2015 [2]. Éste crecimiento acelerado de tráfico, hace necesario que los proveedores de servicio implementen esquemas de calidad de servicio en sus redes de núcleo, con el fin de garantizar el funcionamiento del servicio de IPTV que ofrecen a sus usuarios.

La implementación de esquemas de calidad de servicio no garantiza que la percepción del usuario por la utilización del servicio sea siempre positiva, ya que existen múltiples factores que pueden afectar esta percepción (tamaño de la pantalla, contenido de la programación, tipo de servicio contratado, etc). En consecuencia esta percepción ha evolucionado al concepto de **Calidad de la Experiencia (Quality of Experience - QoE)**, que la UIT-T en su Recomendación G.1080 [3] define como *“la aceptabilidad global de una aplicación o servicio, tal como la percibe subjetivamente el usuario final”*. La calidad de experiencia combina parámetros no técnicos como la percepción, la experiencia y expectativa del usuario, con parámetros técnicos de calidad de servicio a nivel de red.

Por lo tanto, para los proveedores de servicio resulta importante evaluar esta percepción del usuario sobre los servicios que ha contratado con el fin de mejorar

continuamente y mantenerlo satisfecho [4] [5]. Esto motiva la necesidad de establecer y analizar la relación existente entre la QoE percibida por el usuario, con parámetros de **Calidad de Servicio (Quality of Service - QoS)** a nivel de red, con el fin de estimar en cualquier momento la QoE percibida por el usuario.

Después de lo anteriormente expuesto, se plantea la pregunta de investigación a resolver en este proyecto: *¿Cómo generar un modelo para medir QoE en línea para el servicio de IPTV?*

## 1.2 Objetivos

### Objetivo general

- ✓ Proponer un modelo que permita medir QoE en línea<sup>1</sup> para el servicio de IPTV en un entorno residencial

### Objetivos específicos

1. Identificar y seleccionar los parámetros y las métricas asociados a QoS, para evaluar la calidad del servicio de IPTV en un entorno residencial.
2. Seleccionar y adaptar un algoritmo Sin Referencia (No Reference) para la medición de la calidad de video con el fin de obtener una medida de la QoE percibida por el usuario. Identificar e implementar una arquitectura de pruebas en laboratorio, con el fin de obtener los datos necesarios para la generación del modelo.
3. Generar el modelo empleando análisis de regresión multivariable con los datos obtenidos de la variación de los parámetros de QoS y la respuesta entregada por el algoritmo Sin Referencia.

## 1.3 Contribución de la tesis

La contribución de la tesis a la solución del problema consiste en proponer un modelo, utilizando parámetros de calidad de servicio y un algoritmo sin referencia para evaluar la calidad de experiencia al servicio de IPTV. Este tipo de aproximación tiene varias ventajas: (i) No requiere de pruebas subjetivas para obtener datos para generar el modelo. (ii) No se requiere de la decodificación del video y requiere menor ancho de banda para su monitoreo [6], y (iii) Este tipo de modelos permiten evaluar la calidad del video en tiempo real, donde usualmente no se tiene la señal de video original como referencia [7]. Adicional a estas ventajas, el modelo propuesto permitirá a los proveedores de servicio evaluar directamente la calidad de la experiencia en el cliente

---

<sup>1</sup> En línea se refiere a transmisiones en vivo.

IPTV residencial, y con esto los proveedores de servicio podrán ajustar de manera automática parámetros de calidad de servicio para ofrecer un servicio entre los parámetros adecuados al usuario final.

Estas ventajas contrastan con limitaciones de modelos propuestos en los trabajos relacionados, debido a la cantidad de parámetros de calidad de servicio que incluyen en sus análisis y que algunos son generados utilizando pruebas subjetivas. Estas limitaciones se reflejan en los resultados obtenidos al evaluar dichos modelos en entornos reales, ya que los resultados obtenidos por estos modelos se alejan de la calidad de la experiencia percibida por el usuario.

## **1.4 Organización de la tesis**

Este documento está organizado de la siguiente manera:

En el capítulo 2, se presenta el estado del arte, que inicia con los antecedentes donde se explican los modelos subjetivos, objetivos e indirectos. Posteriormente producto de una revisión sistemática se presentan los trabajos relacionados donde se realiza una comparación entre los diferentes trabajos analizados desde tres puntos de vista: Proceso utilizado para obtener los datos para generar el modelo, Tipo de modelo utilizado para evaluar calidad de experiencia y Parámetros de QoS utilizados para generar el modelo. Al finalizar el capítulo se presentan las publicaciones realizadas con base en el estado del arte.

En el capítulo 3, se presentan los componentes para generar los datos para obtener el modelo para medir la QoE al servicio de IPTV. Se empieza describiendo los parámetros de QoS seleccionados, posteriormente se justifica el porqué de la elección del modelo objetivo para realizar las medidas de QoE, se describe la maqueta de experimentación implementada y al finalizar, el proceso para seleccionar los videos fuente para utilizar en el experimento.

En el capítulo 4, se describe detalladamente el diseño de experimento realizado para obtener los datos que permitieron generar el modelo, el análisis estadístico de los datos y al finalizar se presentan los modelos obtenidos con su respectivo análisis de varianza.

En el capítulo 5, se presentarán las pruebas realizadas a los modelos propuestos para verificar su desempeño al momento de evaluar la calidad de la experiencia. Se presentan comparaciones de resultados con dos modelos objetivos para finalizar con la comparación de los resultados de los modelos propuestos con pruebas subjetivas que se realizaron a los videos utilizados durante el desarrollo del experimento.

En el capítulo 6, se presentan las conclusiones del proyecto, los productos obtenidos a lo largo del proyecto, los trabajos futuros y las lecciones aprendidas.

En el anexo 1 se presenta el proceso de configuración de la herramienta para realizar la transmisión de video y la herramienta utilizada para variar el valor de los parámetros de calidad de servicio.

En el anexo 2 se documenta el proceso que se siguió en las pruebas subjetivas que se realizaron a los videos utilizados para generar los datos para obtener los modelos propuestos.

## 2 Estado del arte

En este capítulo se presenta en la primera sección, la metodología que se siguió para analizar el estado del arte que permitió identificar los principales aspectos procedimentales a seguir en el proyecto. En la sección antecedentes, se presentan los conceptos básicos sobre modelos subjetivos, objetivos e indirectos. En la sección de trabajos relacionados se analizan las principales propuestas de modelos para la medición de la calidad de experiencia para el servicio de IPTV bajo el enfoque de modelos indirectos. Al final de dicha sección mediante una tabla comparativa se presentan las características principales de cada propuesta. En la Figura 2, se presenta la descripción del contenido de este capítulo a fin de orientar su lectura.

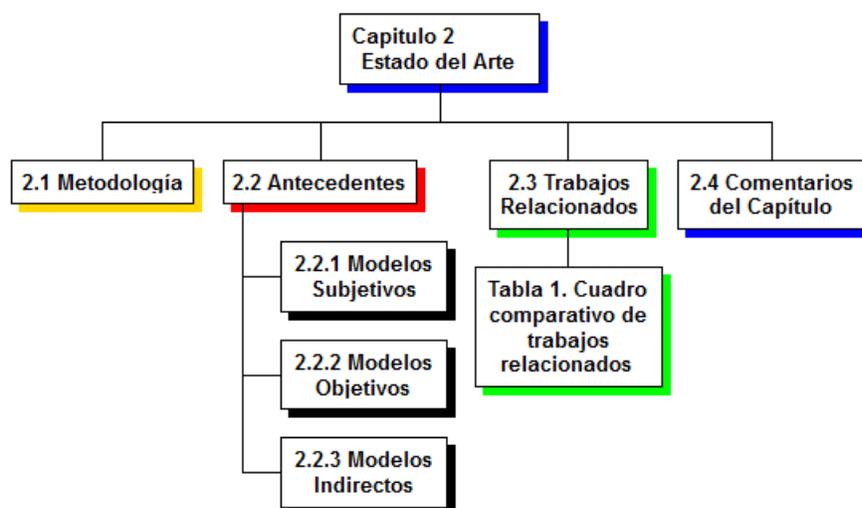


Figura 2. Descripción del contenido del capítulo 2.

### 2.1 Metodología

El estado del arte se construyó a partir de la revisión y análisis de trabajos relacionados basándose en lo presentado en [8] y enfocándose en los siguientes temas y palabras clave: calidad de servicio para IPTV, calidad de la experiencia para IPTV, modelos para medir QoE para el servicio de IPTV y correlación de QoS y QoE para el servicio de IPTV. Se consultaron las bases de datos bibliográficas: IEEEExplore, ACM Digital Library, Springer, Elsevier, ScienceDirect, Scopus, Wiley online Library y el sitio web de la Unión Internacional de Telecomunicaciones (UIT).

La revisión del estado se realizó analizando en las publicaciones los siguientes aspectos: Proceso utilizado para obtener los datos para generar los modelos, tipo de modelo empleado para evaluar la calidad de la experiencia y los parámetros de calidad

del servicio utilizados para generar el modelo. Para lograr identificar estas tres características se revisaron 60 artículos de los cuales, 20 abordaban la temática de calidad de la experiencia y 12 de estos artículos, fueron los elegidos para construir el estado del arte y fueron la base para implementar el diseño del experimento para obtener los datos para generar el modelo propuesto. Producto de esta revisión del estado del arte se publicó el artículo “*Clasificación y Análisis de Métodos para medir Calidad de la Experiencia del Servicio de Televisión sobre Protocolo IP (IPTV)* [9].

La revisión del estado del arte permitió identificar la metodología utilizada en los diferentes modelos propuestos para evaluar la calidad de experiencia para el servicio de IPTV, y aspectos para diseñar la propuesta metodológica para obtener el modelo en esta investigación. Resultado de esta revisión del estado del arte y el artículo [9] publicado, se presentaron trabajos en eventos nacionales e internacionales. Al final de este capítulo se presenta un resumen de las publicaciones realizadas basadas en la revisión del estado del arte.

## 2.2 Antecedentes

Con el fin de monitorear la calidad de experiencia y garantizar una percepción positiva por parte del usuario sobre el servicio prestado, se ha fortalecido la línea de investigación que busca medir la calidad de experiencia al servicio de IPTV.

Por esta razón, para evaluar la QoE, en [10] se proponen tres modelos para este fin: (i) modelos subjetivos, (ii) modelos objetivos y (iii) modelos indirectos, los cuales se explican a continuación.

Estar monitoreando la calidad de la experiencia para el servicio de IPTV residencial se ha convertido en una prioridad para los proveedores de servicio, ya que permite de manera continua monitorear el servicio prestado y ante alguna falla, ajustar parámetros de calidad de servicio con el fin de mantener un servicio de calidad para el usuario final [4].

### 2.2.1 Modelos Subjetivos

En este tipo de modelos se evalúa la calidad de la experiencia utilizando personas que califican la calidad de una secuencia de videos mediante el diligenciamiento de encuestas. Este tipo de experimentos son diseñados, siguiendo los lineamientos planteados por la UIT en [11],[12],[13],[14]. La escala de medición utilizada es el MOS (*Mean Opinion Score*), que es la escala de medida más popular que se usa para analizar y calificar la calidad del video [15]. El usuario cuando ve el video lo califica utilizando una escala discreta de 5 niveles: **1-Malo (Bad)**, **2-Pobre o Mediocre (Poor)**, **3- Justo o Regular (Fair)**, **4-Bueno (Good)** y **5-Excelente (Excellent)** [16].

Este tipo de métodos requieren tiempo y son costosos por la logística necesaria para su implementación, pero son de gran utilidad y un estándar de facto para la validación de modelos propuestos. No es fácil automatizar esta medida, ya que se debe tener en cuenta la influencia de los factores psicológicos, humanos y el sesgo del usuario al momento de calificar cada video. En la referencia [15] se presenta la metodología para

realizar este tipo de pruebas desde la elección del ambiente de pruebas, selección de videos y análisis de datos obtenidos.

Este tipo de modelos no se pueden utilizar para el análisis de la calidad de experiencia asociada al video en línea debido a que las pruebas se deben hacer en un espacio diseñado para tal fin, como también la elección de los videos a utilizar en la prueba subjetiva es de vital importancia.

### 2.2.2 Modelos objetivos

Con el fin de obtener una medida confiable de QoE pero evitando la realización de pruebas subjetivas, se han desarrollado los modelos objetivos. Los modelos objetivos estiman una medida de QoE en función de parámetros de QoS y características asociadas al video. El enfoque de estos modelos no permite que la medida se pueda realizar en vivo, debido a la utilización de una señal de referencia para comparar los resultados del video transmitido, con el video recibido. Solo en los métodos basados en paquetes pueden hacer mediciones en línea, pero en la actualidad este tipo de modelos basados en métodos indirectos está en desarrollo.

Los modelos objetivos se clasifican en cuatro métodos principalmente, y esta clasificación es de acuerdo al tipo de análisis o extracción de información realizada a la señal de video: (i) Métodos con señal de referencia, (ii) Método basado en datos, (iii) Métodos basados en imagen y (iv) Métodos basados en paquetes.

- **Métodos con señal de referencia**

Estos métodos se encuentran basados en la utilización o no de la señal original del video como referencia para ejecutar las mediciones de QoE y se clasifican en:

- *Modelo referencia completa (Full Reference)*: modelo en el cual la señal de video original está disponible en el receptor para comparar la señal original con la señal del video recibida y así, determinar la calidad del video enviado al usuario. La limitación o desventaja de este modelo es que no evalúa la calidad absoluta. Es costoso en tiempo porque requiere poder de procesamiento para ejecutar algoritmos complejos [17],[18].
- *Modelo referencia reducida (Reduced Reference)*: en este modelo está disponible en el receptor la información parcial de la señal de video original transmitida, con el fin de determinar la calidad del video recibido. Este modelo es más económico en tiempo y procesamiento que el modelo referencia completa.[17].
- *Modelo sin referencia (No Reference)*: en este modelo solo se posee la señal de video recibida para determinar la calidad del video. La calidad es medida basándose en parámetros que reflejan la degradación de la señal, muy útil cuando se desean hacer mediciones en tiempo real [17].

- **Método basado en datos**

Este tipo de método evalúa la fidelidad del video sin considerar su contenido, hace una comparación píxel a píxel entre los cuadros (*frames*) de cada imagen [18]. El algoritmo más utilizado que funciona bajo este esquema es el PSNR (Peak-Signal-to-Noise-Ratio) [19],[20].

- **Métodos basados en imagen**

Los métodos basados en imagen analizan la información visual que contiene el video y presentan dos enfoques para evaluar el video:

- *Enfoque de modelado de la visión:* También se conoce como enfoque psicológico [21], debido a que se basa en el análisis de varios componentes del sistema visual humano: contraste, percepción de color, y enmascaramiento temporal y espacial. Entre los algoritmos que trabajan bajo el esquema de modelado de la visión se encuentran: Visual Differences Predictor (VDP) [22], Sarnoff JND (Just Noticeable Differences) [23], Moving Quality Metric (MPQM), [24] y el Perceptual Distortion Metric (PDM) [25]. Los algoritmos basados en modelado de la visión presentan buen rendimiento ya que los modelos imitan el proceso de formación de la visión humana, pero debido a la complejidad de sus algoritmos su procesamiento computacional es elevado [26].
- *Enfoque de ingeniería:* Se basa en extraer los patrones de distorsión del video debidos a compresión y codificación. Entre los algoritmos más utilizados se encuentran: Structural Similarity Index (SSIM) [27]; Multi-scale SSIM [MS-SSIM] [28]; Speed SSIM [29]; MOtion-based Video Integrity Evaluation (Movie) [30]; Visual Signal to Noise Ratio (VSNR) [31] y Video Quality Metric (VQM) [32] [33] que es uno de los modelos más utilizados para evaluar la QoE del video, además VQM adoptado por el North American Estándar (ANSI T1.801.03-2003) y fue incluido en la ITU-TJ.144 e ITU-R BT.1883 [34].

Estos métodos basados en imagen presentan una desventaja ya que requieren tanto la señal fuente, como la señal recibida en el lado del cliente para su análisis. Esta situación hace que deba existir una sincronización entre el video enviado desde el emisor y el video recibido en el cliente, algo complejo de lograr si se desea hacer mediciones en tiempo real. Además algunos de los algoritmos mencionados, como por ejemplo SSIM y VQM, son computacionalmente complejos [26] pero debido a que están estandarizados ya se consiguen implementados en herramientas comerciales.

- **Métodos basados en paquetes**

Los métodos basados en paquetes trabajan con la información que se obtiene de los encabezados de los paquetes (encabezados de nivel 3 y 4), y con parámetros como: retardo y variación del retardo. Estos métodos se clasifican en cinco categorías [4]:

- *Modelos de nivel de red:* Permiten entregar una medición de QoE del video analizando información obtenida del video a nivel de red. En [35][36] se presentan modelos que trabajan bajo este esquema.
- *Modelos paramétricos de nivel paquete:* Entregan la medición de QoE del video mediante información que extraen de los encabezados de los

paquetes del video. En [7] presentan una propuesta que funciona bajo este metodología.

- *Modelos paramétricos de planeación*: Requieren parámetros de planeación para el cálculo de la QoE asociada al video. En [37] se presenta un modelo que funciona bajo este esquema.
- *Modelos de nivel de flujo de bits*: Utilizan la información de los encabezados del paquete y la información de carga útil para calcular la QoE del video. Recientemente en [38] proponen un modelo que trabaja bajo estas características.
- *Modelos híbridos*: Estos modelos basan su funcionamiento en combinar cualquiera de los modelos revisados anteriormente para obtener una medición de QoE del video. En [39] presentan ejemplos de modelos que funcionan bajo esta propuesta.

En la Figura 3 se observan métodos objetivos y su clasificación analizados en esta sección.

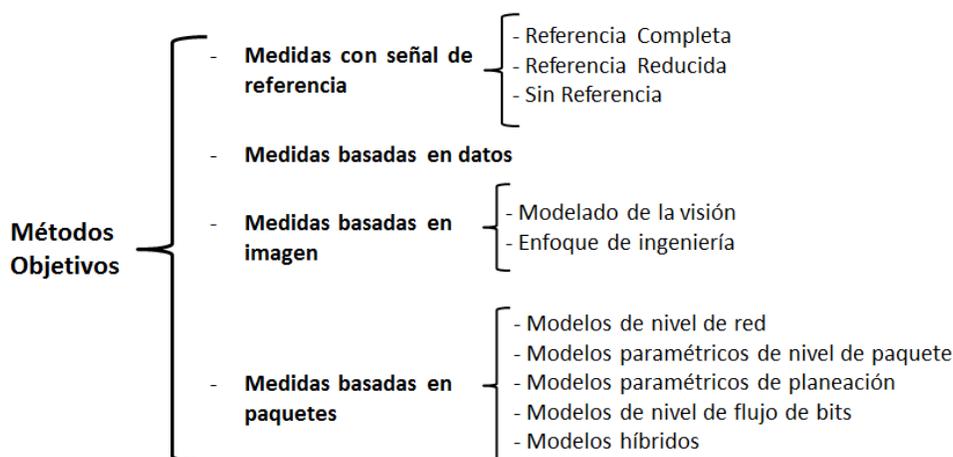


Figura 3. Métodos Objetivos para evaluar la calidad de la experiencia del video. Tomado de [9].

### 2.2.3 Modelos Indirectos

Los modelos indirectos buscan evaluar la QoE del video, mediante el uso de un modelo matemático. Este modelo matemático se obtiene variando uno o más parámetros de QoS con el fin de analizar la respuesta que entregue un modelo subjetivo u objetivo seleccionado para evaluar la QoE para un grupo de videos. Los métodos indirectos son útiles, porque permiten predecir la QoE percibida por el usuario en transmisiones en vivo y sin necesidad de recurrir a pruebas subjetivas.

Con el análisis presentado de los tres tipos de modelos en las secciones anteriores, se observa que cada tipo de modelo tiene sus ventajas y desventajas. Con los modelos subjetivos, se puede obtener la QoE real percibida por el usuario, pero la implementación de pruebas subjetivas es costosa, demanda tiempo, agrega incertidumbre y no es posible hacerlo con video en vivo. Con respecto a los modelos

objetivos, no reflejan la percepción del usuario, debido a que utilizan un número limitado de variables y, además algunos son computacionalmente muy complejos en su cálculo.

Por lo tanto, ante las desventajas que presentan los modelos subjetivos y objetivos, una solución que permite calcular la QoE percibida por un usuario, es mediante la utilización de modelos indirectos. Es por esta razón que múltiples autores –los presentados en los trabajos relacionados–, han planteado una serie de modelos matemáticos para resolver este problema. Cada modelo tiene una orientación que depende de la cantidad de parámetros de QoS que emplea y el método objetivo o subjetivo utilizado para evaluar la calidad de la experiencia.

Basándose en los modelos analizados en las secciones 2.2.1 y 2.2.2, la siguiente sección presenta el resultado de la revisión sistemática realizada a los trabajos relacionados propuestos por diferentes autores que plantean medir QoE utilizando modelos indirectos. Al finalizar la sección se presenta un cuadro comparativo entre diferentes modelos indirectos que permite dar claridad del proceso que siguió cada autor para obtener el modelo para medir la QoE.

## 2.3 Trabajos relacionados

Esta sección plantea un análisis de diferentes propuestas basadas en modelos indirectos, con el fin de entender el procedimiento que cada autor siguió para obtener un nuevo modelo para evaluar la QoE.

Los métodos indirectos relacionan parámetros de Calidad de Servicio -QoS- (retardo, variación del retardo, pérdida de paquetes u otros) con un modelo subjetivo u objetivo, para medir la QoE asociada al video. Esta relación del valor que pueden tomar los parámetros de QoS versus el valor obtenido de la QoE es expresada mediante un modelo matemático o un modelo de correlación.

Después de realizar el análisis de los trabajos relacionados se encontraron tres características que permiten compararlos. Estas características se describen a continuación:

- **Proceso utilizado para obtener los datos para generar el modelo:** Identifica que tipo de experimento se realizó para obtener los datos necesarios para generar el modelo. Es decir, si se utilizaron pruebas subjetivas, una herramienta de simulación o mediante el uso de una maqueta de experimentación en laboratorio.
- **Tipo de modelo utilizado para evaluar la calidad de la experiencia:** Identifica el tipo de modelo utilizado –subjetivo, objetivo, referencia completa, sin referencia, etc. – para evaluar la QoE del video.

- **Parámetros de QoS utilizados para generar el modelo:** Identifica cuáles fueron utilizados para generar el modelo.

En [40] y [41], se propone que la relación entre la calidad de servicio y la calidad de la experiencia está dada por la hipótesis IQX (Exponential Interdependency of Quality of Experience and Quality of Service). Esta hipótesis IQX mediante pruebas subjetivas analiza cómo se afecta la calificación del MOS ante la variación del parámetro pérdida de paquetes. En [42] y [43] involucran en su modelo tres parámetros de calidad de servicio - medidas de falla, medidas exitosas y medidas de recursos- para analizar con mayor precisión la sensibilidad del usuario, basándose en pruebas subjetivas realizadas a un conjunto de usuarios. La limitación de este modelo propuesto con su posterior mejora, radica en que para obtener los datos para generar el modelo, es necesario realizar pruebas subjetivas. Estas pruebas subjetivas, como se ha mencionado, son costosas en tiempo y logística, debido a que si se desea obtener una muestra representativa, se requiere realizar un número proporcional de pruebas a los usuarios.

En [44], [45] y [46], se evalúan cinco parámetros de calidad de servicio: pérdida de paquetes, nivel de ráfaga, variación del retardo, retardo y ancho de banda. Asignan a cada parámetro de QoS un peso relacionado con su grado de importancia relativa. Este grado de importancia relativa se calcula mediante los experimentos presentados en [47]. Emplean el valor de los parámetros de QoS en una ecuación lineal para obtener el valor de QoS asociado al video que posteriormente es evaluado en una ecuación exponencial con el fin de obtener el valor de QoE. En esta ecuación exponencial que evalúa la QoE involucran tres parámetros adicionales: tamaño de la pantalla, clase de servicio contratado por el usuario y un coeficiente determinado por la estructura de los cuadros de acuerdo con el grupo de imágenes (*Group of Picture - GoP*). En [44], [45] y [46] no se especifica cómo se obtiene la ecuación exponencial y además utilizan pruebas subjetivas para asignar a cada parámetro de QoS el peso relacionado con el grado de importancia del mismo.

En [48], los autores proponen un modelo híbrido sin referencia para estimar la calidad del video, utilizando información de los encabezados de los paquetes y de la señal de video. El modelo estima la degradación de la calidad del video debida a la compresión utilizada en su codificación y la pérdida de paquetes. El modelo se obtiene realizando estudios de correlación de datos obtenidos de la señal de video, analizando información espacial y temporal del mismo. El modelo se basa en pruebas subjetivas, realizadas a un conjunto de videos tomados como ejemplo, y solo analizan la calidad de la experiencia teniendo un solo parámetro de calidad de servicio: la pérdida de paquetes. Al igual que modelos presentados en los párrafos anteriores, el modelo propuesto en [48] requiere de pruebas subjetivas para obtener los datos y solo emplean un parámetro de calidad de servicio.

En [49], los autores proponen un esquema de predicción híbrido (Hybrid Quality of Experience [HyQoE]) que evalúa el nivel de calidad del video en tiempo real, basándose en resultados obtenidos a partir de pruebas subjetivas. El esquema está

basado en el aprendizaje estadístico de una red neuronal aleatoria (*Random Neural Network -RNN*) que se entrenó con videos reales generados sobre un escenario simulado inalámbrico. Aunque es una propuesta interesante, requiere del entreno previo de la red neuronal mediante el uso de los datos obtenidos mediante pruebas subjetivas.

En [50], los autores presentan un modelo que se basa en el uso de la función sigmoide [51] como base para estimar la QoE percibida por el usuario. La propuesta se basa en el uso de criterio descendente de la función sigmoide, donde la respuesta de la función disminuye, mientras el valor de entrada utilizado en la función sigmoide aumenta. Para este caso, el valor de entrada de la función sigmoide es la razón de pérdida de bytes. Así, ante mayor pérdida de bytes, menor valor de QoE percibida por el usuario. Para obtener la relación entre la razón de pérdida de bytes y la QoE se utilizó como herramienta de simulación NS-2 y para obtener la calidad de la experiencia percibida por el usuario se utilizó el índice Mean SSIM (MSSIM) [52]. Esta propuesta no requiere de un alto procesamiento computacional, pero solo analiza un parámetro de QoS para obtener la relación de la calidad de la experiencia percibida por parte del usuario.

En [53], los autores proponen cuantificar la calidad de la experiencia utilizando varios parámetros de calidad de servicio, basándose en la hipótesis IQX propuesta en [40] y [41]. En [53] utilizan una regresión lineal entre las diferentes ecuaciones generadas por cada parámetro de QoS para obtener el modelo. Este modelo utiliza más de un parámetro para evaluar la QoE, pero se basa en la expresión en hipótesis IQX obtenida mediante pruebas subjetivas.

En la Tabla 1 se presenta la comparación entre las propuestas analizadas basándose en los aspectos explicados al inicio de esta sección: proceso utilizado para obtener los datos para generar el modelo, tipo de modelo utilizado para evaluar calidad de la experiencia y parámetros de QoS utilizados para generar el modelo.

Tabla 1. Cuadro comparativo de trabajos relacionados que utilizan modelos indirectos en sus propuestas.

<b>Autores</b>	<b>Proceso utilizado para obtener los datos para generar el modelo</b>	<b>Tipo de modelo utilizado para evaluar calidad de experiencia</b>	<b>Parámetros de QoS utilizados para generar el modelo</b>
Hoßfeld et al [40] y Fiedler et al.[41]	Plataforma de experimentación en laboratorio	Pruebas subjetivas	Pérdida de paquetes Variación del retardo.
Kim et al.[44], [45] y [46]	No está definido	Ambiente de computación invasiva	Pérdida de paquetes, nivel de ráfaga, variación del retardo de paquete, retardo de paquete y ancho de banda.
Yamagishi et al.[48]	Plataforma de experimentación en laboratorio	Pruebas subjetivas	Pérdida de paquetes

Autores	Proceso utilizado para obtener los datos para generar el modelo	Tipo de modelo utilizado para evaluar calidad de experiencia	Parámetros de QoS utilizados para generar el modelo
Riker et al.[49]	Utilizan una red Neuronal aleatoria	Entrenan a la red con videos evaluados subjetivamente	No utiliza. La Red Neuronal se entrena con videos previamente evaluados
Abdeljaouad et al.[50]	Simulador NS-2	Función Sigmoide	Proporción de pérdida de bytes.
Aroussi et al.[53]	Plataforma de experimentación en laboratorio	Pruebas subjetivas	Retardo y pérdida de paquetes

Basándonos en la tabla anterior, se puede apreciar que el 60% de las propuestas analizadas evalúan la QoE del video con ayuda de pruebas subjetivas, con el fin de relacionarlas con los parámetros de QoS y obtener un modelo.

Los modelos propuestos en la Tabla 1 presentan una precisión alta (correlación de resultados comparados con pruebas subjetivas superior a 75%) al momento de evaluar la QoE, pero el gran inconveniente que tienen, es la realización de las pruebas subjetivas. Como se ha mencionado anteriormente, realizar pruebas subjetivas es costoso y requiere de una logística compleja al momento de realizar la prueba.

Por otra parte, en [40], [41], [42] y [43], en sus propuestas utilizan una serie de constantes que se ajustan solamente a los experimentos diseñados para generar sus modelos y en algunos casos no especifican de manera clara el proceso para obtener dichas constantes.

Después de analizado lo anterior, y ante el costo en tiempo y logística de realizar pruebas subjetivas, nuestra propuesta se basará en la elección de un modelo objetivo para evaluar la QoE y se seleccionarán tres parámetros de QoS para obtener los datos para generar la propuesta. La justificación sobre la elección del modelo objetivo y los parámetros de QoS elegidos, se presentará en la sección 3.1 y sección 3.2.

Y con respecto a propuestas recientes como la que se presenta en [54] [55] [56], plantean esquemas eficientes de entrega de video P2P (punto a punto), basado en el análisis del comportamiento del usuario. Esta propuesta es novedosa porque permite el análisis del comportamiento del usuario (permitiendo así la mejora de QoE al ofrecer contenido de acuerdo con los requisitos del usuario) mediante el mapeo de dos indicadores (latencia de emisión y carga del servidor) para medir la eficiencia de los sistemas de entrega de video. Actualmente propuestas de modelos de QoE se están enfocando en tecnologías móviles (como Wifi y 5G) debido al alto crecimiento de este tipo de usuarios que crecerá un 46% entre el 2015 y 2020 [2]. Al enfocarse en este tipo de tecnologías los parámetros para analizar la QoE asociada al video cambian como se presenta en [57], donde ya se tienen en cuenta parámetros como el retardo de reproducción del video, congelamiento del video después de reproducirlo y cambio de resolución del video en su reproducción.

Es importante mencionar que en [34], los autores realizan un planteamiento similar al que se presentará en este proyecto. La gran diferencia radica que en [34] los autores plantean un modelo solo utilizando dos parámetros de calidad de servicio ( variación

del retardo y pérdida de paquetes), dejando fijo el retardo a 100mseg. Por otro lado utilizan para la pruebas, videos de calidad estándar de resolución 480x270 y presentan un único modelo para un video con movimiento medio. Además no presentan pruebas de validación estadística al modelo propuesto, ni comparan el modelo propuesto con resultados de modelos objetivos estandarizados.

El modelo a proponer en este proyecto mejora la propuesta planteada en [34] y en los autores analizados en la Tabla 1 en los siguientes aspectos:

- Para obtener datos para generar el modelo se utilizarán videos en alta definición de resolución 1080x1920.
- Se trabaja con tres parámetros de calidad de servicio simultáneamente (retardo, variación del retardo y pérdida de paquetes).
- Al modelo obtenido se le realizan pruebas estadísticas y los resultados se comparan con modelos objetivos estandarizados.

## 2.4 Comentarios del Capítulo

Es claro resaltar, que la revisión del estado del arte permitió identificar la metodología a seguir en el diseño del experimento con el fin de obtener los datos necesarios para la generación del modelo a proponer.

Y como se explicó anteriormente, la propuesta presentada en este trabajo es la única con este tipo de enfoque que será muy útil para los proveedores de servicio a fin de estar analizando la calidad de la experiencia del servicio que ofrecen a sus usuarios.

Para finalizar producto de la revisión del estado del arte presentada en este capítulo, se publicó el artículo:

- **J. C. Cuéllar**, Jesus Hamilton Ortiz, and J. L. Arciniegas, “**Clasificación y Análisis de Métodos para medir Calidad de la Experiencia del Servicio de Televisión sobre Protocolo IP (IPTV)**,” *Información Tecnológica*, Chile. vol. 25, no. 5, pp. 121–128, Enero 2014. [9]

Y se participó en los siguientes eventos:

- **Congreso Nacional:**  
**J.C. Cuéllar**, Jesus Hamilton Ortiz y J.L. Arciniegas. **Modelo Híbrido para Medir QoE al Servicio de IPTV.** Octavo Congreso Colombiano de Computación. Universidad del Quindío. Agosto 21 al 23 de 2013.

**Spring School on Networks**

**Poster: Propuesta Metodológica para Implementar un Modelo para Medir Calidad de la Experiencia para el Servicio de IPTV.** J.C. Cuéllar, Jesus Hamilton Ortiz and J.L. Arciniegas. Santiago de Chile Noviembre 9 y 10 de 2015.

**Hybrid Model to Measure QoE in IPTV: Methodological Proposal**

J.C. Cuéllar, Jesus Hamilton Ortiz and J.L. Arciniegas

**VII Congreso Iberoamericano de Telemática CITA 2015.**

ISBN-13: 978-0-9898393-8-9. Junio 10 al 12 de 2015. Popayán, Colombia.



### 3 Componentes para generar los datos y obtener el modelo para medir QoE para el servicio de IPTV

En este capítulo se presentarán y describirán los componentes para generar los datos para obtener el modelo para medir QoE. Se empieza identificando los parámetros de QoS seleccionados, posteriormente se explica el modelo objetivo utilizado para evaluar la calidad de experiencia, se explican las especificaciones técnicas de la maqueta de experimentación. En la siguiente sección se describen los videos fuentes a utilizar en el experimento y al finalizar se presentan comentarios sobre los resultados obtenidos en el desarrollo de este capítulo. En la Figura 4, se ilustra el contenido esquemático de este capítulo.

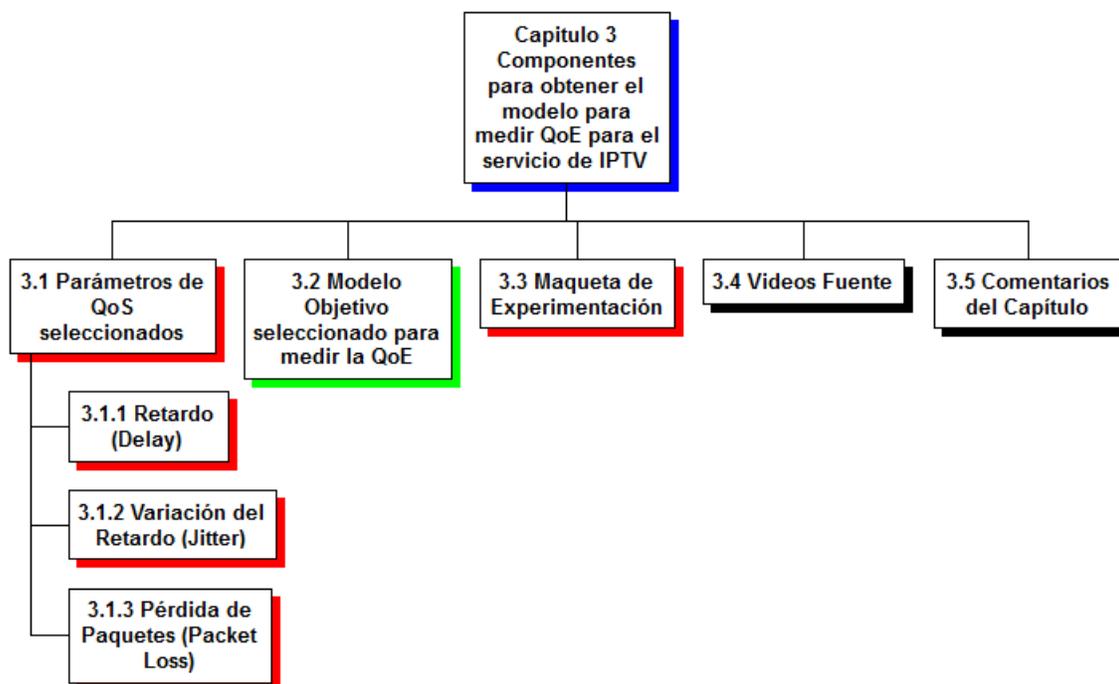


Figura 4. Descripción del contenido del capítulo 3.

## 3.1 Parámetros de calidad de servicio seleccionados

En esta sección se explicará que parámetros de QoS fueron identificados y elegidos para la recopilación de datos para la generación del modelo, partiendo desde su definición hasta explicar qué tipo de distorsión se ve reflejada en el video.

Después de analizar los diferentes modelos propuestos en la sección 2.3 página 30, se eligieron los siguientes parámetros de QoS : retardo (*Delay*), variación del retardo (*Jitter*) y pérdida de paquetes (*Packet loss*), estos parámetros se eligen porque según la revisión sistemática realizada son lo que más influyen en la calidad del video [17].

### 3.1.1 Retardo (*Delay*)

Definido como *IP Packet Transfer Delay*, hace referencia al tiempo que tarda el paquete en pasar por un componente de la red (computador, dispositivo de interconectividad o sección de red) [58]. Cuando el retardo se incrementa a niveles por fuera de la Recomendación UIT-T 1541 (Rango de 0 – 400 msec) [59], la imagen se empieza a ver borrosa (*image pixelation*) y su definición es deficiente [60]. Una solución a este problema es configurar un tamaño de memoria intermedia (*buffer*) adecuado para la reproducción del video en el usuario final.

### 3.1.2 Variación del Retardo (*Jitter*)

Definido como *IP Packet Delay Variation*, hace referencia al tiempo esperado de llegada de cada paquete [58]. El incremento del valor de este parámetro por fuera de los rangos definidos en la Recomendación UIT-T 1541 (Rango de 0 a 50msec) [59], hace que el video se congele y pierda definición [60]. Una solución a este problema es configurando un tamaño de memoria intermedia (*buffer*) adecuado, pero esto puede afectar el retardo.

### 3.1.3 Pérdida de paquetes (*Packet Loss*)

Definido como *IP Packet Loss Ratio*, hace referencia a la tasa de pérdida de paquetes, se obtiene entre el total de paquetes perdidos sobre el total de paquetes transmitidos [58]. Es el parámetro que más influye en la calidad del video [60], causando congelamiento y saltos en el video (Rango de 0 a  $1 \times 10^{-3}$ ). Esto debido a que si hay pérdida de cuadros (*frames*), estos se reemplazan con cuadros previamente decodificados, haciendo que la calidad del video se deteriore.

De acuerdo a las definiciones anteriores, si solo en la transmisión del video se presenta retardo (sin variación del retardo o pérdida de paquetes) el video puede reproducirse lentamente (*smoothly*) con la ayuda de la memoria intermedia (*buffer*). Con la pérdida de paquetes, el cuadro (*frame*) que llegue al usuario se congelará haciendo que haya un salto en la imagen. Esta pérdida de paquetes se puede compensar retransmitiendo el cuadro perdido, pero hará que se incremente el retardo y la variación del retardo. Por su parte, la variación del retardo afecta la reproducción del video presentando congelamiento en la imagen, mientras se espera la llegada de un cuadro tardío. Eligiendo un tamaño adecuado de la memoria intermedia (*buffer*) la variación del

retardo puede disminuir, pero el valor del retardo aumentará. La variación del retardo puede tener casi el mismo efecto en la QoE del video, que la pérdida de paquetes [15].

Con estos parámetros de QoS elegidos, el objetivo específico No.1 se cumple parcialmente: *“Identificar y seleccionar los parámetros y las métricas asociados a QoS, para evaluar la calidad del servicio de IPTV en un entorno residencial”*. En la sección 4.1.4 se presentarán las métricas asociadas a cada parámetro y como fue el proceso para su definición.

Ya con los parámetros seleccionados se explicará, en la siguiente sección, el modelo objetivo de calidad de experiencia elegido para la toma de datos.

## **3.2 Modelo objetivo seleccionado para medir la calidad de la experiencia**

En esta sección se justificará la elección del modelo para evaluar la calidad de experiencia del video y se explicará las dificultades técnicas que se afrontaron durante el desarrollo de esta parte del proyecto.

El objetivo específico No.2, se planteó al inicio del proyecto como: *“Seleccionar y adaptar un algoritmo Sin Referencia (No Reference) para la medición de la calidad de video con el fin de obtener una medida de la QoE percibida por el usuario”*.

Como se presentó en la sección 2.2.2, un modelo sin referencia es muy útil para realizar mediciones en tiempo real, ya que no se requiere la señal de video origen para el cálculo de la calidad de la experiencia.

En la exploración bibliográfica que se realizó al inicio del proyecto, la Recomendación UIT-T P.1202.2 [61] presenta un modelo sin referencia para evaluar la QoE del video. Esta Recomendación especifica los procedimientos para monitorear la calidad del video sobre IP, basándose en la información de los encabezados de los paquetes del flujo de bits (*bitstream*) del video. La Recomendación indica que se puede aplicar para servicios como IPTV, y posee dos modos de operación: Modo 1, donde el video se analiza sin decodificarlo y Modo 2, donde el video se decodifica totalmente con el fin de hacer un análisis a nivel de píxeles. La Recomendación UIT-T P.1202.2 posee 4 bloques funcionales: (i) Extracción de parámetros básicos, (ii) Agregación de parámetros básicos a nivel interno de cuadro, (iii) Agregación de parámetros dentro del modelo de parámetro y (iv) Modelo de estimación de la calidad.

Con el análisis realizado, en su momento a la Recomendación P.1202.2, se decide su implementación debido a que se ajustaba al objetivo específico No.2 que se planteó al inicio del proyecto. Adicionalmente la información que requería cada bloque funcional de la Recomendación P.1202.2 se podía obtener analizando los encabezados de los paquetes del flujo de bits del video como se explicaba en la recomendación. Bajo este supuesto se inició su implementación y en los siguientes párrafos se explicará el proceso que se llevó a cabo.

Durante este proyecto se inició la implementación del bloque, (i) Extracción de parámetros básicos. Para esto en la maqueta de experimentación que se diseñó e implementó, se transmitieron una serie de videos utilizando VLC [62] y capturando el tráfico con ayuda del analizador de tráfico Wireshark [63].

Con el video capturado, se empezó el proceso de decodificación de la cabecera de cada protocolo (Ethernet, IP, UDP, RTP, MPEG-2-TS, PES, H.264) que llevaba empaquetado, y se analizó la carga útil de H.264. En la Figura 5 se aprecia la carga útil de un paquete H.264 hasta el nivel de macrobloque. Realizar esta decodificación hasta este nivel, permitirá obtener información útil para la implementación del bloque de extracción de parámetros de la Recomendación P.1202.2.

A Nivel de Abstracción de Red ( NAL –*Network Abstraction Layer*) se encontró que existían parámetros requeridos por la Recomendación P.1202.2 que se hallaban codificados entrópicamente, según CABAC (Context Adaptative Binary Arithmetic Coder)[64] y CAVLC (Context-Adaptative Variable Length Code) [65]. La codificación entrópica [66] trabaja con probabilidades y tablas de referencia en donde las muestras de subbloques de los macrobloques se encuentran codificadas, en un conjunto de elementos que permiten su posterior decodificación. La interpretación de esta información es necesaria para poder identificar distorsiones en cada cuadro del video.

Llegar a este nivel, para extraer parámetros a usar en la recomendación P.1202.2, se debía identificar como el codificador generaba la codificación entrópica. Por lo cual se hizo necesario analizar el funcionamiento del codificador a nivel de bit, con el fin de obtener los datos requeridos para este módulo. A pesar de que se realizó ingeniería inversa al codificador, no fue posible entender el funcionamiento del algoritmo que generaba esta información. Esta dificultad técnica hacía que el cálculo de la QoE no fuera confiable por no tener certeza de la información obtenida en la implementación de este módulo.

Sin lograr decodificar esta información no se logró implementar la estructura de datos que requería la Recomendación P1202.2 a nivel de cuadro. Esta estructura de datos posteriormente se usaría como entrada para el bloque funcional (ii) Agregación de parámetros básicos a nivel interno de cuadro.

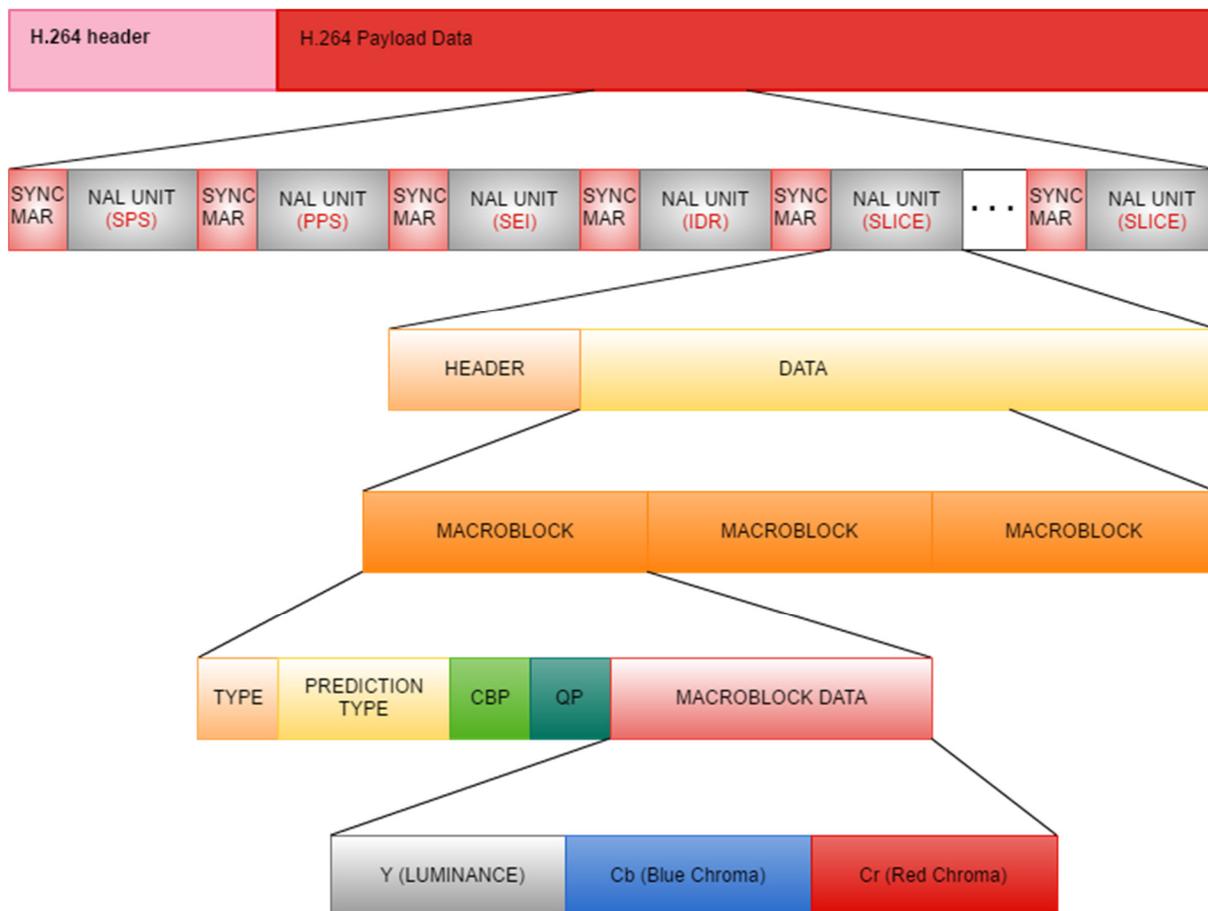


Figura 5. Encapsulamiento de H.264.

Basándose en lo anterior, implementar la Recomendación P.1202.2 no era viable para el proyecto, debido a la dificultad encontrada al momento de identificar parámetros necesarios para su funcionamiento. Adicionalmente diseñar e implementar un codificador propio, técnicamente no era factible, debido a que los datos para proponer el modelo se iban a realizar utilizando videos codificados en H.264, estándar de amplio uso para la codificación de video.

Posteriormente ante la serie de dificultades técnicas explicadas en la implementación de la Recomendación P.1202.2 se realizó una nueva revisión del estado del arte y trabajos relacionados para encontrar referentes que estuvieran trabajando en el tema, pero específicamente en métodos sin referencia. Producto de esta revisión sistemática se encontró el modelo *Video Intrinsic integrity and distortion evaluation oracle* (VIIDEO) propuesto en [67]. VIIDEO incorpora modelos de estadísticas intrínsecas que permiten cuantificar las distorsiones de los videos (congelamiento, baja definición, etc). VIIDEO Maneja un algoritmo capaz de predecir la calidad del video sin necesidad de analizar el video original.

Después de realizar una serie de pruebas con diferentes videos de bases de datos, utilizando videos propios y analizando la estructura del algoritmo se encontraron los siguientes inconvenientes:

- ✓ VIIDEO fue diseñado para analizar videos con una resolución de 432x768.
- ✓ En las pruebas presentadas en [67] utilizaron **LIVE Video Quality Assessment Database** [68], en esta base de datos, los videos ya presentaban distorsiones de cuatro maneras diferentes: distorsión debido al proceso de codificación con MPEG y H.264, distorsión reflejada en errores debido a la trasmisión del video sobre IP y distorsión debido a la transmisión del video sobre redes inalámbricas. No estaba documentado el proceso que se realizó para distorsionar los videos.
- ✓ Se intentó modificar el algoritmo para adecuarlo a la resolución (1920x1080), resolución que sería utilizada en las pruebas, pero los resultados obtenidos no fueron coherentes, es decir las medidas no eran homogéneas para un video con distorsiones, en algunos casos, se obtenía la misma medida que para un video con excelente calidad.
- ✓ El video para ser analizado por VIIDEO debía estar en formato YUV. Este formato presenta dos inconvenientes técnicos: (i) el tamaño del archivo a analizar es del orden de 1 a 2 GB, lo cual hace el tiempo de procesamiento sea muy elevado (alrededor de 40 minutos por video), y que el procesamiento en los experimentos tome mucho tiempo en su desarrollo. (ii) Cuando un video distorsionado se intentaba convertir a YUV, la herramienta empleada (ffmpeg) en algunas ocasiones no podía realizar la conversión, esto porque al perderse cuadros del video no se podía codificar de nuevo en otro formato, Ffmpeg consideraba que el archivo origen estaba defectuoso y no lo procesaba
- ✓ Adicionalmente VIIDEO [67], no documentaba los rangos de las medidas obtenidas de QoE, ni se presentaba un mapeo con la escala MOS. Esto es fundamental con el fin analizar el desempeño de VIIDEO al compararlo con modelos objetivos estandarizados.

Basándonos en lo anterior, no fue posible adaptar el modelo VIIDEO para trabajar con los videos de prueba que fueron grabados a una resolución de 1920x1080 pixeles.

Por las consideraciones anteriores, y en vista que los modelos sin referencia están en fase de desarrollo y hasta el momento no hay modelos que se utilicen de manera estandarizada, se decide identificar un modelo objetivo para realizar las medidas de calidad de experiencia asociadas al video.

Para esto se identifican los algoritmos presentados en la sección 2.2.2: Peak-Signal-to-Noise-Ratio (PSNR), Structural Similarity Index (SSIM), Multi-scale SSIM (MS-SSIM), Video Integrity Evaluation (MOVIE), Visual Signal to Noise Ratio (VSNR) y VQM (Video Quality Metric), con el fin de elegir el algoritmo que se ajuste a las necesidades planteadas en este proyecto.

Después de analizar lo presentado en [69] y [15], y teniendo en cuenta las cuatro métricas de comparación definidas por el Grupo de Expertos de Calidad del Video (*Video Quality Expert Group –VQEG*): (i) Coeficiente de Correlación de Pearson (CCP), (ii) Raíz del Error Cuadrático Medio (RECM), (iii) Coeficiente de correlación de Spearman (CCS), (iv) Porcentaje de Valores atípicos (PVA), que se explicarán en detalle en la sección 5.3.1. Se elige VQM como modelo objetivo para realizar las mediciones de calidad de experiencia asociado al video, debido a que presenta valores

más elevados de CCP y CCS con respecto a los otros algoritmos analizados y los valores RECM y PVA son los más bajos con respecto a los otros algoritmos [69].

Adicionalmente VQM, trabaja con más parámetros de medición que los otros algoritmos (siete, que se explicarán a continuación) y es un modelo estandarizado que utiliza bajo ancho de banda y que extrae datos de los videos origen y destino, fue adoptado por el North American Estándar (ANSI T1.801.03-2003) y fue incluido en la ITU-TJ.144 e ITU-R BT.1883 [34]. Adicionalmente VQM es un método utilizado ampliamente en la comunidad científica para validación y comparación de nuevos modelos propuestos.

El funcionamiento de VQM es basado en modelos de referencia reducida que extraen parámetros de las regiones espacio-temporales de la secuencia de video. VQM requiere un ancho de banda de canal auxiliar de 9.3% del video sin comprimir y requiere de técnicas de calibración asociadas que necesitan de un 4.7% de ancho de banda adicional.

VQM maneja siete parámetros independientes, cuatro de ellos extraen gradientes espaciales del componente de luminancia Y, dos parámetros son basados en características del vector formado por dos componentes de crominancia ( $C_B$ ,  $C_R$ ) y un parámetro obtenido por el producto de características que miden contraste y movimiento, ambas son extraídas del componente de luminancia Y. La descripción de los siete parámetros se presenta a continuación, pero para profundizar en el funcionamiento de VQM se recomienda consultar la referencia [70]:

- Parámetro **si\_loss**: Este parámetro detecta una disminución o pérdida de la información espacial, como por ejemplo definición defectuosa (*Blurring*).
- Parámetro **hv\_loss**: Este parámetro detecta un desplazamiento de los bordes de orientación horizontal y vertical hacia la orientación vertical, esto podría ocurrir si los bordes horizontales y verticales sufren definición defectuosa que los bordes diagonales.
- Parámetro **hv\_gain**: Este parámetro detecta un desplazamiento de los bordes diagonales a bordes horizontales y verticales, como es el caso cuando la imagen presenta falta de definición en sus bordes y se ve defectuosa.
- Parámetro **chroma\_spread**: Este parámetro detecta cambios en la dispersión de la distribución de muestras de color de dos dimensiones.
- Parámetro **si\_gain**: Este es el único parámetro que mide aumento en la calidad del video, midiendo la ganancia de calidad del video mediante el análisis de los bordes de la imagen.
- Parámetro **ci\_ati\_gain**: Este parámetro es calculado como el producto de la característica de contraste, midiendo la cantidad de detalle espacial y la característica de información temporal.
- Parámetro **chroma\_extreme**: Este parámetro detecta errores severos en el color de la imagen causados por errores en la transmisión.

Estos siete parámetros se combinan en la Ecuación 1, permitiendo obtener la medida de la calidad del video, donde una medida cercana a cero indica un video de mejor calidad que un video con una medida cercana a 1 o superior:

$$\begin{aligned}
 VQM = & -0.2097 * si_{loss} + 0.5969 * hv_{loss} + 0.2483 * hv_{gain} + 0.0192 \\
 & * chrom_{spread} - 2.3416 * si_{gain} + 0.0431 * ct_{atigain} \\
 & + 0.0076 * chrome_{extreme}
 \end{aligned}$$

Ecuación 1

### 3.3 Maqueta de Experimentación

En esta sección se presentará el diseño y características técnicas de la maqueta de experimentación (*testbed*) para lograr el cumplimiento a la segunda parte del objetivo específico No.2: “...Identificar e implementar una arquitectura de pruebas en laboratorio, con el fin de obtener los datos necesarios para la generación del modelo”.

En la Figura 6 se puede apreciar la maqueta de experimentación implementada, que consiste en la interconexión de un servidor y cliente VLC Media Player (versión 2.2.2 Wheathenrwx [62]), a través de una máquina (Linux Server NetEm) con dos tarjetas de red y con la herramienta Network Emulator (NetEm) [71] configurada. Las especificaciones de las máquinas utilizadas se pueden apreciar en la Tabla 2.

Esta maqueta se eligió, debido a que en la revisión del estado de arte presentada en la sección 2.3 trabajos relacionados (Página 30), es la manera que diversos autores realizaron sus experimentos a fin de proponer nuevos modelos [72] [73] [74] .

Por otra parte, se decide utilizar todo el ancho de banda de los enlaces Ethernet con el fin de minimizar las variables no controladas en el experimento. Es por esta razón, que no hay aplicaciones adicionales compartiendo en el ancho de banda en la maqueta de experimentación implementada. El servicio de IPTV emulado, es un servicio no gestionado que va sobre protocolo UDP, sin configurar calidad de servicio y sin asignar ancho de banda.

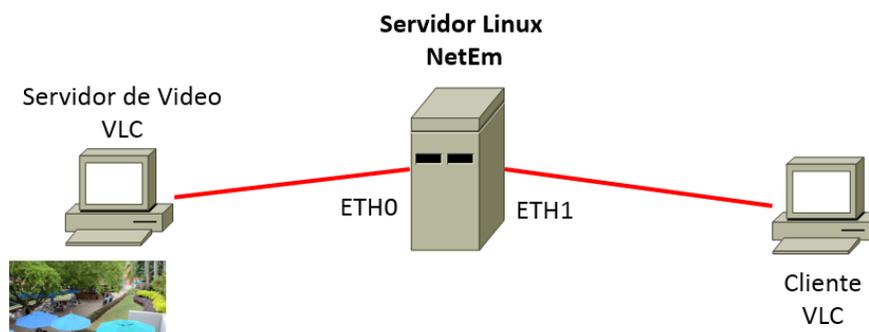


Figura 6. Maqueta de experimentación implementada para obtener los datos para generar el modelo.

Tabla 2. Especificaciones técnicas de las máquinas utilizadas en la maqueta de experimentación.

Maquina	Especificaciones técnicas
Servidor de Video VLC	Procesador: Intel(R) Core™ i3 2330M CPU 2.20GHz Memoria RAM: 8.00 GB Sistema operative: Win 7 – 64 bits
Cliente VLC	Procesador: Intel(R) Core™ i7 870@ 2.93GHz Memoria RAM: 4.00 GB Sistema operative: Win 7 – 64 bits
Network Emulator (NetEm)	Procesador: Intel(R) Core™ 2 Duo CPU 3GHzx2 Memoria RAM: 2.00 GB Sistema operative: Ubuntu 14.04 LTS

Para emular un enlace de red de área amplia entre cliente y servidor se eligió la herramienta NetEm [71]. NetEm es un emulador que permite variar los parámetros de QoS escogidos y adicionalmente como lo plantean en [75], este emulador presenta un buen desempeño y es una herramienta poderosa para trabajar en la emulación de este tipo de ambientes. La configuración utilizada en VLC y NetEm se presenta en el Anexo 1, al final del documento.

Los detalles de los rangos de los parámetros de QoS utilizados en el experimento se explicarán en el capítulo 4, sección 4.1.5, página 62.

### 3.4 Videos Fuente

En esta sección se explicará el proceso desarrollado para seleccionar los videos fuente que se utilizaron para la generación de los datos. Para realizar esta selección, se debe analizar cuidadosamente dos aspectos: contenido y características espacio-temporales del video.

En lo relacionado con el contenido, es importante subrayar que el contenido no debe ser ofensivo, controversial o que cause distracción en el observador [75].

En el aspecto técnico se deben elegir las secuencias de video analizando las características espacio-temporales del video, estas características permiten identificar que tanto movimiento y detalle en los cuadros presenta el video.

La Recomendación P.910 [14] define estas características de la siguiente manera:

- **Información perceptual espacial (*Spatial Information - SI*):** Es una medida que indica la cantidad de detalle espacial de la imagen. Por lo general, es más alto para escenas complejas, es decir, escenas con más objetos o bordes en los cuadros (*frames*) de una secuencia resultarán en valores más altos de SI.
- **Información perceptual temporal (*Temporal Information - TI*):** Es una medida que indica la cantidad de cambios temporales de una secuencia de video. Por lo general, la medida de TI es mayor para las secuencias de alto movimiento.

Ahora bien, para calcular el valor de la información perceptual espacial (IE) se pasa por cada cuadro del video, un filtro Sobel [76], que indica de manera general la cantidad de detalles espaciales de una imagen, el filtro se pasa sobre el valor de luminancia de cada cuadro del video midiendo la desviación estándar del cuadro resultante y luego seleccionando la desviación estándar máxima. En la Ecuación 2 se puede apreciar la expresión que permite calcular IE:

$$IE = \max_{time}\{std_{space}[Sobel(F_n)]\} \quad \text{Ecuación 2}$$

El operador de Sobel calcula el gradiente de cada píxel en función de su intensidad. Esto muestra cómo cambia brusca o suavemente una imagen en cada punto, lo que permite detectar las partes de mayor contraste, o bordes, de una imagen. El filtrado funciona convolucionando dos operadores sobre el marco de video y tomando la raíz cuadrada de la suma de los cuadrados de estas circunvoluciones. Los altos valores de IT en una imagen se interpretan como que tienen más contenido.

El cálculo de la información perceptual temporal (IT) es una medida basada en la diferencia entre los valores de píxeles en la misma ubicación en el espacio, pero en tiempos o marcos sucesivos. Por lo general, indica la cantidad de cambios temporales de una secuencia de video. Suele tener valores más altos para secuencias que representan movimiento elevado. Al igual que con la IE, IT tiene la intención de detectar cambios en cada cuadro posterior, mediante una sustracción píxel a píxel de dos cuadros continuos en el nivel de luminancia, calculando la desviación estándar del cuadro resultante y luego seleccionando la desviación estándar más alta. En la Ecuación 3 se puede apreciar la expresión para calcular IT.

$$IT = \max_{time}\{std_{space}[F_n(i,j) - F_{n-1}(i,j)]\} \quad \text{Ecuación 3}$$

De acuerdo a las definiciones anteriores, en la Recomendación P.910 utilizaron para sus pruebas, una serie de videos identificados con letras de la “a” a la “x” (Ver Figura 7) con diferentes características espacio-temporales. Dependiendo de estas características se puede definir qué tanto movimiento presenta el video. A manera de ejemplo en la Recomendación P.910 presentan la ubicación espacio-temporal para cada uno de estos videos de prueba.

En la Figura 7, en el eje unidad de **información temporal (IT)**  $IT < 20$ , se encuentran videos con escenas con movimiento limitado, como **l, f y a**. En la parte de arriba del eje,  $IT > 50$ , se encuentran videos con movimiento alto como **p, q e i**.

En el eje unidades de **información espacial (IE)**,  $IE < 50$  se encuentran videos con escenas con mínimo detalle espacial, como **l,k,x,u y f**. En la parte derecha de la gráfica con  $IE > 200$  se encuentran videos con escenas con bastante detalle espacial, como **h y s**.

Este tipo de representación de los videos mediante una gráfica IT vs IE (Figura 7) permite de manera eficaz y rápida deducir la cantidad de movimiento y definición que posee el video.

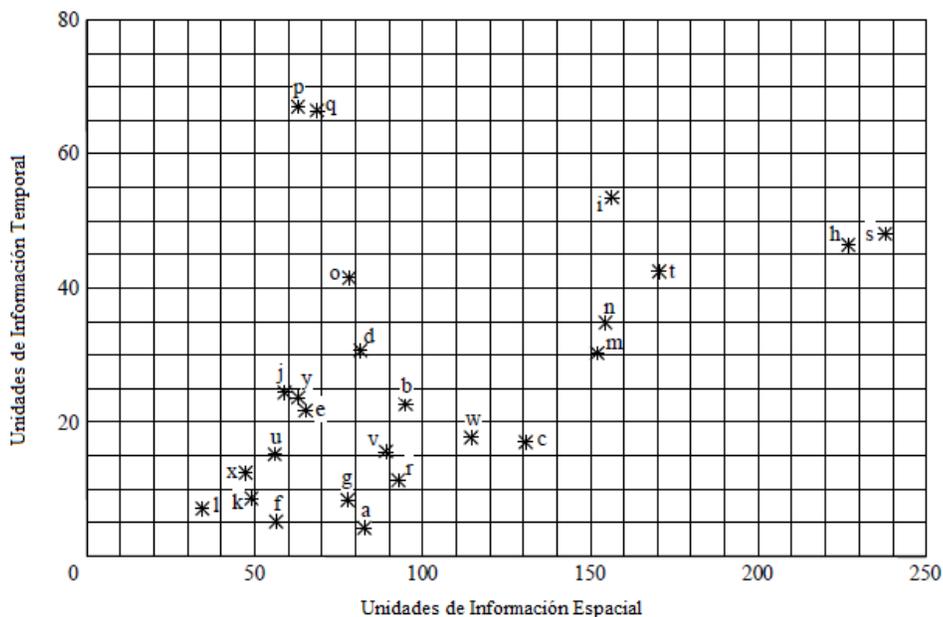


Figura 7. Ejemplos de videos de prueba con sus características espacio-temporales. Tomado de [14].

Ya definidas las características espacio-temporales, se pasó a obtener las secuencias de video a utilizar en el proyecto. Para empezar se realizó una exploración en las diferentes bases de datos de videos y elegir las secuencias para el experimento. Cuando la exploración se realizó se encontraron las siguientes dificultades:

- La gran mayoría de las bases de datos solo tenían secuencias en definición estándar, para nuestro proyecto se requería de alta definición.
- Las secuencias de video de las bases de datos consultadas tenían una duración promedio de 10 a 12 segundos, lo cual impedía su utilización. Esto debido a que en la maqueta de experimentación se debían utilizar videos con una duración de aproximadamente 30 segundos, con el fin de lograr en el lado cliente su decodificación adecuada y su posterior edición.
- En algunas bases de datos ya se les había afectado la calidad a los videos, es decir, eran videos con distorsiones de diferente tipo y los autores no describían de qué manera se había realizado este proceso.

Después de lo arriba mencionado se decidió generar de manera propia las secuencias de video, realizando grabaciones con una cámara digital Nikon D3100 (single-lens reflex). Cada video fue grabado con una resolución de 1920x1080, 24 imágenes por segundo o cuadros por segundo (*frame rate* - 23.976 fps) y codificado en H.264/MPEG-4 en formato MOV (Formato de video y audio creado por Apple). Cada toma se realizó por más de 30 segundos.

En la Tabla 3 se pueden apreciar los tipos de videos elegidos para la recolección de datos. Se presenta el nombre del video, una captura de pantalla (*screenshot*) del mismo, sus características espacio-temporales y la clase de video.

El video Entrevista codificado como **I00**, como su nombre lo indica es una entrevista que está realizando a una persona. Aquí el movimiento es mínimo ya que la persona solo contesta las preguntas y mira a la cámara. Las manos y el cuerpo los mantiene sin movimiento.

El video Cafetería, codificado como **C00**, es un video grabado a personas que van caminando por debajo de los árboles y algunas se detienen a conversar o están haciendo uso de las mesas, la cafetería está ubicada en la primera planta y el video se tomó desde la planta superior.

El vídeo Baloncesto codificado como **B00**, es un entreno de baloncesto, se grabaron jugadores corriendo a lo largo de la cancha.

Se eligieron estos tres tipos de videos, ya que de acuerdo a la Recomendación de la ITU-T P.910 agrupan la descripción descrita en el anexo 1, cuadro A.1 de dicha recomendación, de la siguiente manera: Video Entrevista: Categoría A, Video Cafetería: Categoría B y C, y Video Baloncesto: Categoría D y E. Esto porque los videos definidos para el experimento se ajustan a las características espacio-temporales de cada clase descrita en la Recomendación ITU-T P.910, y en el caso del video Cafetería y Baloncesto, cumplen con la descripción de dos categorías aproximadamente.

Tabla 3. Características espacio temporales de los videos usados en la maqueta de experimentación.

Nombre del video	Inf. Espacial	Inf. Temporal	Clase
<b>Entrevista</b> (I00)		35,9166	5,38516 Movimiento Bajo
<b>Cafetería</b> (C00)		65,9848	6,48074 Movimiento Moderado

Nombre del video		Inf. Espacial	Inf. Temporal	Clase
Baloncesto (B00)		57,41.95	17,4069	Movimiento Alto

En la Figura 8 se presenta la distribución espacio-temporal de los tres videos utilizados en el experimento, Cafetería(C00), Entrevista(I00) y Baloncesto(B00). Estos videos son los videos fuente utilizados en el experimento sin ningún tipo de distorsión.

En el video Entrevista(I00), la característica  $SI=35,91$  indica un detalle espacial bajo, debido a que es una imagen compuesta solo por el entrevistado y el fondo. La característica  $IT=5,38$  indica un cambio temporal mínimo, es decir, el movimiento en cada cuadro es mínimo o muy bajo.

Para el video Cafetería(C00), la característica  $IE=65,98$  presenta un detalle espacial más elevado que para el video I00, indicando que es una imagen compleja porque hay más objetos en cada cuadro (árboles, mesas, sombrillas, personas, sillas, quiosco y plantas). La característica  $IT=6,48074$  indica un cambio temporal bajo es decir, en cada cuadro el movimiento es moderado, debido a que algunas personas están de pie y otras caminan lentamente.

En el video Baloncesto(B00), la característica  $IE=57,41.95$  presenta un detalle espacial más bajo que para el video C00, indicando que es una imagen con menos complejidad debido que hay menos objetos en cada cuadro. La característica  $IT=17,4069$  indica un cambio temporal elevado es decir, en cada cuadro hay movimiento y se puede apreciar porque hay personas corriendo a lo largo de la cancha.

Después de explicado lo anterior, los videos fueron elegidos para la toma de datos en los experimentos a realizar debido a que su información espacio-temporal difiere y permite identificar cada video dependiendo de la cantidad de movimiento que presente como: movimiento bajo, movimiento moderado y movimiento alto. La elección de este tipo de videos se realizó basándose en el análisis y sugerencias procedimentales planteadas en los siguientes trabajos: [68][77] [78] [79] [80] [81] [82].

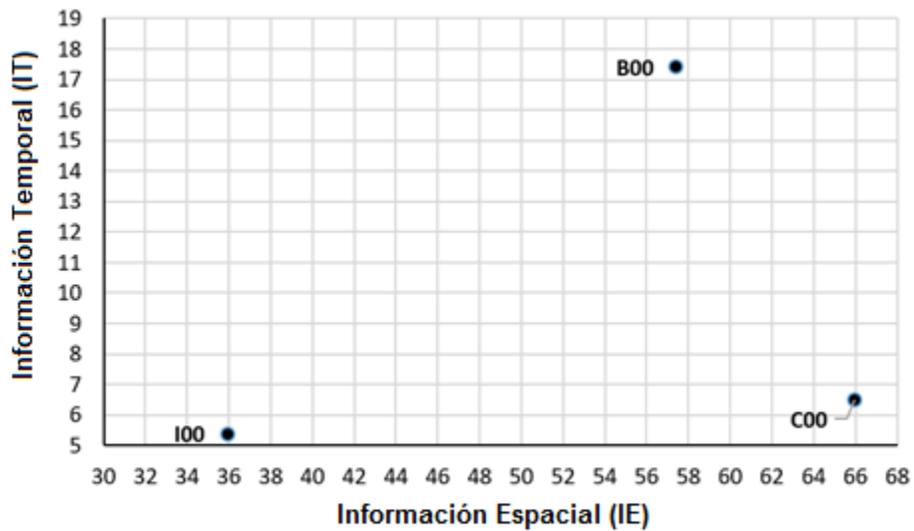


Figura 8. Distribución de los valores IE & IT de los videos originales utilizados en el experimento.

### 3.5 Comentarios del Capítulo

La identificación de los componentes descritos en este capítulo se logró gracias al análisis detallado realizado en la revisión del estado del arte. Y el enfoque utilizado para obtener los modelos propuestos es el primero en este tipo, donde los videos se transmiten y se degradan en línea.

La elección de VQM como algoritmo para medir la calidad de experiencia asociada al video, se debió a que VQM es un método estandarizado, maduro y que analiza el video utilizando mayor número de variables que otros métodos objetivos estandarizados.

En el siguiente capítulo se presentará el diseño del experimento y se explicará detalladamente como se obtuvieron los datos para generar el modelo propuesto.

## 4 Diseño del Experimento para obtener los datos para generar el modelo

En el capítulo anterior se identificaron los componentes para generar los datos para obtener el modelo, en este capítulo, ya con los componentes identificados se presentará el diseño del experimento para obtener el modelo. Se empezará describiendo la prueba de laboratorio que se implementó y los rangos que se asignaron a los parámetros de QoS. Al finalizar se presenta un análisis estadístico de los datos obtenidos y los modelos generados para cada tipo de video. En la Figura 9 se presenta la descripción del contenido esquemático del capítulo.

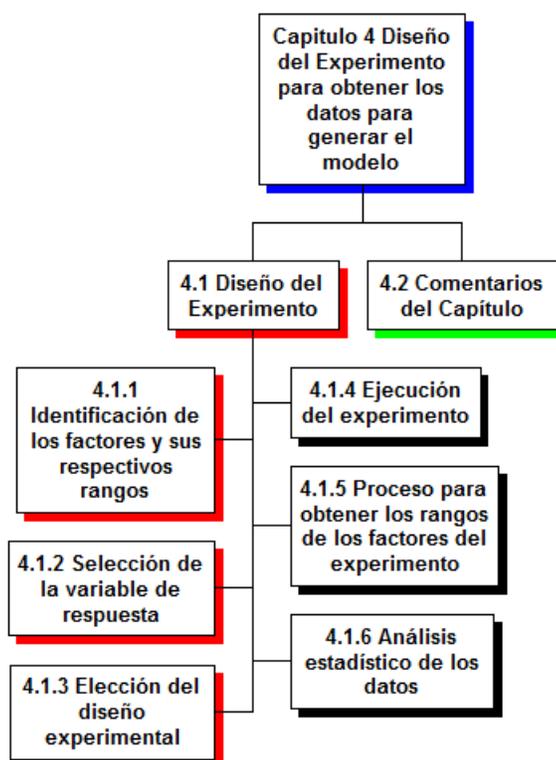


Figura 9. Descripción del contenido del capítulo 4.

## 4.1 Diseño del Experimento

Para obtener los datos necesarios para la generación del modelo se empezó con la planeación y el diseño del experimento. Para diseño del experimento se deben seguir los siguientes pasos [83]:

1. Identificación de los factores con sus respectivos rangos.
2. Selección de variable de respuesta.
3. Elección del diseño experimental
4. Ejecución del experimento
5. Análisis estadístico de los datos
6. Conclusiones

En los siguientes apartados se explicará con detalle cada paso y qué procedimientos se realizaron con el fin de obtener los datos para generar el modelo.

### 4.1.1 Identificación de los factores y sus respectivos rangos

Los factores del diseño de experimentos hacen referencia a la variable independiente que se puede variar dentro de un rango que se define previamente [83].

Para este caso, los factores son los parámetros de QoS que previamente se seleccionaron y los rangos establecidos para su variación se pueden apreciar en la Tabla 4.

Tabla 4. Rango de los parámetros de calidad de servicio configurados en el experimento.

Parámetro de QoS	Rango
Retardo	0 -1300 mseg
Variación del Retardo	0 – 0.018 mseg
Pérdida de paquetes	0 – 0.18%

El proceso para obtener los rangos de los parámetros de QoS presentados en la Tabla 4, será explicado al finalizar la sección 4.1.5 , con el fin de dar claridad sobre el proceso que se siguió y las dificultades que se obtuvieron en ese paso del experimento.

Adicionalmente en el diseño de experimentos se deben tener en cuenta los factores no controlables, que son aquellos que durante el proceso de toma de datos, quedan por fuera del control de la persona que está coordinando el experimento [83].

Para el experimento de toma de datos planteado en esta sección, el factor no controlable fue el proceso de pérdida de paquetes que utilizaba la herramienta NetEm, debido a que la herramienta trabaja de manera aleatoria la pérdida de paquetes. Los paquetes son descartados solo cuando en la memoria intermedia (*buffer*) que usa la herramienta se llena. El tamaño de la memoria interna de la herramienta NetEm se dejó configurada por defecto para no afectar el desarrollo del experimento.

Además para minimizar las fuentes de variación en la toma de datos se trabajó con las siguientes conjeturas:

- No hay configurado en la maqueta de experimentación ningún tráfico de fondo (*background*) entre servidor y cliente. Todo el ancho de banda está dedicado a la transmisión del video. Con esto se descarta que algún tipo de aplicación o tráfico consumiera el ancho de banda y pudiera afectar la transmisión del video, lo cual sería un factor adicional a tener en cuenta.
- El ancho de banda entre cliente y servidor es de 100Mbps, esto debido a que se quería acotar al máximo los factores que pudieran afectar los parámetros de QoS previamente definidos. El modelo a proponer es para el servicio de IPTV residencial, donde los anchos de banda son menores y hay más aplicaciones y/o servicios que el usuario final está consumiendo, el ancho de banda fijo permite analizar cómo la variación de los parámetros de QoS afectan directamente la calidad del video.

#### 4.1.2 Selección de variable de respuesta

La variable de respuesta será la QoE asociada al video, en el experimento esta variable será medida con el algoritmo **VQM** utilizando la herramienta **MSU Video Quality Measurement Tool**, versión 4.4 Professional [84].

En la sección 4.1.4 se explicará el proceso realizado para utilizar esta herramienta y lograr la medida obtenida mediante VQM.

#### 4.1.3 Elección del diseño experimental

El diseño experimental consiste en identificar qué procedimiento se va utilizar para la toma de datos y dependiendo del procedimiento escogido, se define el tamaño de la muestra, número de tratamientos, el orden de las repeticiones y valores que se le asignaran a los factores durante las repeticiones [83].

Identificar el diseño experimental depende del objetivo que se busca, que en este proyecto está dado por el objetivo general: *“Proponer un modelo que evalúe la calidad de la experiencia en línea para el servicio de IPTV en un entorno residencial”*.

Dado que el objetivo es proponer un modelo que evalúe la calidad de experiencia, deben existir valores de los factores (i.e. parámetros de QoS) que permiten obtener la mejor respuesta (i.e. QoE asociada al video), es decir, que el propósito final es establecer los valores de los factores que optimizan el valor de la variable respuesta. Con el modelo obtenido, se espera predecir valores de la respuesta (QoE asociada al video) que se encuentren dentro del rango definido a los factores (valores de los parámetros de QoS) utilizados en el diseño experimental.

De acuerdo a lo anterior, el tipo de diseño experimental que se ajusta a esos requerimientos es la **Metodología de Superficies de Respuesta (MSR)**. MSR es definida como un conjunto de técnicas matemáticas y estadísticas utilizadas en el

tratamiento de problemas en los que una **respuesta** de interés (QoE del video), está influenciada por varios **factores** de carácter cuantitativo (parámetros de QoS) y donde el objetivo es optimizar la respuesta de interés [83].

El propósito inicial de esta técnica es diseñar un experimento bajo condiciones controladas, que proporcione valores de la variable respuesta y, a continuación, determinar el modelo matemático que mejor se ajusta a los datos obtenidos. Para este caso, se debe elegir una distribución satisfactoria de los puntos experimentales sobre la región experimental y el diseño central compuesto (DCC) es el que más se ajusta y el más utilizado por su flexibilidad.[85]. Como se trata de analizar tres parámetros de QoS se definió utilizar un diseño central compuesto (DCC) de tres factores, con un alfa ( $\alpha$ ) igual a 2. El valor del alfa se eligió para facilitar la configuración de los parámetros de calidad de servicio, con el fin de dejarlos en un número entero y homogenizar el número de decimales a utilizar. Se define como tratamiento o ejecución en desarrollo del experimento, el proceso de modificar los valores de los factores (parámetros de QoS) con el fin de obtener una respuesta (QoE del video) que posteriormente será analizada.

Basándose en lo anterior un diseño central compuesto está conformado por tres tipos de puntos, que posteriormente se convertirán en los tratamientos del experimento [85]:

- **Porción factorial:** es una réplica de un diseño factorial que para este caso será  $2^k$ , donde  $k$  es el número de factores, por lo tanto  $2^k = 2^3 = 8$  tratamientos en la porción factorial. La representación geométrica de  $2^3$  tratamientos se presenta en la Figura 10, donde la región experimental es un cubo regular centrado en el origen (0,0,0) cuyos vértices con los 8 tratamientos y al lado izquierdo se presenta la tabla con los factores codificados [85].
- **Puntos axiales:** Dos puntos sobre cada eje a una distancia  $\alpha$  del origen. La elección del  $\alpha$  está definido por  $\alpha = (F)^{1/4}$ , donde  $F$  es el número de puntos de la parte factorial, por lo tanto,  $\alpha = (8)^{1/4} = 1.681$ , pero para este caso se eligió un  $\alpha$  igual a 2, esto con el fin de que los valores de cada factor fueran números enteros y así facilitar procesos de configuración en el emulador (NetEm) tabla 5
- . En la Figura 11 se pueden apreciar la ubicación de los puntos axiales del DCC.
- **Puntos centrales:** número de repeticiones al centro del diseño ( $n_o$ ), que según lo define [83] para 3 factores es  $n_o = 6$ .

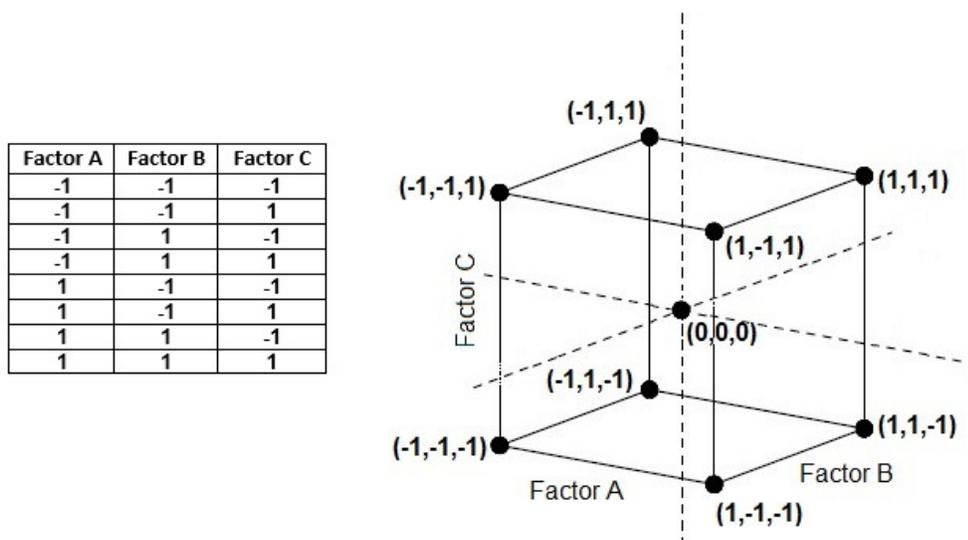


Figura 10. Tabla de codificación porción factorial del DCC para tres factores y su representación geométrica.

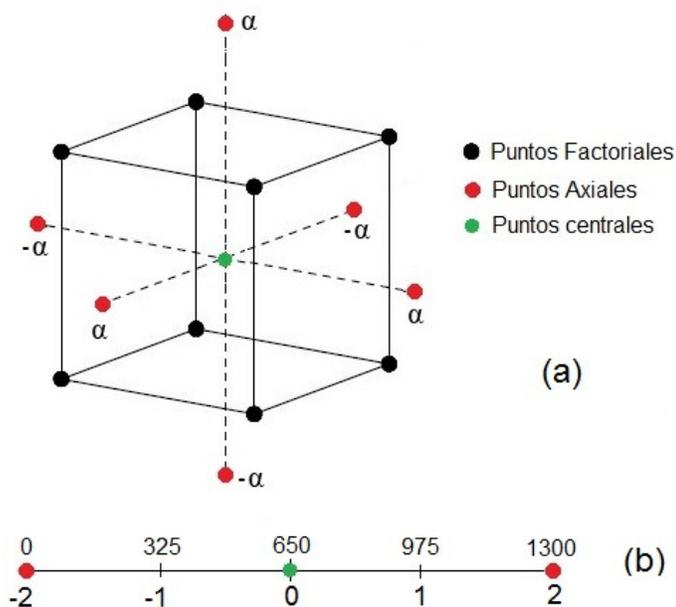


Figura 11. (a) Puntos Axiales y factoriales del DCC. (b) Valores asociados a la codificación del eje X, parámetro retardo.

Ya presentado lo anterior, el número de tratamientos para el DCC en cuestión, está definido por  $N = 2^k + 2k + n_0$ , [85]. Por lo tanto  $N = 2^3 + 2(3) + 6 = 20$  tratamientos.

Para determinar el valor de los factores para cada tratamiento se parte de los rangos presentados en la Tabla 4 de la página 52. Estos rangos determinaran el área de experimentación donde se ubicará el cubo de la Figura 11, por lo tanto el rango de cada factor se debe dividir en 4, esto por el valor del alfa definido (puntos axiales) y los puntos factoriales que están a una distancia de uno del alfa definido.

Para entender mejor cómo se generan los valores presentados en la Tabla 5, se tomará como ejemplo el eje X de la Figura 11(a), correspondiente al Retardo (R). Por lo tanto como el alfa es 2, dicho eje irá desde -2 hasta 2, siendo intersectado por los otros ejes, en -1, 0 y 1, hasta llegar a 2, como se aprecia en la figura 12(b).

El rango definido para el retardo es de 0 hasta 1300 mseg, al dividir dicho rango entre 4, se obtienen los valores asociados a cada punto en la codificación como se aprecia en la Figura 11(b).

De igual manera se calculan los valores para cada tratamiento para la variación del retardo y la pérdida de paquetes. En la Tabla 5, se puede apreciar los valores codificados con su respectivo valor real.

Por lo tanto, el valor para el punto codificado (-1,-1,-1) sería (325,0.0045, 0.045) que está a una distancia de  $\frac{1}{4}$  del valor más bajo de cada factor. El punto codificado (-2,0,0) tendría un valor de (0, 0.009, 0.09) y el punto (0,0,0) estaría ubicado en la mitad del valor de cada factor, (650, 0.009, 0.09).

En la Tabla 5 se relacionan los puntos factoriales del cubo con sus valores codificados, al igual que los puntos axiales y punto central presentados en la Figura 11. En la sección 4.1.4 se explicará el proceso que se realizó para definir los rangos de los factores (parámetros de QoS) para el diseño de las pruebas.

El DCC diseñado permite obtener un modelo para cada tipo de video, donde el modelo obtenido es un modelo jerárquico de segundo orden. Este tipo de modelo tiene un comportamiento más estable, que facilita la exploración de la superficie de respuesta que él representa. Además permite que los términos no significativos permanezcan en el modelo para lograr la jerarquía [85].

Tabla 5. Tratamientos para el experimento con los valores codificados y valores reales asociados a cada factor.

No. de Tratamiento	Valores Codificados			Valores Reales			
	R	V <sub>R</sub>	P <sub>P</sub>	Retardo (R)	Variación del Retardo (V <sub>R</sub> )	Pérdida de paquetes (P <sub>P</sub> )	
1	-1	-1	-1	325	0,0045	0,045	<b>Puntos Factorial (cubo)</b>
2	1	-1	-1	975	0,0045	0,045	
3	-1	1	-1	325	0,0135	0,045	
4	1	1	-1	975	0,0135	0,045	
5	1	-1	1	325	0,0045	0,135	
6	-1	-1	1	975	0,0045	0,135	
7	1	1	1	325	0,0135	0,135	
8	-1	1	1	975	0,0135	0,135	
9	-2	0	0	<b>0</b>	0,009	0,09	<b>Puntos axiales</b>
10	2	0	0	<b>1300</b>	0,009	0,09	
11	0	-2	0	650	<b>0</b>	0,09	
12	0	2	0	650	<b>0,018</b>	0,09	
13	0	0	-2	650	0,009	<b>0</b>	
14	0	0	2	650	0,009	<b>0,18</b>	
15	0	0	0	650	0,009	0,09	<b>Puntos Centrales</b>
16	0	0	0	650	0,009	0,09	
17	0	0	0	650	0,009	0,09	
18	0	0	0	650	0,009	0,09	
19	0	0	0	650	0,009	0,09	
20	0	0	0	650	0,009	0,09	

El modelo a obtener es de la forma presentada en la Ecuación 4, donde  $X_i$  representa los factores (parámetros de QoS: Retardo, Variación del Retardo y Pérdida de paquetes), y  $Y$  representa la respuesta de interés (QoE del video) y  $\varepsilon$  es el error aleatorio:

$$Y = \beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i=1}^k \beta_i X_i^2 + \sum_{i=1}^k \sum_{j=1}^k \beta_{ij} X_i X_j + \varepsilon \quad \text{Ecuación 4}$$

Para estimar los coeficientes ( $\beta_i$ ) en la Ecuación 4, el método de mínimos cuadrados es usado. Suponiendo que  $n > k$  tratamientos están disponibles, y  $X_{ij}$  expresa el  $j$ -ésimo valor del factor  $X_i$ . Los datos de los tratamientos o ejecuciones se presentan en la Tabla 6. Este procedimiento de estimación requiere que el error aleatorio tenga  $E(\varepsilon) = 0$ ,  $V(\varepsilon) = \sigma^2$ , y que el  $\{\varepsilon\}$  no estén correlacionados [85].

Tabla 6. Datos para el modelo presentado en la Ecuación 4 .

$Y$	$X_1$	$X_2$		$X_k$
$Y_1$	$X_{12}$	$X_{21}$	...	$X_{k1}$
$Y_2$	$X_{21}$	$X_{22}$	...	$X_{k2}$
.	.	.		.
.	.	.		.
.	.	.		.
$Y_n$	$X_{1n}$	$X_{2n}$	...	$X_{kn}$

El modelo presentado en la Ecuación 4 puede ser escrito en términos de los datos presentados en la Tabla 6, como:

$$\begin{aligned}
 Y_i &= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i \\
 &= \beta_0 + \sum_{j=1}^k \beta_j X_{ij} + \varepsilon_i \quad i=1,2, \dots, n
 \end{aligned}$$

Ecuación 5

El método de mínimos cuadrados consiste en elegir los  $\beta$  de la Ecuación 5 de tal modo que la suma de los cuadrados de los errores  $\varepsilon_i$ , se minimice. Por lo tanto, la función de mínimos cuadrados se define como:

$$\begin{aligned}
 S &= \sum_{i=1}^n \varepsilon_i^2 \\
 &= \sum_{i=1}^n \left[ Y_i - \beta_0 - \sum_{j=1}^k \beta_j X_{ij} \right]^2
 \end{aligned}$$

Ecuación 6

La función  $S$ , Ecuación 6, debe minimizarse con respecto a  $\beta_1, \beta_2, \dots, \beta_k$ . Los estimadores de mínimos cuadrados, por ejemplo:  $\widehat{\beta}_0, \widehat{\beta}_1, \dots, \widehat{\beta}_k$ , deben satisfacer:

$$\left. \frac{\partial S}{\partial \beta_j} \right|_{\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k} = -2 \sum_{i=1}^n \left( Y_i - \hat{\beta}_0 - \sum_{j=1}^k \hat{\beta}_j X_{ij} \right) X_{ij} = 0 \quad \text{Ecuación 7}$$

$$j=1, 2, \dots, k$$

Al simplificar la

Ecuación 7 se obtiene  $k+1$  ecuaciones y la solución de las  $k+1$  ecuaciones simultáneas, son los estimadores de los mínimos cuadrados  $\hat{\beta}_j$ . Resolver este conjunto de ecuaciones simultáneas es más sencillo si se utiliza la notación matricial. Por lo tanto el modelo presentado en la Ecuación 5, puede escribirse en notación matricial como:

$$y = X\beta + \varepsilon$$

Donde,

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix},$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

El vector de estimadores de mínimos cuadrados,  $\beta$ , se obtiene minimizando la Ecuación 8, ahora presentada en forma matricial:

$$S = \sum_{i=1}^n \varepsilon_i^2 = \varepsilon' \varepsilon = (y - X\beta)'(y - X\beta) \quad \text{Ecuación 8}$$

S puede ser expresado de la siguiente manera:

$$\begin{aligned} S &= y'y - \beta'X'y - y'X\beta + \beta'X'X\beta \\ &= y'y - 2\beta'X'y + \beta'X'X\beta \end{aligned} \quad \text{Ecuación 9}$$

De la Ecuación 9, los estimadores de los mínimos cuadrados deben satisfacer la siguiente expresión:

$$\left. \frac{\partial S}{\partial \beta} \right|_{\hat{\beta}} = -2X'y + 2X'X\hat{\beta} = 0$$

Esto implica que

$$X'X\hat{\beta} = X'y \quad \text{Ecuación 10}$$

Para resolver la Ecuación 10 en términos de  $\hat{\beta}$ , multiplique ambos lados por la inversa  $X'X$ . Así, el estimador de mínimos cuadrados de  $\beta$  es:

$$\hat{\beta} = (X'X)^{-1}X'y \quad \text{Ecuación 11}$$

Por lo tanto el modelo ajustado está dado por:

$$\hat{y} = X'\hat{\beta} \quad \text{Ecuación 12}$$

Y en notación escalar, el modelo ajustado es

$$\hat{y}_i = \hat{\beta}_0 + \sum_{j=1}^k \hat{\beta}_j x_{ij} \quad i=1,2,\dots,n \quad \text{Ecuación 13}$$

Ahora partiendo de los valores  $\hat{\beta}$  obtenidos en la Ecuación 13 se puede expresar el modelo de la Ecuación 4 como:

$$Y_{(QoE)} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_1^2 + \beta_5 X_2^2 + \beta_6 X_3^2 + \beta_7 X_1 X_2 + \beta_8 X_1 X_3 + \beta_9 X_2 X_3 \quad \text{Ecuación 14}$$

Por lo tanto en el modelo propuesto en la Ecuación 14,  $X_1$ =Retardo,  $X_2$ =Variación del Retardo,  $X_3$ =Pérdida de Paquetes. Y con el fin de hacer el modelo más amigable para su aplicación e interpretación, los factores representados por  $X_i$ , serán reemplazados por:  $R$ =Retardo,  $V_R$ =Variación del Retardo,  $P_P$ =Pérdida de Paquetes. El resultado se presenta en la Ecuación 15:

$$QoE = \beta_0 + \beta_1 R + \beta_2 V_R + \beta_3 P_P + \beta_4 R^2 + \beta_5 V_R^2 + \beta_6 P_P^2 + \beta_7 R V_R + \beta_8 R P_P + \beta_9 V_R P_P \quad \text{Ecuación 15}$$

Como se puede apreciar el modelo presentado en la Ecuación 15 es un modelo de segundo orden con interacción de los tres parámetros de QoS. El modelo permite obtener la medida de QoE asociada al video solo identificando el valor de los tres parámetros de QoS del video a medir. En la sección 4.1.6, se presentarán los modelos

obtenidos para cada tipo de video y se presentará el valor asociado a cada constante  $\beta$ .

#### 4.1.4 Ejecución del experimento

Basándose en los tratamientos de la Tabla 5 el experimento se planteó de la siguiente manera:

1. Para cada tratamiento en el NetEM se configuraban los valores para los factores (parámetros de QoS) de acuerdo a lo planteado en la Tabla 5. Por ejemplo para el tratamiento 1, los parámetros de QoS se configuraron de la siguiente manera:
  - Retardo: 325 msec
  - Variación del Retardo: 0.0045 msec
  - Pérdida de paquetes: 0.045 %
2. Con los factores (parámetros de QoS) configurados, se procede a enviar el vídeo desde el servidor hasta el cliente VLC. En el cliente VLC el video es almacenado para posteriormente editarlo y recortarlo a 10 segundos.
3. El video original que tiene una duración de casi 30 segundos, se envía desde el servidor de video al cliente. En el cliente el video es almacenado para editarlo y recortarlo a 10 segundos, esto con el fin de medir la QoE de dicho video con la herramienta MSU Video Quality Measurement Tool.
4. Se procede a medir la QoE del video utilizando la herramienta MSU Video Quality Measurement Tool, comparando el video original y el video recibido en el cliente VLC, ambos editados a 10 segundos. Con los valores configurados en el ítem 2 y la medida de QoE realizada, se almacena como se aprecia en la Tabla 7, los valores asociados a un tratamiento. Este proceso se repite para los 20 tratamientos de cada video (ver Tabla 5), y posteriormente se replica 2 veces más todo el experimento. Es decir que para cada video se realizan 60 tratamientos.

Tabla 7. Medición de QoE obtenida para el tratamiento No.1

No. de Tratamiento	Retardo	Variación del Retardo	Pérdida de paquetes	QoE_Cafeteria VQM
1	325	0,0045	0,045	4,782823333

5. Adicional a los 20 tratamientos realizados a cada video, se realizó un tratamiento de control (tratamiento identificado como No.0) que consistió en configurar los tres parámetros de QoS en 0 (cero). Esto con el fin de obtener el video transmitido sin ningún tipo de distorsión, posteriormente se explicará en qué tipo de prueba se usó este tratamiento de control.

#### 4.1.5 Proceso para obtener los rangos de los factores del experimento

Para obtener los rangos de los factores del experimento se requirió de varias iteraciones, realizando réplicas de prueba de todo el experimento explicado en la sección anterior 4.1.4.

Las pruebas iniciales se realizaron basándose en los rangos definidos en la Recomendación UIT-T 1541 [59] para los parámetros elegidos de QoS, donde:

- Retardo: 0 - 400mseg
- Variación del retardo: 0 - 50mseg
- Pérdida de paquetes: 0 –  $1 \times 10^{-3}$

Los datos obtenidos en las réplicas iniciales no fueron estadísticamente válidos, es decir, los resultados obtenidos después del análisis de varianza realizado, indicaban una fidelidad muy baja (cálculo del  $R_{cuadrado}$ ) en el modelo obtenido. Adicionalmente presentaban los siguientes inconvenientes a nivel técnico:

- Los videos obtenidos en las primeras pruebas con los rangos de la Recomendación UIT-T 1541 presentaban distorsiones generalizadas que no permitían su correcta visualización y por ende la medición obtenida de VQM era elevada (calidad deficiente del video).
- Para algunos valores de los parámetros de QoS los videos no se podían almacenar, esto debido a que la pérdida de paquetes impedía decodificar el video de manera correcta en el cliente.
- Para algunos datos generados, se realizaron pruebas estadísticas que no permitían generar un modelo o en ciertos casos, modelos con niveles de aproximación estadísticamente muy bajos. Esto debido a que los videos de una prueba no presentaban variaciones, por ejemplo, de los 20 tratamientos, el 90% de los videos obtenidos tenían una distorsión muy alta, lo cual hacia que el valor de QoE fuera muy bajo.

Por las razones expuestas anteriormente, se realizaron en la maqueta experimental una serie de pruebas ajustando los valores de los parámetros de QoS que permitieran la transmisión de cada video y adicionalmente que se pudieran apreciar distorsiones significativas mientras se transmitía cada uno de ellos. Después de numerosas pruebas, variando los parámetros de QoS, se lograron obtener los valores presentados en la Tabla 4. Basándonos en lo anterior, los rangos de los parámetros de calidad de servicio obtenidos se generaron de manera empírica, y son un aporte adicional de esta investigación. Ya que se pudo demostrar experimentalmente que los rangos obtenidos permiten transmitir video en alta definición sin distorsiones, hasta obtener distorsiones perceptibles para el usuario.

Con los valores de los parámetros de QoS obtenidos (Tabla 4) en las pruebas realizadas, se ejecutaron análisis de varianza a los modelos obtenidos y las respuestas

obtenidas fueron estadísticamente válidas. Este proceso y estos análisis de varianza se presentarán en la siguiente sección.

Además los videos transmitidos, después de la modificación de los parámetros de QoS se podían decodificar en el cliente. La decodificación del video en el cliente era de suma importancia, debido a que el video se tenía que almacenar para posteriormente editarlo a 10 segundos.

En la Figura 12 se presenta de manera resumida el proceso general para obtener los datos para la generación del modelo. Empezando por la identificación y rangos de los parámetros de QoS, luego la identificación del modelo para evaluar la QoE de los videos, después la implementación de la arquitectura de pruebas y definición del experimento, para finalizar con el análisis de los datos que se explicarán en la siguiente sección.

En este punto vale la pena mencionar la cantidad de pruebas y archivos de video que se procesaron durante toda la etapa experimental para obtener los datos necesarios para generar los modelos.

Para cada video, una prueba consistía en ejecutar los 20 tratamientos que se han explicado anteriormente, pero para obtener un video adecuado en un tratamiento, se debían realizar al menos 5 transmisiones de video para dicho tratamiento. Esto con el fin de trabajar con el video que se pudiera decodificar y almacenar adecuadamente en el cliente VLC.

Basándose en lo anterior para una sola prueba se requería procesar aproximadamente 100 videos o realizar 100 transmisiones del mismo. Y para que los datos tuvieran validez estadística para cada modelo propuesto se realizaron tres pruebas, por lo tanto para cada modelo se procesaron más de 900 videos.

Es decir, que para los tres modelos propuestos se procesaron más de 2700 videos, desde las pruebas iniciales para definir los rangos de los parámetros de QoS, hasta las pruebas que permitieron obtener datos válidos para generar cada modelo.

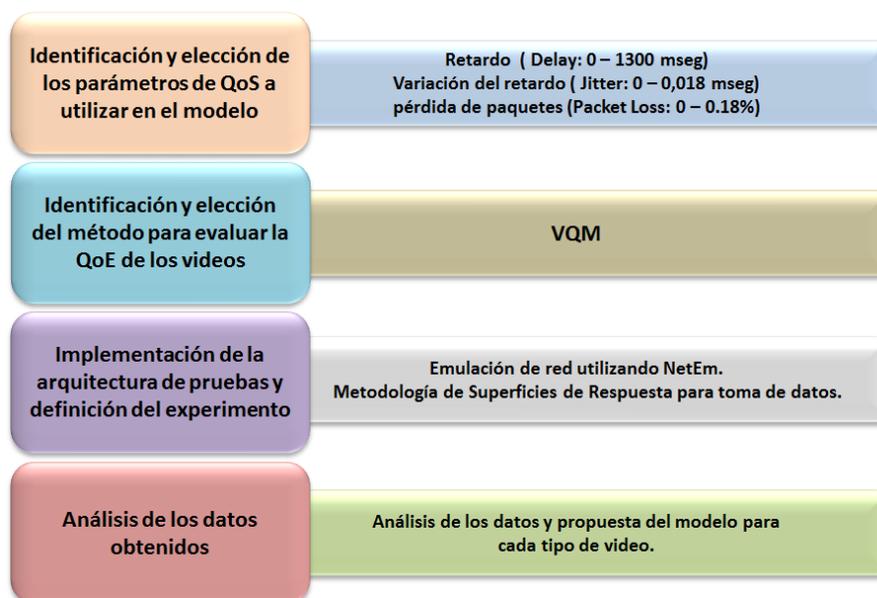


Figura 12. Proceso para lo obtener los datos para la generación de los tres modelos

#### 4.1.6 Análisis estadístico de los datos

En esta sección se presentará el modelo obtenido para cada tipo de video (cafetería, entrevista, baloncesto), con su respectivo análisis estadístico que permitirá analizar la precisión en la interpretación de los datos en cada modelo.

##### 4.1.6.1 Modelo obtenido para el video Entrevista

Como se explicó en la sección anterior, en la ejecución del experimento se ejecutaron 20 tratamientos. En cada uno de ellos se transmitió el video que fue grabado en el cliente, posteriormente el video grabado se editó y se le hizo la medida de QoE utilizando algoritmo VQM. Este procedimiento es el mismo para los videos Cafetería y Baloncesto.

Para el video entrevista los datos obtenidos se presentan en la Tabla 8. En la columna *No. Tratamiento* está indicado el número de tratamiento, posteriormente aparecen tres columnas con los valores de los factores (parámetros de QoS: Retardo, Variación del Retardo y Pérdida de Paquetes) y la última columna es la medida de QoE obtenida con el algoritmo VQM. El mismo formato de tabla se utilizará para presentar los datos de los videos cafetería y baloncesto.

Tabla 8. Datos obtenidos para generar la modelo del video entrevista.

No. Tratamiento	Retardo (mseg)	Variación del Retardo (mseg)	Pérdida de paquetes (%)	QoE_Entrevista VQM
1	325	0,0045	0,045	1,27934
2	975	0,0045	0,045	1,07103
3	325	0,0135	0,045	1,40668
4	975	0,0135	0,045	1,14887
5	325	0,0045	0,135	1,65073
6	975	0,0045	0,135	3,40487
7	325	0,0135	0,135	1,56809
8	975	0,0135	0,135	2,56156
9	<b>0</b>	0,009	0,09	1,69293
10	<b>1300</b>	0,009	0,09	1,75235
11	650	<b>0</b>	0,09	1,40056
12	650	<b>0,018</b>	0,09	1,27615
13	650	0,009	<b>0</b>	1,3631
14	650	0,009	<b>0,18</b>	3,24128
15	650	0,009	0,09	1,29295
16	650	0,009	0,09	2,66169
17	650	0,009	0,09	2,38688
18	650	0,009	0,09	1,74842
19	650	0,009	0,09	1,66789
20	650	0,009	0,09	2,40491

En la Figura 13 se presenta el diagrama de barras de los datos obtenidos en la Tabla 8, con su coeficiente de variación (Coef. Variación=37,24), la media de los datos y una desviación estándar por encima y por debajo de la media (Desviación estándar = 0,689). Para el caso del valor de VQM, entre menor sea el valor obtenido, mejor será la calidad del video, para este caso un valor de VQM bajo (1,071) indica una calidad del video mejor en comparación de un valor superior de VQM (3,405) que indica un video de calidad deficiente para este caso. En el siguiente capítulo se explicará cómo se mapeo el valor VQM con la escala MOS con el objetivo de realizar pruebas comparativas de desempeño.

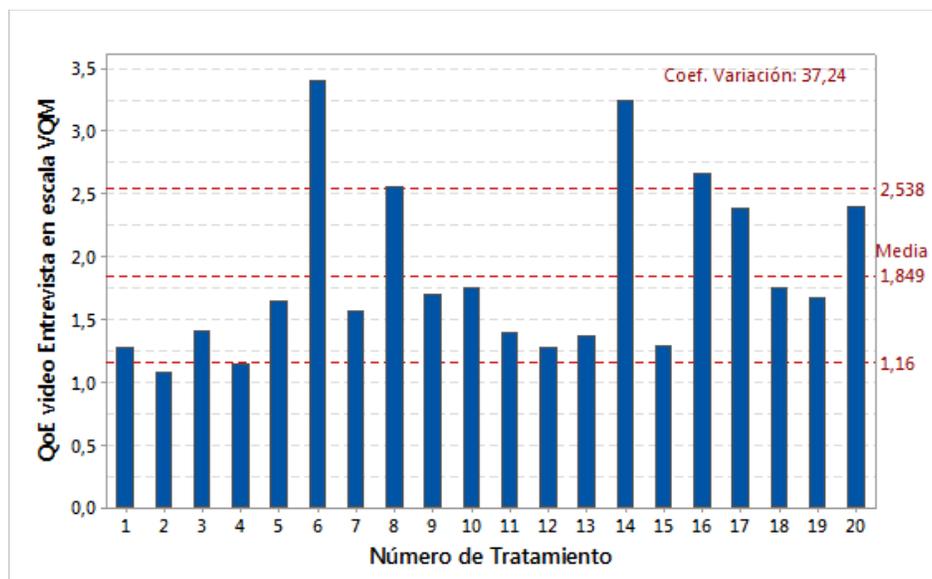


Figura 13. Diagrama de barras de la QoE obtenida al video Entrevista en escala VQM.

En la Tabla 9 se presenta el análisis de varianza de los datos presentados en la Tabla 8, estos valores se obtuvieron mediante el uso de la herramienta Minitab. Para el análisis de los datos de los tres videos se trabajó con un nivel de significancia de 0.05 ( $\alpha = 0.05$ ). Este nivel de significancia en su definición básica, es la probabilidad de rechazar la hipótesis nula cuando es verdadera [86]. Es decir, el nivel de significancia determina la probabilidad de error inherente a la prueba. Para este caso un nivel de significancia de 0.05 indica un riesgo de 5%, de concluir que existe una diferencia cuando no existe una diferencia real.

El valor del nivel de significancia se utiliza en el análisis de varianza a fin de definir qué parámetros tienen influencia significativa en el modelo. Esto se hace comparando el *valor p* con el nivel de significancia ( $\alpha$ ), cuando el *valor p* es menor que el  $\alpha$ , ( $p < \alpha$ ) se concluye que el parámetro tiene influencia significativa en el modelo.

Tabla 9. Análisis de Varianza para los datos obtenidos al video entrevista.

Fuente	GL	SC Ajust.	MC. Ajust.	Valor F	Valor p
Modelo	9	7,17118	0,79680	4,33	0,016
Lineal	3	4,45463	1,48488	8,07	0,005
R (Retardo)	1	0,36010	0,36010	1,96	0,192
VR (Variación del Retardo)	1	0,05876	0,05876	0,32	0,584
PP (Pérdida de Paquetes)	1	4,03577	4,03577	21,94	0,001
Cuadrado	1	0,17722	0,17722	0,96	0,349
R*R	1	0,17722	0,17722	0,96	0,349
VR*VR	1	0,81487	0,81487	4,43	0,062
PP*PP	1	0,09335	0,09335	0,51	0,493
Interacción de 2 factores	3	1,53299	0,51100	2,78	0,096
R*VR	1	0,08205	0,08205	0,45	0,519
R*PP	1	1,29101	1,29101	7,02	0,024
VR*PP	1	0,15993	0,15993	0,87	0,373

Error	10	1,83952	0,18395		
Falta de ajuste	5	0,41896	0,08379	0,29	0,897
Error puro	5	1,42056	0,28411		
Total	19	9,01070			

GL: Grados de Libertad.

SC Ajust.: Suma de Cuadrados

MC Ajust.: Media de los Cuadrados

De los datos de la Tabla 9 se puede concluir que el parámetro que tiene influencia significativa en el modelo es la pérdida de paquetes y la interacción entre los factores Retardo y Pérdida de Paquetes. Esto se obtiene analizando la última columna (valor p), donde este valor debe ser inferior a 0,05 para que el parámetro sea considerado como influyente en el modelo. Igual ocurre para los otros dos videos que se analizaran posteriormente.

Otro aspecto que se debe tener en cuenta en el análisis de varianza es el **coeficiente de determinación o R-cuadrado**, definido como el porcentaje de variación de la variable de respuesta (QoE) que explica su relación con los factores (parámetros de QoS) [87]. El R-cuadrado se encuentra entre 0-100% y entre mayor sea, mejor será el ajuste del modelo a los datos.

Para el análisis de varianza de los datos de la Tabla 9, el R-cuadrado obtenido fue de 79,59%, lo cual indica que el modelo explica el 79,59% la variabilidad de los datos con que se generó el modelo.

Por lo tanto, el modelo obtenido mediante el uso de la herramienta Minitab para el video Entrevista es el presentado en la Ecuación 16, donde Retardo es representado por **R**, Variación del Retardo por **V<sub>R</sub>** y Pérdida de paquetes por **P<sub>P</sub>**. Este modelo se identificará como **Modelo Video Movimiento Bajo (MVMB) –video entrevista**.

$$\begin{aligned}
 QoE = & 0,65 + 0,35 \times 10^{-3} R + 254 V_R - 5,8 P_P - 1 \times 10^{-6} R^2 \\
 & - 0,069 R V_R - 8890 V_R^2 + 0,0275 R P_P - 698 V_R P_P \\
 & + 30,1 P_P^2
 \end{aligned}
 \quad \text{Ecuación 16}$$

De manera similar al modelo obtenido para el video de cafetería el modelo presentado en la Ecuación 16 (MVMB – video entrevista), se le incluyeron todos los factores que tenían influencia significativa y los que no tenían influencia o su influencia era muy baja. Esto se hizo debido a que al retirar factores de influencia significativa baja del análisis de varianza y generar de nuevo el modelo, el R-cuadrado bajaba a un 65,29%. Por tal razón se decide dejar el modelo con todos los factores, ya que los factores que se pueden retirar del modelo aportan de alguna manera en mejorar el R-cuadrado y mejorar la precisión del modelo.

En la Figura 14 y Figura 15 se pueden apreciar las gráficas de contorno del parámetro que influye estadísticamente en el modelo, (eje Y) y los dos parámetros que no influyen estadísticamente en el modelo (retardo y variación del retardo). Se puede apreciar que para el retardo y la variación del retardo, la variación de la QoE con respecto a la pérdida de paquetes es casi constante. Teniendo en cuenta que en cada

figura hay un sitio donde se obtiene la mejor medición de QoE y es cuando la variación del retardo está próxima a 0,0050 mseg y el retardo cercano a 1000 mseg.

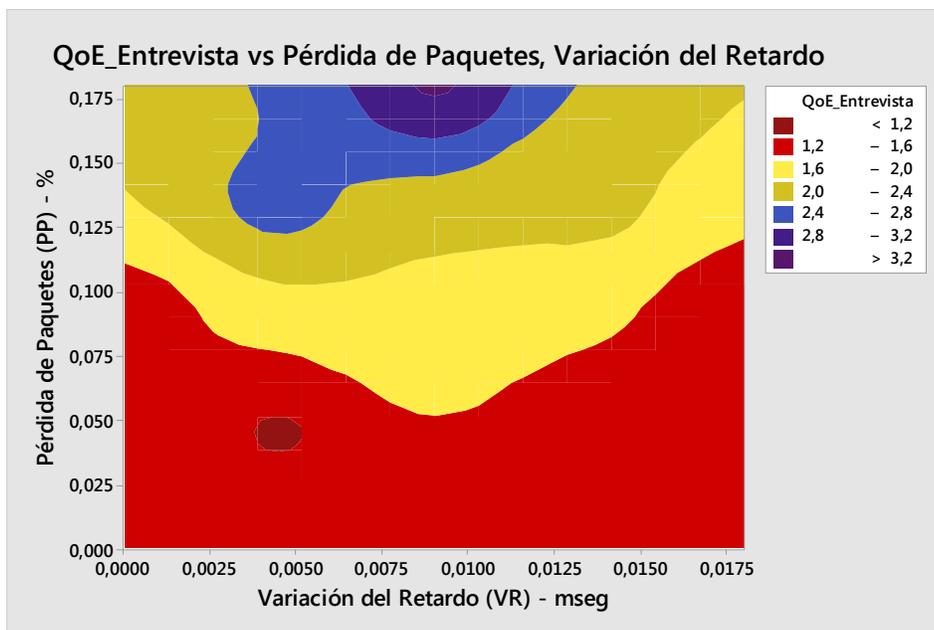


Figura 14. Gráfica de contorno QoE\_Entrevista vs Pérdida de paquetes, Variación del Retardo.

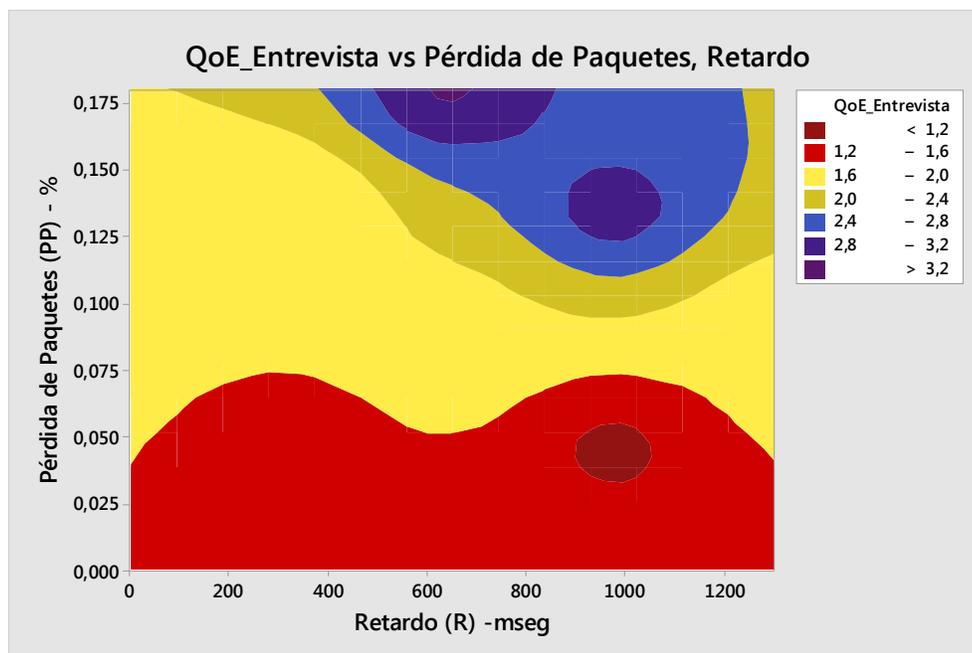


Figura 15. Gráfica de contorno QoE\_Entrevista vs Pérdida de Paquetes, Retardo.

En la Figura 16 se graficaron los dos parámetros que no influyen estadísticamente en el modelo (retardo y variación del retardo), se puede apreciar que para ciertos

intervalos se puede lograr obtener una medición de QoE hasta un valor de 1.6. Esto se debe nuevamente al tipo de video, en el cual ante los pocos cambios que tiene ciertas distorsiones no se reflejan en la medida de VQM.

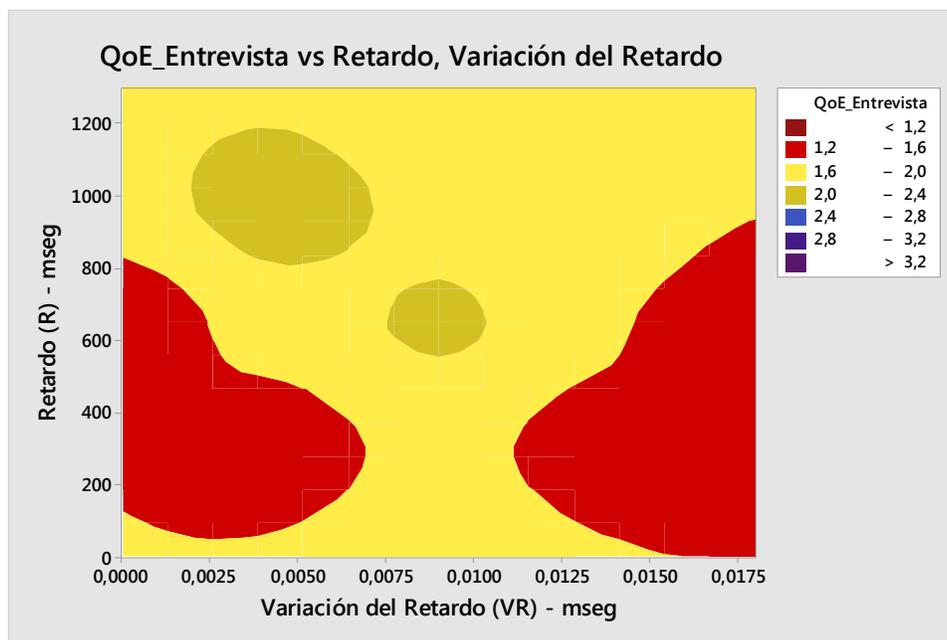


Figura 16. Gráfica de contorno QoE\_Entrevista vs Retardo, Variación del Retardo.

De igual manera que para el video cafetería el parámetro que más influye es la pérdida de paquetes, y se puede apreciar que después de 0.1% la QoE del video se empieza a degradar. Aunque se debe tener en cuenta que por el tipo de video (video sin movimiento) los valores que entrega el modelo propuesto son mucho más bajos comparados con los obtenidos con el modelo obtenido para el video cafetería, que se explicará en la siguiente sección.

#### 4.1.6.2 Modelo obtenido para el video Cafetería

En la Tabla 10 se presentan los resultados de la ejecución del experimento para el video Cafetería. Se puede apreciar la diferencia en el valor de VQM obtenido con respecto al video entrevista y esto se debe a las características espacio-temporales del video entrevista, que es un video de movimiento bajo.

En la Tabla 10 se puede apreciar que para un valor en retardo de 1300 mseg, la calidad del video se degrada a su valor máximo en escala VQM (11,43) y para un valor de retardo de 325 mseg se obtiene el valor mínimo de VQM (3,65) indicando que el video es de mejor calidad.

Tabla 10. Datos obtenidos para generar el modelo del video cafetería.

No. Tratamiento	Retardo (mseg)	Variación del Retardo (mseg)	Pérdida de paquetes (%)	QoE_Cafetería VQM
1	325	0,0045	0,045	4,782823333
2	975	0,0045	0,045	6,239113333
3	325	0,0135	0,045	3,65498
4	975	0,0135	0,045	5,744786667
5	325	0,0045	0,135	4,913223333
6	975	0,0045	0,135	6,77963
7	325	0,0135	0,135	6,536403333
8	975	0,0135	0,135	6,80659
9	0	0,009	0,09	5,94472
10	1300	0,009	0,09	11,43939667
11	650	0	0,09	5,08529
12	650	0,018	0,09	5,018236667
13	650	0,009	0	3,97794
14	650	0,009	0,18	6,778456667
15	650	0,009	0,09	5,700313333
16	650	0,009	0,09	5,728156667
17	650	0,009	0,09	5,98116
18	650	0,009	0,09	7,200743333
19	650	0,009	0,09	4,366896667
20	650	0,009	0,09	5,928826667

En la Figura 17 se presenta el diagrama de barras de los datos obtenidos en la Tabla 10, con su coeficiente de variación (Coef. Variación=27,38), la media de los datos (5,93) y una desviación estándar por encima (7,55) y por debajo (4,31) de la media (Desviación estándar = 1,624).

Un valor de VQM bajo (cercano a 3.65) indica una calidad del video mejor en comparación de un valor alto de VQM (cercano a 11,43) que indica un video de calidad deficiente. En el siguiente capítulo se explicará cómo se mapeo el valor VQM con la escala MOS, con el objetivo de realizar pruebas comparativas de desempeño.

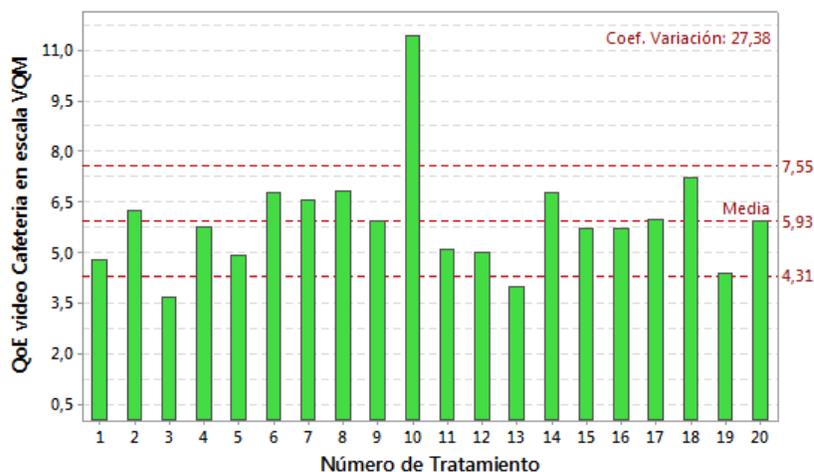


Figura 17. Diagrama de barras de la QoE obtenida al video Cafetería en escala VQM.

En la Tabla 11 se presenta el análisis de varianza de los datos presentados en la Tabla 10, estos valores se obtuvieron mediante el uso de la herramienta Minitab.

Tabla 11. Análisis de Varianza para los datos obtenidos al video Cafetería.

Fuente	GL	SC Ajust.	MC. Ajust.	Valor F	Valor p
Modelo	9	42,2409	4,6934	5,97	0,005
Lineal	3	23,8949	7,9650	10,14	0,002
<b>R (Retardo)</b>	<b>1</b>	<b>17,3723</b>	<b>17,3723</b>	<b>22,11</b>	<b>0,001</b>
VR (Variación del Retardo)	1	0,0007	0,0007	0,00	0,977
PP (Pérdida de Paquetes)	1	6,5219	6,5219	8,30	0,016
Cuadrado	3	16,6433	5,5478	7,06	0,008
<b>R*R</b>	<b>1</b>	<b>11,2310</b>	<b>11,2310</b>	<b>14,30</b>	<b>0,004</b>
VR*VR	1	1,4692	1,4692	1,87	0,201
PP*PP	1	0,6446	0,6446	0,82	0,386
Interacción de 2 factores	3	1,7027	0,5676	0,72	0,561
R*VR	1	0,1158	0,1158	0,15	0,709
R*PP	1	0,2483	0,2483	0,32	0,586
VR*PP	1	1,3385	1,3385	1,70	0,221
Error	10	7,8557	0,7856		
Falta de ajuste	5	3,7772	0,7554	0,93	0,533
Error puro	5	4,0785	0,8157		
Total	19	50,0966			

GL: Grados de Libertad.

SC Ajust.: Suma de Cuadrados

MC Ajust.: Media de los Cuadrados

De los datos de la tabla anterior se puede concluir que los parámetros que tienen influencia significativa en el modelo son el retardo, la pérdida de paquetes y el retardo al cuadrado.

Para el análisis de varianza de los datos de la Tabla 11, el R-cuadrado obtenido fue de 84.23%, lo cual indica que el modelo explica el 84,23% la variabilidad de los datos obtenidos.

Por lo tanto, el modelo obtenido mediante el uso de la herramienta Minitab y los cálculos presentados en la sección 4.1.3, para el video Cafetería es el presentado en la Ecuación 17, donde Retardo es representado por  $R$ , Variación del Retardo por  $V_R$  y Pérdida de paquetes por  $P_P$ . Este modelo se identificará como **Modelo Video Movimiento Moderado (MVMM)**.

$$\begin{aligned}
 QoE = & 3,89 - 3,19 \times 10^{-3} R + 85 V_R + 18,1 P_P + 6 \times 10^{-6} R^2 \\
 & - 0,082 R V_R - 11,937 \times 10^3 V_R^2 + 0,012 R P_P \\
 & + 2020 V_R P_P - 79,1 P_P^2
 \end{aligned}
 \tag{Ecuación 17}$$

El modelo presentado en la Ecuación 17 (MVMM –video cafetería), se le incluyeron todos los factores que tenían influencia significativa y lo que no tenían influencia o su influencia era muy baja. Esto se hizo debido a que al retirar factores de influencia significativa baja del análisis de varianza y generar de nuevo el modelo, el R-cuadrado bajaba a un 79.63%. Por tal razón se decide dejar el modelo con todos los factores, ya que los factores que se pueden retirar del modelo aportan de alguna manera en mejorar el R-cuadrado y mejorar la precisión del modelo.

Para analizar el comportamiento del MVMM-video cafetería se presenta en la Figura 18 la gráfica de contorno de los parámetros que influyen en el modelo (Retardo y Pérdida de paquetes). Como se puede apreciar hay dos partes en la gráfica que se obtienen valores de QoE inferiores a 4.5, y con valores inferiores en Retardo a 1000 mseg y perdida de paquetes de 0.1% se logra obtener valores de VQM inferiores a 5.7. Como se mencionó anteriormente entre menor sea el valor obtenido de VQM mejor será la calidad del video.

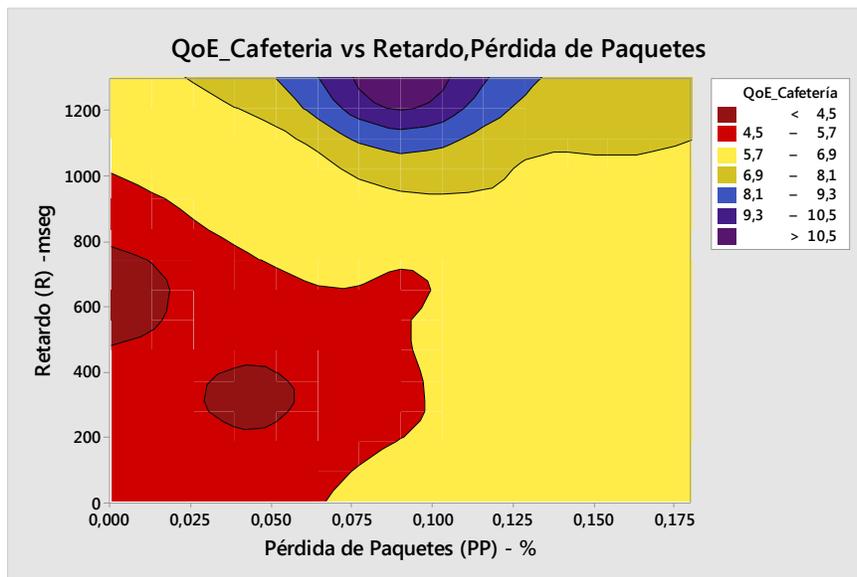


Figura 18. Gráfica de contorno QoE\_Cafetería vs Retardo, Perdida de Paquetes.

En la Figura 19 y Figura 20 se puede apreciar que la variación del retardo tiene un aporte mínimo al modelo ya que ante cambios de su valor la QoE permanece

constante, como se puede apreciar en el área roja de las figuras. Adicionalmente se aprecia que la medida obtenida de QoE está entre 4.5 y 5.7 indicando que el video es de buena calidad.

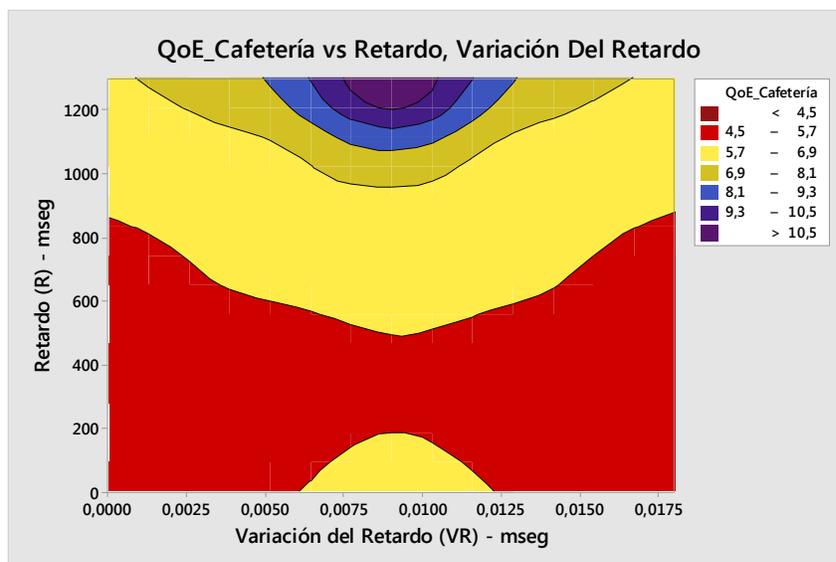


Figura 19. Gráfica de contorno QoE\_Cafetería vs Retardo, Variación del Retardo.

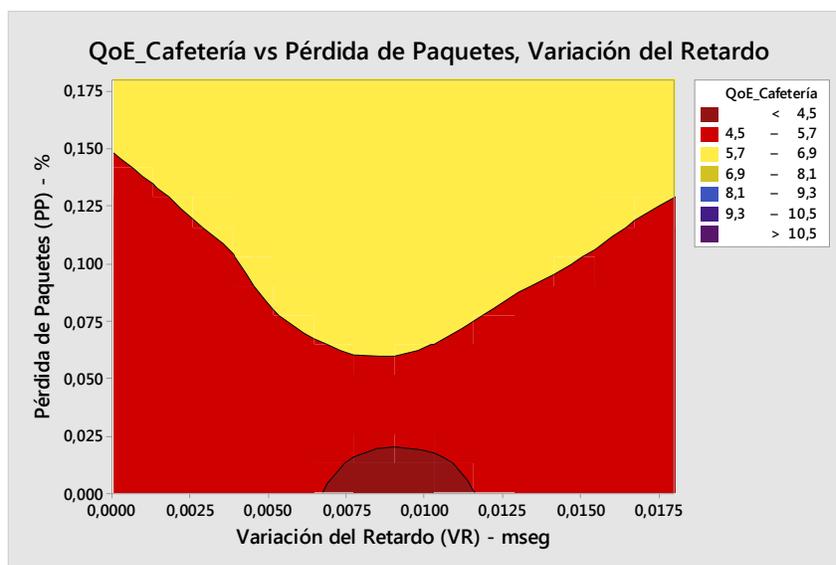


Figura 20 . Gráfica de contorno QoE\_Cafetería vs Perdida de Paquetes, Variación del Retardo.

Adicionalmente, se puede apreciar en la Figura 18, que los parámetros que tienen influencia significativa en el modelo, son los que hacen que el valor de QoE se mejore o se degrade. Para este video la pérdida de paquetes influye en los resultados a

obtener con respecto a la QoE. Por su parte en la Figura 19 y la Figura 20, los parámetros que no tienen influencia sobre el modelo, hacen que la QoE se mantenga constante, así el valor del parámetro este variando a lo largo de su rango definido.

#### 4.1.6.3 Modelo obtenido para el video Baloncesto

De igual manera que para los videos cafetería y entrevista, los datos obtenidos en la ejecución del experimento para el video Baloncesto se presentan en la Tabla 12. Nuevamente se presentan valores de QoE obtenidos diferentes a los otros dos videos debido a las características espacio-temporales del video Baloncesto, el cual es un video con alto movimiento.

Tabla 12. Datos obtenidos para generar el modelo del video Baloncesto.

No. Tratamiento	Retardo (mseg)	Variación del Retardo (mseg)	Pérdida de paquetes (%)	QoE_Basket VQM
1	325	0,0045	0,045	4,543043333
2	975	0,0045	0,045	5,245216667
3	325	0,0135	0,045	5,873283333
4	975	0,0135	0,045	6,579756667
5	325	0,0045	0,135	7,375166667
6	975	0,0045	0,135	6,468343333
7	325	0,0135	0,135	7,220733333
8	975	0,0135	0,135	6,587893333
9	0	0,009	0,09	6,402543333
10	1300	0,009	0,09	6,351253333
11	650	0	0,09	6,00095
12	650	0,018	0,09	5,267866667
13	650	0,009	0	4,170376667
14	650	0,009	0,18	6,783943333
15	650	0,009	0,09	6,617116667
16	650	0,009	0,09	6,258183333
17	650	0,009	0,09	6,548003333
18	650	0,009	0,09	6,857236667
19	650	0,009	0,09	6,5089
20	650	0,009	0,09	6,4944

En la Figura 21 se presenta el diagrama de barras de los datos obtenidos en la Tabla 12, con su coeficiente de variación (Coef. Variación=13,27), la media de los datos y una desviación estándar por encima y por debajo de la media (Desviación estándar =0,824). Para este video el coeficiente de variación obtenido fue bajo, inferior al de los otros dos videos, lo que indica que los resultados obtenidos en cada tratamiento son similares y esto se reflejará en el análisis de varianza que se presentará más adelante.

Estos resultados también se deben al tipo de video, donde la pérdida de paquetes hace que el video se degrade ante la pérdida de cuadros.

Un valor de VQM bajo (4,17) indica una calidad del video mejor en comparación de un valor superior de VQM (7,375) que indica un video de calidad deficiente. En el siguiente capítulo se explicará cómo se mapeo el valor VQM con la escala MOS con el objetivo de realizar pruebas comparativas de desempeño.

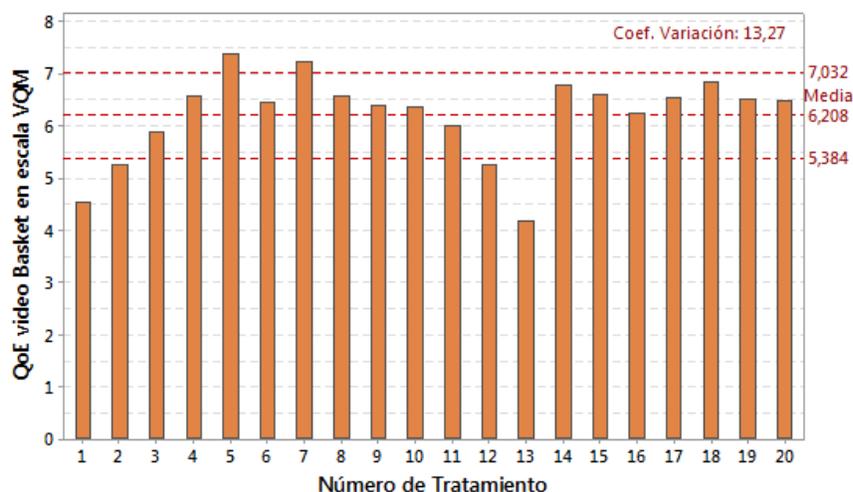


Figura 21. Diagrama de barras de la QoE obtenida para el video Entrevista en escala VQM.

En la Tabla 13 se presenta el análisis de varianza de los datos presentados en la Tabla 12, estos valores se obtuvieron mediante el uso de la herramienta Minitab.

Tabla 13. Análisis de Varianza para los datos obtenidos al video Baloncesto.

Fuente	GL	SC Ajust.	MC. Ajust.	Valor F	Valor p
Modelo	9	11,4226	1,26918	8,60	0,001
Lineal	3	7,1610	2,38698	16,18	0,000
R	1	0,0034	0,00341	0,02	0,882
VR	1	0,0846	0,08464	0,57	0,466
<b>PP</b>	1	7,0729	7,07290	47,95	0,000
Cuadrado	3	2,2544	0,75145	5,09	0,021
R*R	1	0,0121	0,01206	0,08	0,781
VR*VR	1	1,0828	1,08278	7,34	0,022
PP*PP	1	1,5319	1,53187	10,39	0,009
Interacción de 2 factores	3	2,0073	0,66909	4,54	0,030
R*VR	1	0,0097	0,00968	0,07	0,803
R*PP	1	1,0866	1,08657	7,37	0,022
VR*PP	1	0,9110	0,91102	6,18	0,032
Error	10	1,4750	0,14750		
Falta de ajuste	5	1,2862	0,25723	6,81	0,028
Error puro	5	0,1888	0,03776		
Total	19	12,8976			

GL: Grados de Libertad.  
 SC Ajust.: Suma de Cuadrados  
 MC Ajust.: Media de los Cuadrados

De los datos de la tabla anterior se puede concluir que el parámetro que tiene influencia significativa en el modelo es la pérdida de paquetes, el cuadrado de la variación del retardo y las interacciones entre los factores retardo y pérdida de paquetes y la interacción entre variación del retardo y la pérdida de paquetes.

Para el análisis de varianza de los datos de la Tabla 13, el R-cuadrado obtenido fue de 88,56%, lo cual indica que el modelo explica el 88,56% la variabilidad de los datos con que se generó el modelo.

Por lo tanto, el modelo obtenido mediante el uso de la herramienta Minitab para el video Baloncesto es el presentado en la Ecuación 18, donde Retardo es representado por  $R$ , Variación del Retardo por  $V_R$  y Pérdida de paquetes por  $P_P$ . Este modelo se identificará como **Modelo Video Movimiento Alto (MVMA) video Baloncesto**.

$$QoE = 0,55 + 2,28 \times 10^{-3}R + 355V_R + 68,1P_P + 0,0238RV_R - 10248V_R^2 - 0,02520RP_P - 1666V_RP_P - 121,9P_P^2 \quad \text{Ecuación 18}$$

De manera similar al modelo obtenido para el video cafetería e entrevista, el modelo presentado en la Ecuación 18 (MVMA – video baloncesto), se le incluyeron todos los factores, así estos por el análisis de varianza realizado no aportaran al modelo. Esto se hizo debido a que al retirar dichos factores del análisis de varianza y generar de nuevo el modelo, el R-cuadrado bajaba a un 87,88%. Aunque el valor del R-cuadrado casi no varía se decide dejar, como en los otros modelos, todos los factores con el fin de poder realizar más adelante comparaciones entre los resultados obtenidos con los diferentes modelos.

En la Figura 22 y Figura 23 se graficó el parámetro pérdida de paquetes que es el que estadísticamente influye en el modelo y se puede apreciar que se logra una medición de QoE entre 4.5 y 5 cuando el retardo está en un rango de 400 a 800mseg y la variación del retardo se encuentra entre 0.005 e inferior a 0.0125.

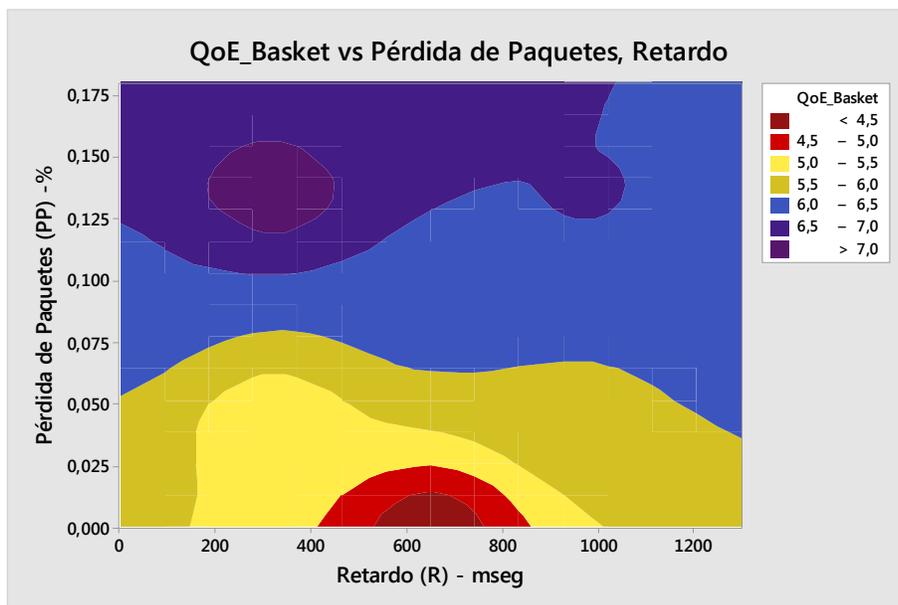


Figura 22. Gráfica de contorno QoE\_Baloncesto vs Pérdida de Paquetes, Retardo.

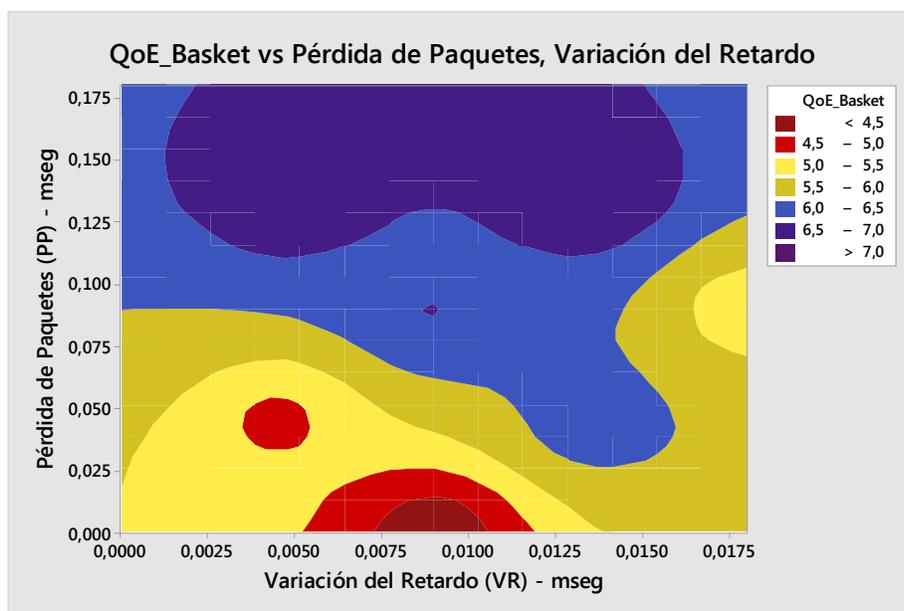


Figura 23. Gráfica de contorno de QoE\_Baloncesto vs Pérdida de paquetes, Variación del retardo.

En la Figura 24 se puede apreciar que los dos parámetros (retardo y variación del retardo) que no influyen estadísticamente en el modelo no presentan intervalos donde

se logre apreciar una QoE inferior a 5.0, indicando una influencia muy baja en el modelo planteado en la Ecuación 18.

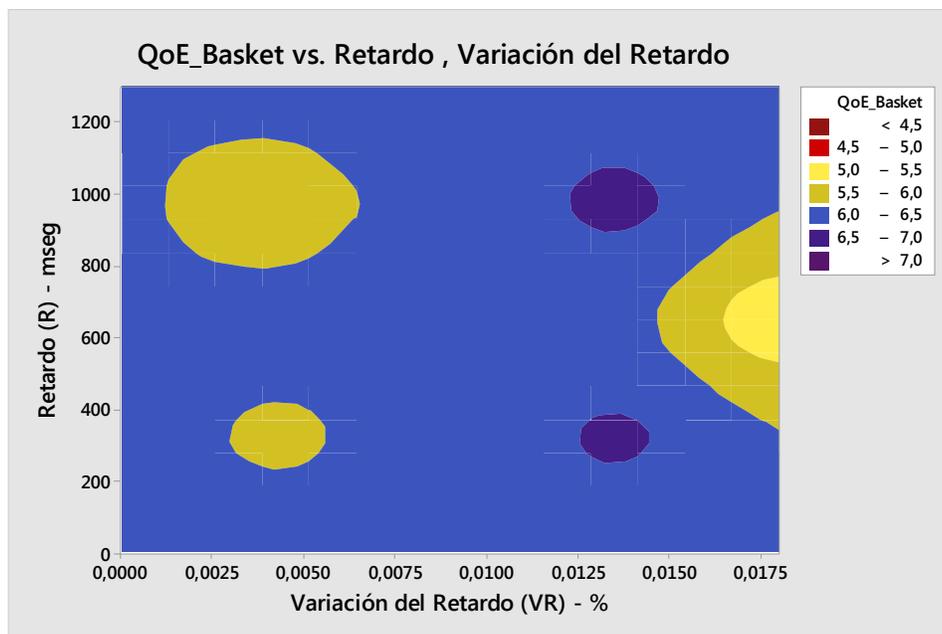


Figura 24. Gráfica de contorno de QoE\_Baloncesto vs Retardo, Variación del retardo.

Al igual que en el video de cafetería y entrevista, el parámetro que más influye en la QoE es la pérdida de paquetes, solo que para este tipo de video de alto movimiento el rango para obtener una QoE aceptable es mucho menor en comparación a los otros dos videos. Esto debido a las características espacio – temporales del video Baloncesto, donde la pérdida de un cuadro I afecta considerablemente la QoE asociada al video.

## 4.2 Comentarios del capítulo

En la Tabla 14 se presenta el resumen del modelo obtenido para cada video. En las filas se presentan los coeficientes asociados a cada término del modelo, los factores que estadísticamente influyen y el respectivo R-cuadrado calculado para cada modelo.

Se aprecia que el factor que influye estadísticamente en los tres videos es la pérdida de paquetes y como se ha mencionado anteriormente este factor es de los más críticos al momento de transmitir IPTV. Esto se debe porque al no existir retransmisión de cuadros, el codificador debe recuperar los cuadros perdidos basándose en la información que llega en el cliente. Esta recuperación de cuadros en algunos casos no es eficaz haciendo que en el video del cliente se produzca congelamiento y pérdida de nitidez. Causando que se degrade la QoE asociada al video.

Por otra parte dependiendo de las características espacio-temporales de cada video, en unos afectará más la pérdida de paquetes, como es el caso del video de alto movimiento, donde el rango para obtener una QoE aceptable es menor al de los otros dos videos.

Tabla 14. Coeficientes de los tres modelos obtenidos.

Factor - Modelo	Entrevista	Cafetería	Baloncesto
<b>Término independiente</b>	0,65	3,89	0,55
<b>R</b>	$0,35 \times 10^{-3}$	$-3,19 \times 10^{-3}$	$2,28 \times 10^{-3}$
<b>V<sub>R</sub></b>	254	85	355
<b>P<sub>P</sub></b>	-5,8	18,1	68,1
<b>R<sup>2</sup></b>	$-1 \times 10^{-6}$	$6 \times 10^{-6}$	0,0
<b>V<sub>R</sub><sup>2</sup></b>	-8890	-11937	-10248
<b>P<sub>P</sub><sup>2</sup></b>	30,1	-79,1	-121,9
<b>R*V<sub>R</sub></b>	-0,069	-0,082	0,0238
<b>R*P<sub>P</sub></b>	0,0275	0,012	-0,0252
<b>V<sub>R</sub>*P<sub>P</sub></b>	-698	2020	-1666
<b>Factores que influyen estadísticamente en el modelo</b>	<b>P<sub>P</sub>, R*P<sub>P</sub></b>	<b>R, P<sub>P</sub>, R<sup>2</sup></b>	<b>P<sub>P</sub>, V<sub>R</sub><sup>2</sup>, R*P<sub>P</sub>, V<sub>R</sub>*P<sub>P</sub></b>
<b>R_cuadrado</b>	<b>79,59%</b>	<b>84,23%</b>	<b>88,56%</b>

Por otra parte, proponer un único modelo que permita evaluar la QoE para el servicio de IPTV, hace que el modelo se deba ajustar introduciendo una serie de constantes. Estas constantes son necesarias para lograr que el R-cuadrado del modelo siga siendo alto, y en algunos casos las constantes calculadas se basan en los videos utilizados para generar los datos o en condiciones del entorno del usuario.

La situación anteriormente descrita se puede apreciar en [45], donde los autores ponderan el peso de cada parámetro de QoS con una constante que es calculada utilizando un ambiente de computación invasiva [47] y adicionalmente tienen en cuenta otras constantes dentro del modelo como:

- Constante de acuerdo al tamaño de la pantalla del usuario.
- Constante de acuerdo del tipo de servicio que el cliente ha contratado con el proveedor de servicios.
- Constante de acuerdo al tipo de cuadro (I, B o P) perdido. Donde al perderse un cuadro I hace que la QoE se degrade.

En [48], los autores para realizar cálculos para obtener la QoE final, se basan en una serie de 9 constantes calculadas utilizando pruebas subjetivas. Estas constantes se incorporan al modelo y solo son válidas para los videos utilizados en el experimento.

Con los modelos propuestos planteados en este proyecto no es necesario considerar ningún tipo de variable externa (tamaño de tv, tipo de servicio, variables calculadas mediante pruebas subjetivas, etc.). Esto permitiría que el modelo se aplique para

cualquier tipo de video en tiempo real solo obteniendo los valores de los parámetros de QoS (retardo, variación del retardo y pérdida de paquetes).

Basándonos en lo presentado en la sección 3.4 página 45 y Tabla 14, para utilizar cada uno de los modelos propuestos se deben conocer las características espacio-temporales del video a evaluar, con el fin de elegir el modelo más adecuado para evaluar la QoE. En la revisión sistemática que se realizó a más de 14 bases de datos, y por las especificaciones técnicas de la cámara utilizada para grabar los videos fuente para generar cada modelo, se definieron los siguientes rangos para para elegir el modelo adecuado al tipo de video (movimiento en el video):

- Movimiento Bajo IE:30-40 IT:5-10
- Movimiento Medio: IE:60-70 IT:5-10
- movimiento Alto: IE:50-60 IT:15-20

Es claro mencionar que los valores obtenidos por el modelo hasta este momento, solo permiten dar indicios de la calidad del video, es decir, un video con un valor de QoE (QoE en escala VQM) bajo (0-4) tendrá menos distorsiones que un video con un valor alto de QoE (5-10). En el siguiente capítulo se presentará la escala con la cual se mapea el valor obtenido por nuestros modelos con la escala MOS.

Esto permitirá obtener una medición de QoE fácil de interpretar en escala de MOS, y posibilita comparar las respuestas de los modelos propuestos con modelos estandarizados.

## 5 Análisis de Evaluación de Desempeño de los Modelos Propuestos

En este capítulo se presenta el resultado de las pruebas para evaluar el desempeño de los modelos propuestos al momento de medir la calidad de la experiencia asociada al video. En la primera sección se presenta la estandarización de las medidas que entregan los modelos con el fin de poder realizar pruebas comparativas de desempeño. En la segunda sección se presentan los resultados de las pruebas comparativas de desempeño con dos modelos objetivos (PSNR y SSIM). En la tercera sección se presentan las pruebas de desempeño realizadas basándose en el protocolo de pruebas propuesto por el Grupo de Expertos de Calidad del Video (*Video Quality Expert Group –VQEG*) [88].

Al finalizar el capítulo, se presentan pruebas estáticas realizadas a cada modelo que pretenden analizar su comportamiento y estabilidad. En la Figura 25, se ilustra el contenido esquemático de este capítulo.

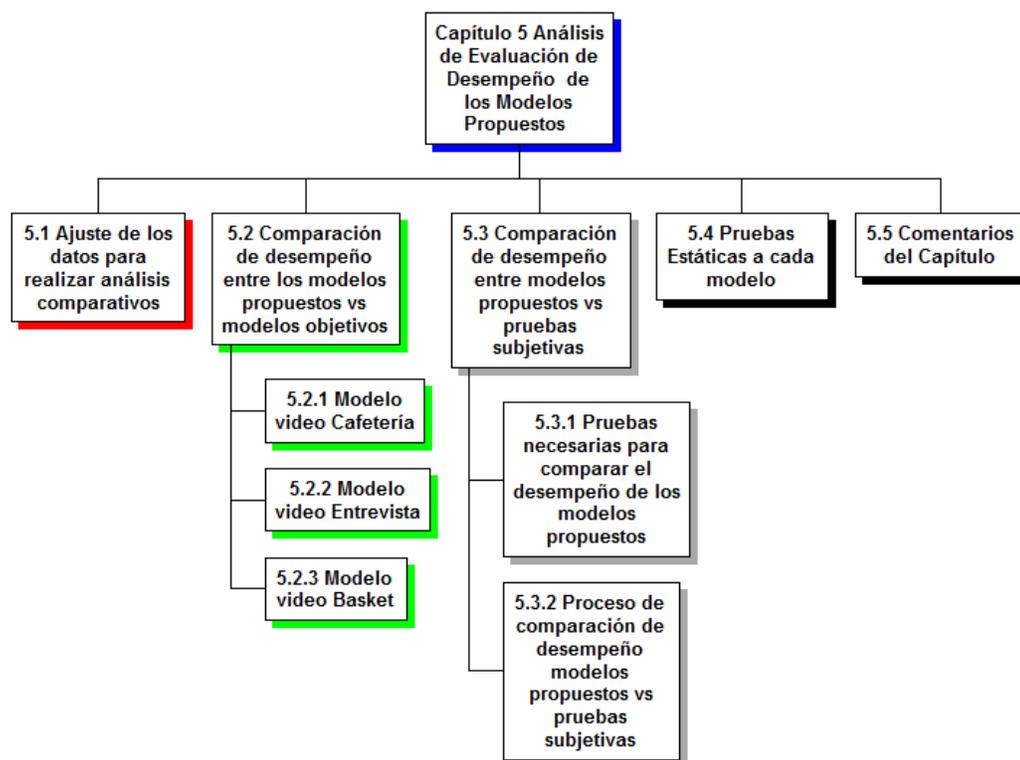


Figura 25. Descripción del contenido del capítulo 5.

## 5.1 Ajuste de los datos para realizar los análisis comparativos

De acuerdo a la revisión sistemática realizada en el capítulo 2 y lo que plantean en [89] se propone el siguiente procedimiento para efectuar los análisis de desempeño de los modelos propuestos: (i) seleccionar el conjunto de datos con los cuales se van a desarrollar las comparaciones, (ii) estandarizar o normalizar todos los datos a comparar a una escala común, (iii) ejecutar las comparaciones utilizando herramientas estadísticas que permitan deducir el desempeño o precisión del modelo.

De acuerdo al procedimiento descrito en el párrafo anterior se define trabajar con los primeros 15 tratamientos realizados a cada video, esto porque del tratamiento 16 a 20, la configuración de los parámetros de QoS es igual. Adicionalmente, esta elección va alineada con las pruebas subjetivas que se realizaron a los videos utilizados para generar los datos y obtener los modelos. Estas pruebas subjetivas se explicarán posteriormente y están documentadas en el Anexo 1.

Dentro de las pruebas subjetivas se incluye el tratamiento cero que es el tratamiento donde el video no tiene ningún tipo de distorsión.

En la Tabla 15 se presentan los resultados que arroja cada modelo propuesto (Ecuación 16, Ecuación 17, Ecuación 18) al momento de evaluar los valores de parámetros de QoS en cada tratamiento. Se pueden observar los resultados para el tratamiento cero y los 15 tratamientos seleccionados. Los valores obtenidos por cada modelo presentados en la Tabla 15, son valores en escala VQM, es decir no se han mapeado a la escala MOS para su análisis.

Tabla 15. Resultados obtenidos al evaluar cada modelo con los parámetros de QoS definidos para cada tratamiento. La QoE obtenida está en escala VQM.

No. Tratamiento	Retardo (mseg)	Variación del Retardo (mseg)	Pérdida de paquetes (%)	QoE MVMB - video Entrevista	QoE MVMM - video Cafetería	QoE MVMA - video Baloncesto
0	0	0	0	0,650	3,890	0,550
1	325	0,0045	0,045	1,353	4,396	4,828
2	975	0,0045	0,045	0,884	6,801	5,642
3	325	0,0135	0,045	1,715	3,805	5,757
4	975	0,0135	0,045	0,841	5,731	6,711
5	325	0,0045	0,135	1,841	5,210	7,570
6	975	0,0045	0,135	2,980	6,914	6,910
7	325	0,0135	0,135	1,637	6,256	7,150
8	975	0,0135	0,135	2,372	7,480	6,630
9	0	0,009	0,09	1,372	6,313	6,707
10	1300	0,009	0,09	1,638	9,942	7,001

No. Tratamiento	Retardo (mseg)	Variación del Retardo (mseg)	Pérdida de paquetes (%)	QoE MVMB - video Entrevista	QoE MVMM - video Cafetería	QoE MVMA - video Baloncesto
11	650	0	0,09	1,331	4,638	5,699
12	650	0,018	0,09	1,084	4,613	6,349
13	650	0,009	0	1,162	3,670	4,536
14	650	0,009	0,18	3,180	6,233	7,197
15	650	0,009	0,09	1,927	5,592	6,854

Se puede apreciar en la Tabla 15 que el tratamiento cero, es decir, sin afectar los valores en los parámetros de calidad de servicio se obtiene un valor de QoE bajo, indicando que dicho video tiene una calidad superior a videos con un valor de QoE superior. En el caso de la QoE obtenida para el MVMB-video entrevista se puede apreciar que es menor que para los otros dos modelos, esto se debe a las características espacio-temporales del video, donde debido a su composición la degradación del video no es tan notoria como para los otros dos tipos de videos analizados.

Como se describió al inicio de este capítulo, ya seleccionados el conjunto de datos (videos de las pruebas subjetivas) con los cuales se van a desarrollar las comparaciones, se procede a estandarizar todos los datos (medida de QoE que entrega cada modelo) a una escala común. Esto con el fin de interpretar fácilmente la respuesta que entrega el modelo, es decir, definir la QoE asociada al video en escala de MOS y realizar pruebas de fidelidad con otros modelos objetivos.

Para este proceso de estandarización de las medidas de QoE que entrega cada modelo se analizaron y ejecutaron dos técnicas: (i) el análisis de conglomerados (clusters) [90] y (ii) VQAMap [89].

Con respecto al **análisis de clusters**, que es una técnica que busca agrupar variables de un conjunto de datos tratando de lograr homogeneidad en cada grupo, se encontró que esta técnica no era la adecuada para realizar este proceso de estandarización. Debido que se requería utilizar un  $K=5$  (los 5 niveles de la escala MOS), pero la medida de ajuste identificando el número de clusters original, no se minimiza para  $K=5$ , si no para  $K=2$  o  $K=3$ , dependiendo de la dispersión de los valores de QoE que se obtienen para cada tipo de video. Adicionalmente las pruebas de correlación que se realizaron entre las pruebas subjetivas y los resultados que arrojan esta técnica fueron inferiores a las obtenidas con el procedimiento VQAMap. Por esta razón la técnica de clusters no se empleó para realizar la estandarización de los valores de QoE asociadas a cada video.

Después de analizado lo anterior, se procede a utilizar el procedimiento planteado en [89] llamado **VQAMap**, que permite crear escalas de manera genérica con el fin de mapear valores de medidas que entreguen diferentes modelos propuestos a la escala MOS. Este procedimiento se diseñó de manera específica para ser utilizado en evaluación de video y es enfocado para modelos basados en referencia completa, pero

los autores indican que también se puede emplear para modelos basados en referencia reducida y sin referencia.

Para empezar el proceso para obtener el mapeo entre los valores que entrega los modelos propuestos a la escala MOS, se debe realizar el siguiente procedimiento:

1. Realizar una regresión no lineal entre los valores que entrega cada modelo propuesto (en escala VQM, ver Tabla 15) y las medidas subjetivas ( $MOS_K$ ) de los videos utilizados para generar los datos.

Una función de mapeo no lineal es la que se utiliza y se ha demostrado empíricamente que funciona bien en diferentes bases de datos donde se ha empleado [91] [92]. Esta función es un polinomio cubico de la forma presentada en la Ecuación 19.

$$MOS_K^{(p)} = a + b VQM_K^3 + c VQM_K^2 + d VQM_K \quad \text{Ecuación 19}$$

Donde  $MOS_K^{(p)}$  representa el valor calculado de la calidad percibida por el usuario y  $VQM_K$  es la medida que entrega el modelo propuesto para el tratamiento  $K$ .  $a$ ,  $b$ ,  $c$  y  $d$  son coeficientes calculados para minimizar la suma de los errores cuadrados entre el vector de valores subjetivos  $MOS_K$  y el valor de  $MOS_K^{(p)}$  calculados.

2. Calcular el R\_cuadrado de la bondad de ajuste del modelo obtenido mediante la regresión no lineal, este valor del R\_cuadrado se utiliza para indicar el grado por el cual las variaciones en los valores MOS son contabilizadas por el modelo de regresión no lineal [91]. Un valor de R\_cuadrado más cercano a 1 indica un mejor ajuste, pero dependiendo del número de puntos de datos, incluso un R\_cuadrado de 0,25 podría considerarse aceptable [93]. El modelo también debe ser estadísticamente significativo (por ejemplo,  $p < 0,05$  a un nivel de confianza del 95%), es decir cuando se utiliza el software Minitab para realizar la regresión no lineal el nivel de confianza debe configurarse al 95%.
3. Realizar una interpolación inversa al polinomio de la regresión no lineal, con el fin de calcular los límites inferior y superior de la escala a utilizar para el modelo propuesto. Adicional a esto, calcular el valor para cada nivel de MOS y con estos valores establecer la escala para mapear los valores  $VQM_K$  a MOS.

En la Tabla 16, se presentan los valores obtenidos de las pruebas subjetivas realizadas a cada tipo de video y en la columna siguiente el valor que entrega en escala VQM cada modelo propuesto. Basándose en estos datos se procede a realizar el procedimiento descrito arriba, se empieza realizando la regresión no lineal para cada tipo de video utilizando el software estadístico Minitab.

Tabla 16. Valores de las pruebas subjetivas realizadas a los tres videos con su correspondiente valor en escala VQM obtenido por cada modelo propuesto.

No. Tratamiento	MOS Prob subj Entrevista	QoE MVMB Entrevista	MOS Prob subj Cafetería	QoE MVMM Cafetería	MOS Prob subj Baloncesto	QoE MVMA Baloncesto
0	4,500	0,650	4,211	3,890	3,640	0,550
1	4,061	1,353	4,123	4,396	4,896	4,828
2	4,096	0,884	3,509	6,801	4,430	5,642
3	3,474	1,715	2,640	3,805	3,070	5,757
4	3,632	0,841	3,044	5,731	3,123	6,711
5	4,228	1,841	3,544	5,210	3,044	7,570
6	2,667	2,980	3,939	6,914	3,229	6,910
7	3,825	1,63	2,939	6,256	2,798	7,150
8	3,491	2,372	3,298	7,480	3,254	6,630
9	4,333	1,372	2,693	6,313	3,553	6,707
10	3,509	1,638	2,193	9,942	3,254	7,001
11	4,237	1,331	3,640	4,638	3,228	5,699
12	4,342	1,08	3,009	4,613	3,211	6,349
13	3,693	1,162	3,798	3,670	4,746	4,536
14	3,044	3,180	2,947	6,233	3,307	7,197
15	4,211	1,927	3,079	5,592	3,474	6,854

En la Tabla 17 se presenta el valor de las constantes  $a$ ,  $b$ ,  $c$ ,  $d$  y valor del  $R_{\text{cuadrado}}$  calculado para cada modelo después de realizar la regresión no lineal para los datos de la Tabla 16.

Tabla 17. Valores de las constantes de  $a, b, c, d$  y valor del  $R_{\text{cuadrado}}$  de la regresión no lineal para cada modelo propuesto.

Modelo	Video	a	b	c	d	$R_{\text{cuadrado}}$
MVMB	Entrevista	4,35632	-0,0599065	0,142379	-0,339642	0.59
MVMM	Cafetería	10,7662	-0,0289677	0,562472	3,58745	0.32
MVMA	Baloncesto	31,5951	-0,0735298	1,58238	-11,5316	0.72

Se puede apreciar que el  $R_{\text{cuadrado}}$  más bajo corresponde al modelo video movimiento moderado (MVMM) asociado al video Cafetería debido a la dispersión de los datos. Esto se logra corroborar al calcular la varianza de los datos presentados en la Tabla 16 donde la varianza más alta corresponde a las medidas VQM entregadas por el MVMM-video cafetería. En la Figura 26, Figura 27, y Figura 28 se pueden apreciar las gráficas ajustadas en la regresión no lineal realizada a los datos de cada modelo. Como se comentó anteriormente el modelo con menor bondad de ajuste es el MVMM (video Cafetería). Los otros dos modelos presentan una bondad de ajuste superior a 0.5, lo cual indica que el modelo se ajusta de manera adecuada a los datos.

Con respecto a la regresión no lineal presentada en la Figura 26 hubo la necesidad de descartar la pareja de datos del tratamiento cero, este dato era un dato atípico que hacía que el valor de  $R^2$  cuadrado fuera inferior al 20% lo cual hacía que el modelo obtenido no fuera estadísticamente válido. Este dato atípico fue producto de la evaluación subjetiva que se le realizó a dicho video, donde por circunstancias ajenas al experimento, el video fue calificado de manera regular por las personas que participaron de la prueba subjetiva.

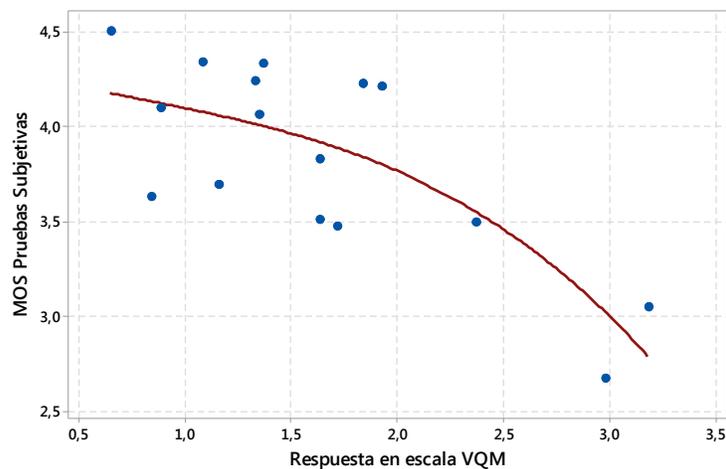


Figura 26. Gráfica de la regresión no lineal correspondiente al MVMB - video Entrevista.

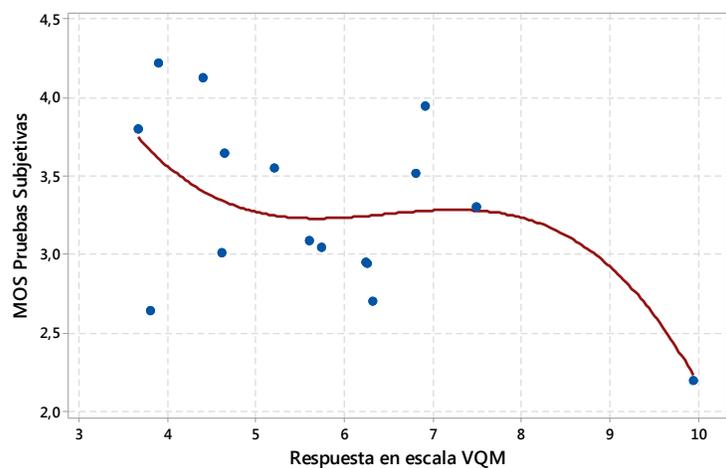


Figura 27. Gráfica de la regresión no lineal correspondiente al MVMM - video Cafetería.

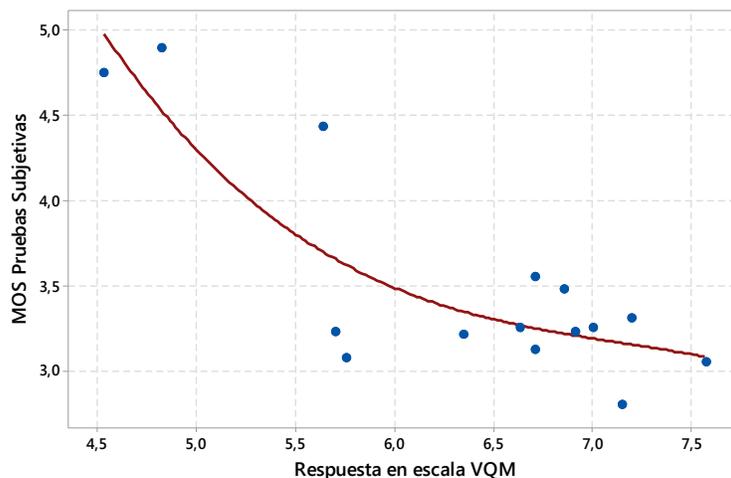


Figura 28. Gráfica de la regresión no lineal correspondiente al MVMA - video Baloncesto.

Después de obtener la regresión lineal para cada modelo (ver Tabla 17), se procede a realizar la interpolación inversa, eligiendo 4 parejas de datos por cada modelo. Posterior a esto, a cada grupo de datos se le aplica el método de interpolación de Newton y el resultado obtenido permite definir los rangos para realizar el mapeo de escala VQM a escala MOS.

En la Tabla 18 se presentan las escalas definidas para mapear el resultado de cada modelo propuesto a escala MOS. La variación de los valores de cada escala, depende del tipo de video en lo referente a las características espacio-temporales.

Tabla 18. Escalas para mapear MOS vs VQM para cada video.

MOS	Modelo - Video		
	MVMB - Entrevista	MVMM -Cafetería	MVMA -Baloncesto
5,0	$\geq 0$ & $< 1,38$	$\geq 2,3969$ & $< 3,3145$	$\geq 0$ & $< 5,279$
4,0	$\geq 1,38$ & $< 3,00$	$\geq 3,3146$ & $< 8,8325$	$\geq 5,279$ & $< 7,8857$
3,0	$\geq 3,00$ & $< 3,71$	$\geq 8,8325$ & $< 10,1588$	$\geq 7,8857$ & $< 9,3556$
2,0	$\geq 3,71$ & $< 4,59$	$\geq 10,1588$ & $< 10,8856$	$\geq 9,3556$ & $< 9,9935$
1,0	$\geq 4,59$	$\geq 10,8856$	$\geq 9,9935$

Ya con los intervalos de cada escala definidos se procede a mapear el valor de QoE obtenido en la Tabla 15, a escala MOS y el resultado de este proceso se puede apreciar en la Tabla 19. A manera de ejemplo el mapeo se realiza de la siguiente manera partiendo de los datos consignados en la Tabla 15:

- Para modelo MVMB-Entrevista, el tratamiento No.1 de valor 0,65 de QoE le corresponde una calificación de 5,0 en escala MOS.
- Para el modelo MVMM-Cafetería el tratamiento No.6 de valor 6,914 le corresponde un valor de 4,0 en escala MOS.

- Para el modelo MVMA-Baloncesto, el tratamiento No.15 de valor 6,349 le corresponde un valor de 4,0 en escala MOS.

Tabla 19. Medidas obtenidas de cada modelo en escala MOS.

No. Tratamiento	MOS MVMB Entrevista	MOS MVMM Cafetería	MOS MVMA Baloncesto
0	5,0	4,0	Descartado
1	5,0	4,0	5,0
2	5,0	4,0	4,0
3	4,0	4,0	4,0
4	5,0	4,0	4,0
5	4,0	4,0	4,0
6	4,0	4,0	4,0
7	4,0	4,0	4,0
8	4,0	4,0	4,0
9	5,0	4,0	4,0
10	4,0	3,0	4,0
11	5,0	4,0	4,0
12	5,0	4,0	4,0
13	5,0	4,0	5,0
14	3,0	4,0	4,0
15	4,0	4,0	4,0

Ya con los resultados que entregan los modelos propuestos estandarizados a la escala MOS, en las siguientes secciones se explicará el proceso realizado para analizar el desempeño de los modelos propuestos comparándolos contra dos modelos objetivos y posteriormente comparándolos con las pruebas subjetivas siguiendo las recomendaciones del Grupo de Expertos de Calidad del Video (*Video Quality Expert Group –VQEG*) [88][94].

## 5.2 Comparación de desempeño entre modelos propuestos vs modelos objetivos

En esta sección se presentará la comparación de desempeño entre las respuestas que entregan los tres modelos propuestos y dos modelos objetivos seleccionados. Se dará una breve introducción a los dos modelos objetivos seleccionados, para posteriormente normalizarlos para efectos de la comparación.

Los modelos objetivos seleccionados para estas comparaciones de desempeño fueron PSNR (*Peak Signal to Noise Ratio*) y SSIM (*Structural SIMilarity*), estos modelos son de amplio uso para la evaluación de video ([95][75][15][96][97][77][78][69][98]

[99][49][68] ) y son los más usados al momento de comparar el desempeño de nuevos modelos propuestos.

PSNR es un modelo de referencia completa basado en comparar pixel a pixel de cada cuadro, se evalúa mediante una función simple del Error medio cuadrado (*Mean Squared Error* – MSE) entre el video de referencia y el video a evaluar [100]. La ventaja de este modelo es su simplicidad en el cálculo, aunque presenta valores de correlación muy bajos con pruebas subjetivas [15].

SSIM fue propuesto in [101] y desarrollado en [52], captura la estructura de la imagen. Esto permite medir la diferencia de estructura entre la imagen original y la distorsionada. Tiene en cuenta tres factores: luminancia, contraste y estructura. La luminancia y el contraste son afectados principalmente por la iluminación del entorno, mientras que la estructura es la característica intrínseca del objeto [15].

Ya con los modelos objetivos seleccionados, se procede a realizar las mediciones a los 16 videos (tratamiento cero y 15 tratamientos adicionales) utilizando la herramienta *MSU Video Quality Measurement Tool*. Las mediciones obtenidas para PSNR se puede apreciar en la Tabla 20 y las mediciones obtenidas para SSIM están en la Tabla 21.

Tabla 20. Resultados obtenidos al evaluar los 16 tratamientos para los tres videos con PSNR.

No. Tratamiento	Medición PSNR Videos:		
	Entrevista	Cafetería	Baloncesto
0	45,116	38,10627	29,706
1	40,847	32,06711	29,706
2	40,326	32,38903	28,180
3	37,079	30,1646	23,946
4	41,268	29,48377	24,584
5	35,970	29,0406	23,598
6	33,568	32,05032	24,818
7	38,152	27,23993	21,824
8	34,503	27,35013	24,882
9	35,813	27,70518	27,251
10	35,700	23,30279	25,856
11	36,647	29,34001	26,125
12	37,859	28,177	26,698
13	40,384	31,72738	29,551
14	33,593	27,21137	25,157
15	38,736	27,06382	24,560

Tabla 21. Resultados obtenidos al evaluar los 16 tratamientos para los tres videos con SSIM.

No. Tratamiento	Medición SSIM Videos:		
	Entrevista	Cafetería	Baloncesto
0	0,992	0,971	0,951
1	0,987	0,946	0,951
2	0,986	0,947	0,938
3	0,977	0,939	0,872
4	0,988	0,909	0,909
5	0,967	0,911	0,900
6	0,952	0,952	0,911
7	0,979	0,880	0,834
8	0,967	0,876	0,902
9	0,967	0,874	0,923
10	0,970	0,800	0,912
11	0,972	0,909	0,917
12	0,978	0,908	0,930
13	0,986	0,953	0,948
14	0,946	0,891	0,910
15	0,982	0,872	0,904

Los valores de las mediciones obtenidas y presentadas en la Tabla 20 y la Tabla 21 no permiten ningún tipo de comparación, debido a que no están en una escala normalizada. Para realizar las comparaciones de desempeño entre las respuestas que entregan los modelos propuestos con las mediciones obtenidas con PSNR y SSIM, los valores obtenidos por estos dos últimos se deben mapear a la escala MOS. En la Tabla 22 se presenta el mapeo entre la escala MOS y los rangos de medición de PSNR [102] y SSIM [99].

Tabla 22. Mapeo de escala MOS a PSNR y SSIM.

MOS	PSNR (dB)	SSIM
5 (Excelente)	>37	> 0,99
4 (Bueno)	31-37	$\geq 0,95$ & $0,99$
3 (Justo/Regular)	25-31	$\geq 0,88$ & $< 0,95$
2 (Pobre/Mediocre)	20-25	$\geq 0,5$ & $< 0,88$
1 (Malo)	<20	< 0,5

Para proceder con las comparaciones de desempeño entre los tres modelos propuestos y los dos modelos objetivos seleccionados (PSNR y SSIM) se aplica a las respuestas obtenidas de los 16 tratamientos de la Tabla 21 el mapeo planteado en la Tabla 22.

Posteriormente se hace un análisis de correlación entre los resultados de cada modelo propuesto y los valores arrojados por PSNR y SSIM, con el fin de verificar el desempeño de cada modelo propuesto [15]. En este análisis de correlación consiste en calcular el coeficiente de correlación de Pearson (se designa con la letra  $r$ ) y

describe la fuerza de la relación entre dos conjuntos de variables o el grado de covariación entre dos conjuntos de variables relacionadas linealmente [86]. Su valor oscila entre -1,00 y +1,00, donde un -1,00 indica una correlación negativa perfecta y un +1,00 indica una correlación positiva perfecta. En la Figura 29 se presentan los valores que pueden tomar el coeficiente de correlación y su interpretación.



Figura 29. Interpretación de los valores que entrega el coeficiente de correlación de Pearson. Adaptado de [86].

En las siguientes secciones se presentarán los análisis para el modelo generado para cada tipo de video.

### 5.2.1 Modelo Video Movimiento Bajo para el video Entrevista

Para el video entrevista los resultados se presentan en la Tabla 23. La primera columna (MOS Modelo Entrevista) presenta el resultado mapeado a escala MOS que entrega el modelo propuesto al evaluar cada tratamiento, las otras dos columnas entregan los valores de PSNR y SSIM obtenidos al evaluar para cada tratamiento y mapeados a escala MOS.

Se aclara que para lograr el mapeo de escala PSNR y SSIM a MOS, se procede de la siguiente manera:

- Se obtiene la medida en PSNR y SSIM para cada cuadro del video analizado.
- Se mapea la medida obtenida para cada cuadro a escala MOS.
- Se calcula el promedio de los valores obtenidos en escala MOS para todos los cuadros que componen el video.

Es por esta razón que las medidas obtenidas no son en números enteros como se presentó en la Tabla 22.

Tabla 23. Datos obtenidos para evaluar el modelo video movimiento bajo para el video Entrevista.

No. Tratamiento	MOS MVMB Entrevista	MOS PSNR Entrevista	MOS SSIM Entrevista
0	5,0	5,000	4,876
1	5,0	4,888	4,606
2	5,0	4,834	4,544
3	4,0	4,618	4,502
4	5,0	4,967	4,423
5	4,0	4,465	3,979
6	4,0	4,050	3,481
7	4,0	4,855	4,025
8	4,0	4,344	3,905
9	5,0	4,382	4,004
10	4,0	4,407	4,187
11	5,0	4,415	4,315
12	5,0	4,593	4,349
13	5,0	4,959	4,224
14	3,0	4,033	3,382
15	4,0	4,772	4,241

En la Tabla 24 se presenta el resumen del análisis estadístico realizado a las respuestas obtenidas. Se puede apreciar que los datos que arroja el modelo propuesto, con respecto al coeficiente de correlación indican una correlación positiva fuerte. Esto permite indicar que los resultados del modelo propuesto covarían casi en más de un 70% de manera similar con las respuestas obtenidas con PSNR y SSIM. Adicionalmente se puede apreciar que el coeficiente de variación del modelo propuesto es significativamente mayor que los coeficientes calculados para las respuestas obtenidas con PSNR y SSIM. Esto se debe a la manera de evaluar la QoE que efectúa PSNR y SSIM, además se debe tener en cuenta que para el video Entrevista no hay ningún cambio de imagen o movimiento en el video.

Tabla 24. Resultados estadísticos para comparar el desempeño del modelo propuesto para video Entrevista.

Modelo	Media	Desv. Estándar	Coef. de Variación	Coef. Correl. Pearson - PSNR	Coef. Correl. Pearson - SSIM
MVMB	4,438	0,629	14,15	0,611	0,71
PSNR	4,598	0,314	6,84		
SSIM	4,190	0,392	9,36		

### 5.2.2 Modelo Video Movimiento Moderado - Video Cafetería

Para el video cafetería los resultados se presentan en la Tabla 25. La primera columna (MOS MVMM) presenta el resultado mapeado a escala MOS que entrega el modelo

propuesto al evaluar cada tratamiento, las otras dos columnas entregan los valores de PSNR y SSIM obtenidos al evaluar para cada tratamiento y mapeados a escala MOS.

Tabla 25. Datos obtenidos para evaluar el modelo Cafetería.

No. Tratamiento	MOS MVMM Cafetería	MOS PSNR Cafetería	MOS SSIM Cafetería
0	4,0	4,776	4,000
1	4,0	4,154	3,598
2	4,0	4,029	3,593
3	4,0	3,216	3,199
4	4,0	3,311	2,929
5	4,0	3,286	3,112
6	4,0	3,855	3,510
7	4,0	3,133	2,660
8	4,0	3,083	2,357
9	4,0	3,183	2,232
10	3,0	2,000	2,000
11	4,0	3,274	2,909
12	4,0	3,232	3,004
13	4,0	3,780	3,456
14	4,0	3,112	3,000
15	4,0	3,332	2,531

En la Tabla 26 se presenta el resumen del análisis estadístico realizado a las respuestas obtenidas. Se puede apreciar el valor obtenido en el coeficiente de correlación, indican una correlación positiva moderada. Este valor obtenido del coeficiente de correlación indica que hay una correlación positiva moderada entre los resultados que entrega el modelo y las respuestas obtenidas con PSNR y SSIM.

Tabla 26. Resultados estadísticos para comparar el desempeño del modelo video movimiento moderado para el video cafería.

Modelo	Media	Desv. Estándar	Coef. de Variación	Coef. Correl. Pearson - PSNR	Coef. Correl. Pearson - SSIM
MVMM	3,9375	0,2500	6,35	0,623	0,485
PSNR	3,422	0,609	17,80		
SSIM	3,006	0,553	18,39		

### 5.2.3 Modelo Video Movimiento Alto para el Video Baloncesto

Para el video Baloncesto los resultados se presentan en la Tabla 27. La primera columna (MOS MVMA) presenta el resultado mapeado a escala MOS que entrega el

modelo propuesto al evaluar cada tratamiento, las otras dos columnas entregan los valores de PSNR y SSIM obtenidos al evaluar para cada tratamiento y mapeados a escala MOS.

Tabla 27. Datos obtenidos para evaluar el modelo Baloncesto.

No. Tratamiento	MOS MVMA Baloncesto	MOS PSNR Baloncesto	MOS SSIM Baloncesto
1	5,0	3,996	3,548
2	4,0	3,419	3,232
3	4,0	2,685	2,639
4	4,0	2,643	3,062
5	4,0	2,303	2,776
6	4,0	2,714	3,116
7	4,0	2,178	2,232
8	4,0	2,627	2,851
9	4,0	3,174	2,975
10	4,0	2,954	2,971
11	4,0	2,710	2,913
12	4,0	3,141	3,299
13	5,0	3,934	3,539
14	4,0	2,710	2,942
15	4,0	2,689	2,763

En la Tabla 28 se presenta el resumen del análisis estadístico realizado a las respuestas obtenidas. Se puede apreciar que el cálculo del coeficiente de correlación arroja una correlación positiva fuerte entre los resultados del modelo propuesto y los datos obtenidos por PSNR y SSIM. Lo cual indica que los datos que entrega el modelo propuesto covaría un 74.9% con respecto a los resultados que se obtuvo con PSNR y con respecto a SSIM los resultados covarían en un 57.2%.

En el caso del resultado obtenido con PSNR se observa una correlación positiva fuerte debido a las distorsiones causadas al momento de transmitir el video y esto se relaciona directamente con la manera de medir de PSNR que compara pixel a pixel.

Tabla 28. Resultados estadísticos para comparar el desempeño del modelo video movimiento alto para el video Baloncesto.

Modelo	Media	Desv. Estándar	Coef. de Variación	Coef. Correl. Pearson - PSNR	Coef. Correl. Pearson - SSIM
MVMA	3.333	0.617	18.52	0.749	0,572
PSNR	2,925	0.527	18.00		
SSIM	2.9905	0.3392	11.34		

Es claro anotar, que el tratamiento cero no se tuvo en cuenta para estos análisis por ser un dato atípico, como se explicó anteriormente.

Al final del capítulo se presentarán de manera resumida los resultados obtenidos en esta sección y se profundizará más en los análisis y las conclusiones de los datos obtenidos.

## 5.3 Comparación de desempeño entre modelos propuestos vs pruebas subjetivas

En esta sección se compararán los resultados de los tres modelos propuestos con los resultados obtenidos en una serie de pruebas que se implementaron para medir subjetivamente los 16 tratamientos de cada video. El proceso de ejecución de estas pruebas subjetivas se encuentra descrito en el Anexo 1, y en esta sección solo se realizará las comparaciones necesarias para verificar el desempeño de los modelos propuestos.

### 5.3.1 Pruebas para comparar el desempeño de los modelos propuestos

Para verificar el desempeño de los modelos para evaluar QoE, los resultados obtenidos por estos deben ser correlacionados con resultados obtenidos mediante pruebas subjetivas. Para esto, el Grupo de Expertos de Calidad del Video (*Video Quality Expert Group –VQEG*) ha planteado un plan de pruebas para validar el desempeño de los diferentes modelos propuestos por diferentes autores [92].

La relación entre los resultados obtenidos por los modelos propuestos y los resultados obtenidos a partir de las pruebas subjetivas se realiza mediante una regresión no lineal, permitiendo evaluar el desempeño del modelo propuesto teniendo en cuenta tres aspectos: (i) Precisión de Predicción (*Prediction Accuracy*), (ii) Predicción de Monotocidad (*Prediction Monotonicity*) y (iii) Predicción de Consistencia (*Prediction Consistency*) [15]. Cada aspecto se explicará a continuación y el proceso que se realiza en cada aspecto consiste en comparar los resultados obtenidos por el modelo propuesto con datos obtenidos mediante pruebas subjetivas.

- **Precisión de Predicción:** Se representa por el coeficiente de correlación de Pearson, explicado en la sección 5.1, ( Figura 29 página 91), y por la raíz del error cuadrático medio (*Root Mean Square Error - RMSE*), donde  $x_i = 1, \dots, N$  y  $\bar{x}$  es la media del conjunto de datos y está definida por la Ecuación 20 :

$$RMSE = \sqrt{\frac{1}{N} \sum (x_i - \bar{x})^2} \quad \text{Ecuación 20}$$

- **Predicción de Monotocidad:** Se representa por el coeficiente de correlación de Spearman (también conocido como Rho de Spearman) que evalúa la

relación monótona entre dos variables. En una relación monótona, las variables tienden a cambiar de forma conjunta, pero no necesariamente de manera constante, lineal o logarítmica, etc. El coeficiente de Spearman se calcula basándose en parejas de datos  $X, Y$  y el rango de su valor está entre  $[-1, 1]$ , donde  $-1$  significa que  $X$  puede ser representado como una función monótona decreciente de  $Y$  y  $1$  significa que  $X$  puede ser presentado como una función creciente monótona de  $Y$  [15].

- **Predicción de Consistencia:** Es representado por el porcentaje de valores atípicos (*Outlier Ratio* - *OR*) que está definido como el porcentaje del número de predicciones por fuera del rango de  $\pm 2$  desviaciones estándar de los resultados de las pruebas subjetivas [69]. Si  $N$  es el número total de datos y  $N'$  es el número de valores atípicos, el porcentaje de valores atípicos está dado por:  $OR = N'/N$ .

Basándonos en estos tres aspectos, en la siguiente sección se presentarán los resultados de desempeño obtenidos para los tres modelos propuestos.

### 5.3.2 Proceso de comparación de desempeño: modelos propuestos vs. pruebas subjetivas

En esta sección se presentarán los resultados de las pruebas de desempeño realizadas a cada modelo propuesto. Para cada modelo se calcula lo siguiente:

- Coeficiente de Correlación de Pearson (CCP)
- Raíz del Error Cuadrático Medio (RECM)
- Coeficiente de correlación de Spearman (CCS)
- Porcentaje de Valores atípicos (PVA)

#### 5.3.2.1 Modelo video movimiento bajo para el video entrevista

En la Tabla 29, se presentan los resultados obtenidos en escala MOS del MVMB y los resultados de los 16 tratamientos evaluados mediante pruebas subjetivas. Se puede apreciar que para el tratamiento cero, el MVMB-video entrevista entrega un valor de MOS de 5,0 y las pruebas subjetivas entregaron el valor más alto (4,5) y solo en el tratamiento 14 se puede observar una calificación de 3,0 por parte del modelo propuesto en comparación a la obtenida con las pruebas subjetivas que fue de 3,044. Se puede deducir que los valores obtenidos por el modelo propuesto están relacionados con la escala de mapeo diseñada, que para este caso no fue ajustada, ni se diseñó buscando que los resultados de las pruebas y coeficientes calculados fueran más altos.

Tabla 29. Resultados en escala de MOS del modelo propuesto y las pruebas subjetivas realizadas a los 16 tratamientos del video Entrevista.

No. Tratamiento	MOS MVMB Video Entrevista	MOS Pruebas Subjetivas
0	5,0	4,500
1	5,0	4,061
2	5,0	4,096
3	4,0	3,474
4	5,0	3,632
5	4,0	4,228
6	4,0	2,667
7	4,0	3,825
8	4,0	3,491
9	5,0	4,333
10	4,0	3,509
11	5,0	4,237
12	5,0	4,342
13	5,0	3,693
14	3,0	3,044
15	4,0	4,211

Y en la Tabla 30 se presentan los datos estadísticos del modelo propuesto y los resultados de las pruebas subjetivas. Se puede apreciar que los valores obtenidos de la desviación estándar indican que los resultados no están tan dispersos de la media y esto se puede corroborar observando la Figura 30.

Tabla 30. Datos estadísticos de los resultados del MVMB – video entrevista y los resultados obtenidos mediante las pruebas subjetivas.

Modelo	Media	Desv. Estándar	Coef. de Variación
MVMB - video Entrevista	4,438	0,629	14,18
Pruebas Subjetivas	3,834	0,512	13,35

En la Figura 30 se aprecia los resultados que arroja el MVMB cuando se evalúan los valores de los parámetros de QoS para los 16 tratamientos (Figura 30 (a)) y los resultados de las pruebas subjetivas para cada tratamiento (Figura 30 (b)). Las líneas rojas horizontales en la figura (a) indican el límite de dos desviaciones estándar por encima y por debajo del promedio de los datos obtenidos mediante las pruebas subjetivas. Con esos límites se calcula el porcentaje de valores atípicos (PVA). Y en la figura (b) la línea horizontal roja presenta la media de los datos.

Aunque los coeficientes de variación del modelo propuesto y las pruebas subjetivas sean de valor parecido, la variabilidad de las respuestas en algunos casos es diferente como se puede observar en los tratamientos 3, 6 y 14.

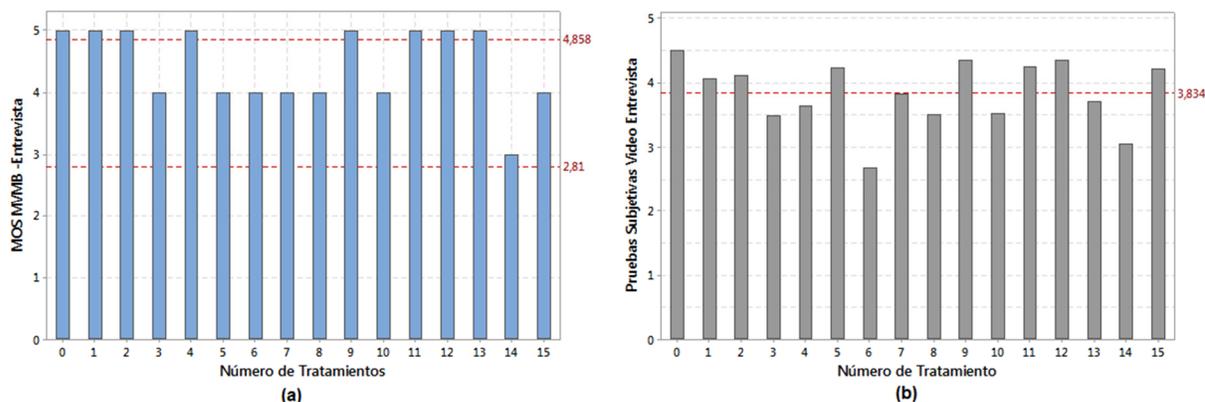


Figura 30. (a) Medidas en escala MOS obtenidas al evaluar los parámetros de QoS del Modelo Entrevista. (b) Medidas obtenidas al realizar las pruebas subjetivas a los 16 tratamientos del video Entrevista.

En la Tabla 31 se presentan los resultados de las pruebas de desempeño del modelo Entrevista. Se puede apreciar que la fidelidad del modelo Entrevista es moderada por los valores obtenidos en el CCP y CCS, el PVA es del 50%. Esto indica que el modelo propuesto explica en 62.4% de la variabilidad de los resultados obtenidos con las pruebas subjetivas. Aunque el PVA está muy elevado (50%), es decir, la mitad de las respuestas están por fuera de dos desviaciones estándar de la media de las pruebas subjetivas, esto puede ser debido a la escala de mapeo diseñada.

En la siguiente sección se presentará un resumen de los datos obtenidos en las pruebas de desempeño del MVMB.

Tabla 31. Resultados de pruebas de desempeño del MVMB.

Aspecto a evaluar	MVMB Video Entrevista
Coeficiente de Correlación de Pearson (CCP)	0,624
Coef. de correlación de Spearman (CCS)	0,630
Raíz del Error Cuadrático Medio (RECM)	0,607
Porcentaje de Valores atípicos (PVA)	0,5

### 5.3.2.2 Modelo video movimiento moderado para el video cafetería

En la Tabla 32, se presentan los resultados en escala MOS del MVMM para el video Cafetería y los resultados de los 16 tratamientos evaluados mediante pruebas

subjetivas. Se puede apreciar con respecto a las pruebas subjetivas que para este tipo de video se obtienen calificaciones que empiezan en 2,1 y llegan hasta 4,2. Esto permite apreciar que las personas detectaron distorsiones del video al momento de la evaluación. Con respecto a los resultados obtenidos con el modelo solo se obtuvo una calificación de 3,0 y esto se debe al diseño de la escala que mapea VQM a MOS, donde cerca del 50% de la escala correspondía a la calificación de 4,0 en MOS.

Tabla 32. Resultados en escala de MOS del modelo propuesto y las pruebas subjetivas realizadas a los 16 tratamientos del video Cafetería.

No. Tratamiento	MOS MVMM Video Cafetería	MOS Pruebas Subjetivas
0	4,0	4,211
1	4,0	4,123
2	4,0	3,509
3	4,0	2,640
4	4,0	3,044
5	4,0	3,544
6	4,0	3,939
7	4,0	2,939
8	4,0	3,298
9	4,0	2,693
10	3,0	2,193
11	4,0	3,640
12	4,0	3,009
13	4,0	3,798
14	4,0	2,947
15	4,0	3,079

En la Tabla 33, se presentan los datos estadísticos del modelo propuesto y los resultados de las pruebas subjetivas. Se puede apreciar que el coeficiente de variación de los resultados del MVMM-video cafetería es bajo (6,35%) esto se debe a la escala diseñada para este modelo. Y de manera consecuente la desviación estándar es mucho más baja en comparación con la desviación estándar obtenida en las pruebas subjetivas.

Tabla 33. Datos estadísticos de los resultados del modelo video movimiento moderado para el video Cafetería y los resultados obtenidos mediante las pruebas subjetivas.

Modelo	Media	Desv. Estándar	Coef. de Variación
MVMM-video Cafetería	3,9375	0,25	6,35
Pruebas Subjetivas	3,288	0.569	17,32

En la Figura 31 se aprecia los resultados que arroja el modelo video movimiento moderado cuando se evalúan los valores de los parámetros de QoS para los 16 tratamientos (Figura 31 (a)) y los resultados de las pruebas subjetivas para cada tratamiento (Figura 31(b)). Las líneas rojas horizontales en la figura (a) indican el límite de dos desviaciones estándar por encima y por debajo del promedio de los datos obtenidos mediante las pruebas subjetivas. Con esos límites se calcula el porcentaje de valores atípicos (PVA). Y en la figura (b) la línea horizontal roja presenta la media de los datos.

Se puede apreciar variabilidad reducida de los resultados obtenidos por el MVMM-video cafetería, en comparación con los resultados de las pruebas subjetivas, como se mencionó anteriormente esto se debe al diseño de la escala para mapear y a que las personas para este tipo de video fueron más críticas o detectaron más distorsiones al momento de calificar el video.

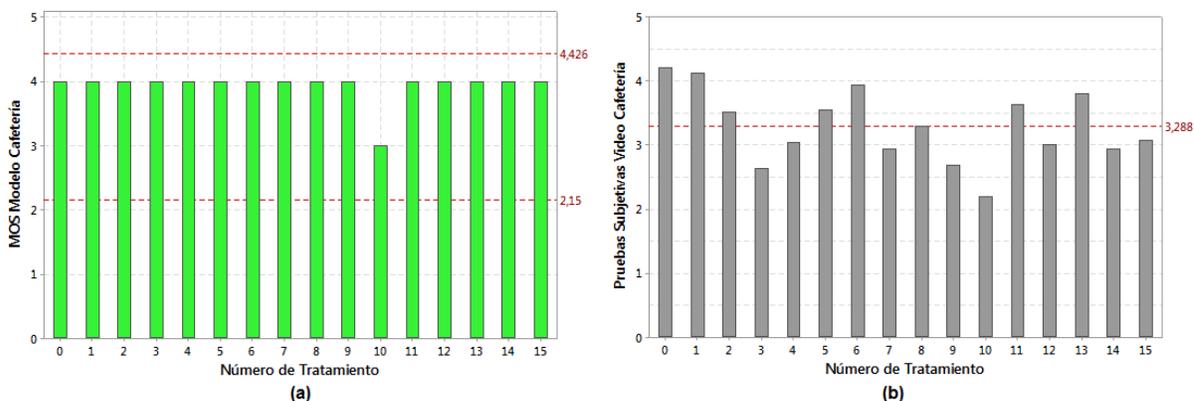


Figura 31. (a) Medidas en escala MOS obtenidas al evaluar los parámetros de QoS del MVMM. (b) Medidas obtenidas al realizar las pruebas subjetivas a los 16 tratamientos para el video cafetería.

En la Tabla 34 se presentan los resultados de las pruebas de desempeño del MVMM – video cafetería, se puede apreciar que el CCP es 0,513 lo cual indica una correlación positiva moderada. Y también se puede apreciar que el PVA es cero esto debido a la escala propuesta para mapear los resultados del modelo con la escala MOS. El análisis de fidelidad se completará en la siguiente sección. El RECM es baja en comparación con el MVMB-video entrevista, esto debido a la escala de mapeo

propuesta, que para este modelo hace que las respuestas que entrega el modelo tenga un coeficiente de variación bajo.

Tabla 34. Resultados de pruebas de desempeño del MVMM –video cafetería.

Aspecto a evaluar	MVMM Video Cafetería
Coeficiente de Correlación de Pearson (CCP)	0,513
Coef. de correlación de Spearman (CCS)	0,420
Raíz del Error Cuadrático Medio (RECM)	0,24
Porcentaje de Valores atípicos (PVA)	0,0

### 5.3.2.3 Modelo video movimiento alto para el video baloncesto

En la Tabla 35 se presentan los resultados obtenidos en escala MOS del MVMA y los resultados de los 15 tratamientos evaluados mediante pruebas subjetivas. Donde se puede apreciar que el valor más alto de calificación obtenido en ambas pruebas es en el tratamiento 1. Para este video se pueden apreciar un 73% de las calificaciones entre 3,0 y 3,9, esto indica que las personas al evaluar el video detectaban con mayor facilidad las distorsiones. El modelo propuesto entrega una calificación de 5,0 en el tratamiento 1 y 13, donde la pérdida de paquetes es cero. Esto corrobora que para este tipo de videos de movimiento alto la pérdida de paquetes es un parámetro que afecta claramente la calidad de la experiencia asociada al video.

Tabla 35. Resultados en escala de MOS del modelo propuesto y las pruebas subjetivas realizadas a los 16 tratamientos del video Baloncesto.

No. Tratamiento	MOS Modelo Propuesto Baloncesto	MOS Pruebas Subjetivas
1	5,0	4,896
2	4,0	4,430
3	4,0	3,070
4	4,0	3,123
5	4,0	3,044
6	4,0	3,229
7	4,0	2,798
8	4,0	3,254
9	4,0	3,553
10	4,0	3,254
11	4,0	3,228
12	4,0	3,211
13	5,0	4,746

No. Tratamiento	MOS Modelo Propuesto Baloncesto	MOS Pruebas Subjetivas
14	4,0	3,307
15	4,0	3,474

Y en la Tabla 36 se presentan los datos estadísticos del modelo propuesto y los resultados de las pruebas subjetivas. Se puede apreciar que los resultados obtenidos para el MVMA-video baloncesto la desviación estándar es más baja que la obtenida con las pruebas subjetivas, y esto va relacionado con el coeficiente de variación obtenido por los datos arrojados por el modelo (8,51%).

Tabla 36. Datos estadísticos de los resultados del MVMA y los resultados obtenidos mediante las pruebas subjetivas.

Modelo	Media	Desv. Estándar	Coef. de Variación
MVMA- video Baloncesto	4,133	0,351	8,51
Pruebas Subjetivas	3,508	0,643	18,33

En la Figura 32 se aprecia los resultados que arroja el MVMA cuando se evalúan los valores de los parámetros de QoS para los 16 tratamientos (Figura 32 (a)) y los resultados de las pruebas subjetivas para cada tratamiento (Figura 32 (b)). Las líneas rojas horizontales en la figura (a) indican el límite de dos desviaciones estándar por encima y por debajo del promedio de los datos obtenidos mediante las pruebas subjetivas. Con esos límites se calcula el porcentaje de valores atípicos (PVA). Y en la figura (b) la línea horizontal roja presenta la media de los datos. Se puede apreciar la variabilidad baja en las respuestas que entrega el MVMA-video baloncesto asociado al coeficiente de variación obtenido y ligado con el diseño de la escala de mapeo entre VQM y MOS.

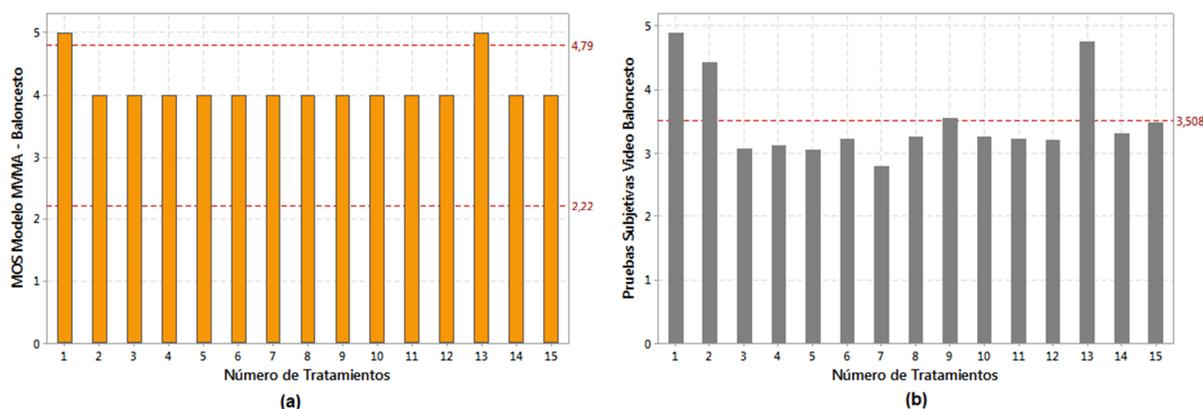


Figura 32. (a) Medidas en escala MOS obtenidas al evaluar los parámetros de QoS del MVMA. (b) Medidas obtenidas al realizar las pruebas subjetivas a los 16 tratamientos del video Baloncesto.

En la Tabla 37 se presentan los resultados de las pruebas de desempeño del MVMA. Se puede apreciar que CCP y CCS entregan una correlación positiva moderada, siendo más alta la correlación obtenida con el CCP que llega al 82,9%, el PVA es del 13% lo cual es bajo y con el CCP obtenido, indica que el modelo posee una fidelidad alta. En la siguiente sección se presentará de manera conjunta las pruebas con PSNR y SSIM que se le realizaron al modelo.

Tabla 37. Resultados de pruebas de desempeño del modelo Baloncesto.

Aspecto a evaluar	Modelo Baloncesto
Coeficiente de Correlación de Pearson (CCP)	0,829
Coef. de correlación de Spearman (CCS)	0,591
Raíz del Error Cuadrático Medio (RECM)	0,3415
Porcentaje de Valores atípicos (PVA)	0,13

## 5.4 Pruebas estáticas de cada modelo

En esta sección se presentarán los resultados de pruebas estáticas realizadas a cada modelo con el fin de analizar su comportamiento ante diferente cantidad de tratamientos. Las pruebas se realizaron generando valores aleatorios para dos parámetros de QoS y dejando el otro fijo para 1000 tratamientos.

Para el MVMB -video entrevista- se dejó constante el retardo a un valor de 1300ms y se variaron de manera aleatoria pérdida de paquetes y la variación del retardo. Se puede apreciar en la Figura 33 que la pérdida de paquetes influye en la QoE que entrega el modelo, debido a que para valores inferiores a 0.1% se puede obtener una QoE superior a 4 en escala MOS. Aunque se puede apreciar que para valores inferiores a 0.005 ms en la variación del retardo la QoE cae drásticamente a 2, en escala MOS, esto se debe a la influencia de la pérdida de paquetes en el modelo propuesto. Adicionalmente debido a que el video es estático, la QoE se ve afectada porque el usuario puede percibir la degradación del video en partes específicas del mismo.

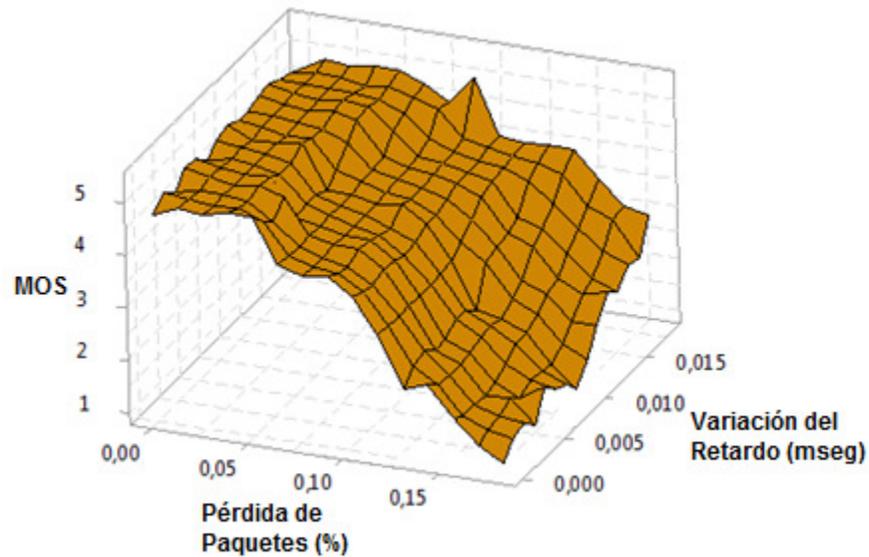


Figura 33. MOS vs Pérdida de paquetes y variación del retardo para el MVMB para el video entrevista.

Para el MVMM la variación del retardo se dejó configurado en 0.018 ms, Retardo y Pérdida de paquetes se variaron de manera aleatoria. La Figura 34 muestra que para valores de pérdida de paquetes cercanos a 0.18 y retardo superior a 1100 ms la QoE del video baja a niveles de 1 a 2 en la escala MOS. Mientras que para valores inferiores a 0.1% en pérdida de paquetes y retardo inferior a 600ms, la QoE del video se puede mantener en valores de 4 (Buena) a 5(Excelente) en escala MOS. Este resultado es importante para el proveedor de servicio porque le permitirá ajustar los parámetros de QoS entre los rangos descritos anteriormente a fin de mantener la QoE del video en un nivel superior a 4, en escala MOS.

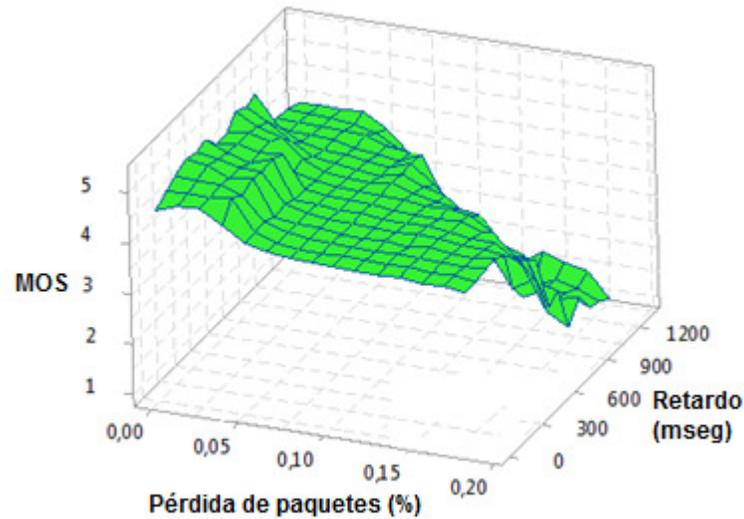


Figura 34. MOS vs Pérdida de Paquetes y retardo, para el MVMM para el video cafetería.

Para el MVMA –video baloncesto- se dejó constante el retardo a un valor de 400ms (según la Rec. 1541 [59]) y los otros dos parámetros se variaron de manera aleatoria. Se puede apreciar en la Figura 35 que la pérdida de paquetes influye en el modelo y a partir de valores superiores a 0.1%, el MOS obtenido empieza a descender a un valor de 3. Para este tipo de video la pérdida de paquetes influyen notoriamente en la QoE del video debido al alto movimiento. Es por esta razón que solo para valores muy bajos de pérdida de paquetes (inferiores a 0.05%) se logran obtener medidas de MOS superiores a 4.

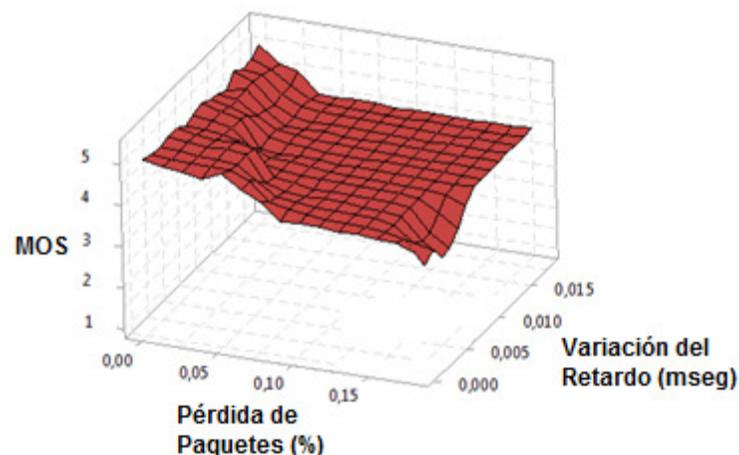


Figura 35. MOS vs Pérdida de paquetes y variación del retardo para el MVMA para el video baloncesto.

## 5.5 Comentarios del capítulo

En esta sección se presentará el resumen de las pruebas realizadas a los tres modelos y se analizarán los datos obtenidos.

En la Tabla 38 se presenta un resumen de las pruebas realizadas, donde cada modelo propuesto está en una columna identificado por su nombre y video asociado. En las filas se presenta el coeficiente de variación del modelo propuesto, coeficiente de correlación de Pearson (CCP) calculado entre cada modelo y los resultados de las mediciones realizadas con PSNR y SSIM. Y se presentan los resultados de las pruebas de desempeño de cada modelo basándose en los resultados obtenidos en las pruebas subjetivas.

Tabla 38. Resumen de pruebas de desempeño de los tres modelos propuestos.

<b>Parámetro/Modelo</b>	<b>MVMB Entrevista</b>	<b>MVMM Cafetería</b>	<b>MVMA Baloncesto</b>
Coef. de Variación Modelo Propuesto	14,18	6,35	8,51
CCP -PSNR	0,624	0,623	0,802
CCP-SSIM	0,607	0,485	0,662
<b>Pruebas Subjetivas</b>			
Coefficiente de Correlación de Pearson (CCP)	0,624	0,513	0,829
Coef. de correlación de Spearman (CCS)	0,630	0,420	0,591
Raíz del Error Cuadrático Medio (RECM)	0,607	0,24	0,341
Porcentaje de Valores atípicos (PVA)	0,5	0	0,13

Analizando los resultados obtenidos de los tres modelos con respecto a CCP-PSNR y CCP-SSIM, se puede decir que los modelos tienen una correlación positiva moderada donde el modelo que mejor explica la variabilidad de los datos es el modelo Baloncesto y el modelo con una correlación positiva débil es el asociado al video cafetería.

Pero al analizar los resultados de desempeño con respecto a las pruebas subjetivas se encuentra lo siguiente:

- El modelo que mejor explica la variabilidad de los datos es el MVMA para el video baloncesto, con un CCP de 0,829, es decir, el modelo explica la variabilidad de los datos en un 82,9% y con un PVA bajo de un 13%.
- El modelo con menor fidelidad es el MVMM –video cafetería con un CCP de 0,513 en las pruebas subjetivas, lo cual indica una correlación positiva moderada de tan solo un 51,7%. Esto se puede explicar debido a que en las pruebas subjetivas los sujetos evaluaban con más rigurosidad las distorsiones en el video, y al momento de comparar los resultados que entrega el modelo

con las pruebas subjetivas se puede notar las diferencias que se reflejan al calcular el CCP.

Aunque las pruebas subjetivas sean un estándar de facto para probar modelos propuestos, se debe analizar con mucho cuidado el proceso de su implementación como también el análisis de los resultados obtenidos. Ya que en algunos casos un diseño inadecuado de una prueba subjetiva puede hacer que se obtengan conclusiones erradas con respecto a la precisión y fidelidad de modelos propuestos.

Con respecto a las pruebas estáticas se puede apreciar que los resultados son consistentes con los análisis realizados en el capítulo 3 y 4, no se encontraron comportamientos atípicos en las respuestas que entrega cada modelo. Y se corroboró los resultados que entrega el modelo Baloncesto que entrega medidas máximas de 4,0 en las 1000 iteraciones que se realizaron.

Por otra parte, se puede apreciar que el parámetro que más influye en la QoE asociada al video es la pérdida de paquetes, este parámetro es crítico en el bucle del usuario debido a los anchos de banda que se manejan en los usuarios residenciales. Los modelos propuestos permitirán que los proveedores de servicio estén continuamente evaluando QoE asociada al video y tomar acciones correctivas con respecto a ofrecer mayor ancho de banda o mejorar los esquemas de calidad de servicio que ofrece al usuario.

Cómo el objetivo es que el modelo propuesto sea utilizado por los proveedores de servicio que ofrezcan el servicio de IPTV, a continuación se plantea el protocolo de pruebas para que el proveedor pueda utilizar este modelo propuesto en su plataforma de servicios:

1. Se deben utilizar los videos asociados al tratamiento cero I00, C00 y B00 para realizar las medidas de QoE en la red del proveedor.
2. En el Set-Top-Box (STB) se instala una aplicación que permita ejecutar los modelos asociados a cada tipo de video, y se verifica que se pueda obtener en el setup-box, la medición de los parámetros de calidad de servicio.
3. Se envían los videos I00, C00 y B00 por la red hasta el usuario que se desea realizar la medida de QoE cuando el proveedor lo estime conveniente.
4. Se calcula la QoE asociada a cada video y la información se envía al nodo central de la plataforma de gestión para su procesamiento.
5. Si se llegan a presentar inconvenientes en la instalación de la aplicación en el STB debido a que son arquitecturas cerradas y propietarias, una opción sería extraer los valores de los parámetros de calidad de servicio utilizando Simple Network Management Protocol (SNMP). Ya con los valores de los parámetros de calidad de servicio obtenidos, mediante una aplicación web en la nube, calcular el valor de calidad de experiencia asociado al video utilizando el modelo propuesto.

De esta manera los proveedores de servicio podrían hacer uso de esta primera versión del modelo propuesto en este trabajo de investigación.

Con el fin de dejar claro el proceso para obtener el modelo, se presenta el diagrama de flujo del proceso, esto permitirá a otros investigadores reproducir el proceso, utilizando otro tipo de parámetros de calidad de servicio. El diagrama de flujo se presenta en la figura Figura 36 y la explicación se presenta a continuación. Adicionalmente en la Figura 37 se presenta el algoritmo para obtener el modelo.

**Paso 1:** En este paso, los investigadores deben escoger si grabaran sus videos con recursos propios (Paso 1A), o seleccionarán videos de una base de datos pública (Paso 2).

**Paso 3:** Seleccionar el diseño del experimento. Este paso es fundamental porque permitirá definir el tipo de modelo a obtener, número de tratamientos a ejecutar y/o rangos de las variables a usar.

**Paso 4:** Seleccionar el número y tipo de parámetros de calidad de servicio a incluir en el modelo. Otros tipos de parámetros de red también pueden ser usados.

**Paso 5:** Distorsión de los videos dependiendo de los recursos disponibles. Si se poseen recursos de hardware y software, los investigadores pueden implementar una maqueta de pruebas (Paso 5A) con el fin de distorsionar el video de diferente manera. Si solo se poseen recursos de software las distorsiones a los videos se deben realizar con una herramienta de simulación (Paso 6).

**Paso 7:** Seleccionar el método para evaluar la QoE asociada al video. Método subjetivo o métodos objetivos pueden ser elegidos.

**Paso 8:** Ejecutar el método elegido en el paso anterior y empezar la recolección de datos. El número de tratamientos fueron definidos en el paso 3.

**Paso 9:** Realizar el análisis de datos utilizando herramientas estadísticas.

**Paso 10:** Modelo Preliminar obtenido. A este modelo se le deben realizar pruebas de desempeño.

**Paso 11:** Elegir el tipo de evaluación de desempeño a realizar al modelo propuesto. Pruebas subjetivas (Paso 11A) o comparación con métodos objetivos (Paso 11B).

**Paso 12:** Umbral de desempeño. Factores de desempeño como CCP y CCS son analizados. Si se obtiene una métrica mayor o igual a 0.5, el investigador ha obtenido la versión final de su modelo (Paso 13). De otra manera, el procedimiento se debe ejecutar de nuevo desde el Paso 3.

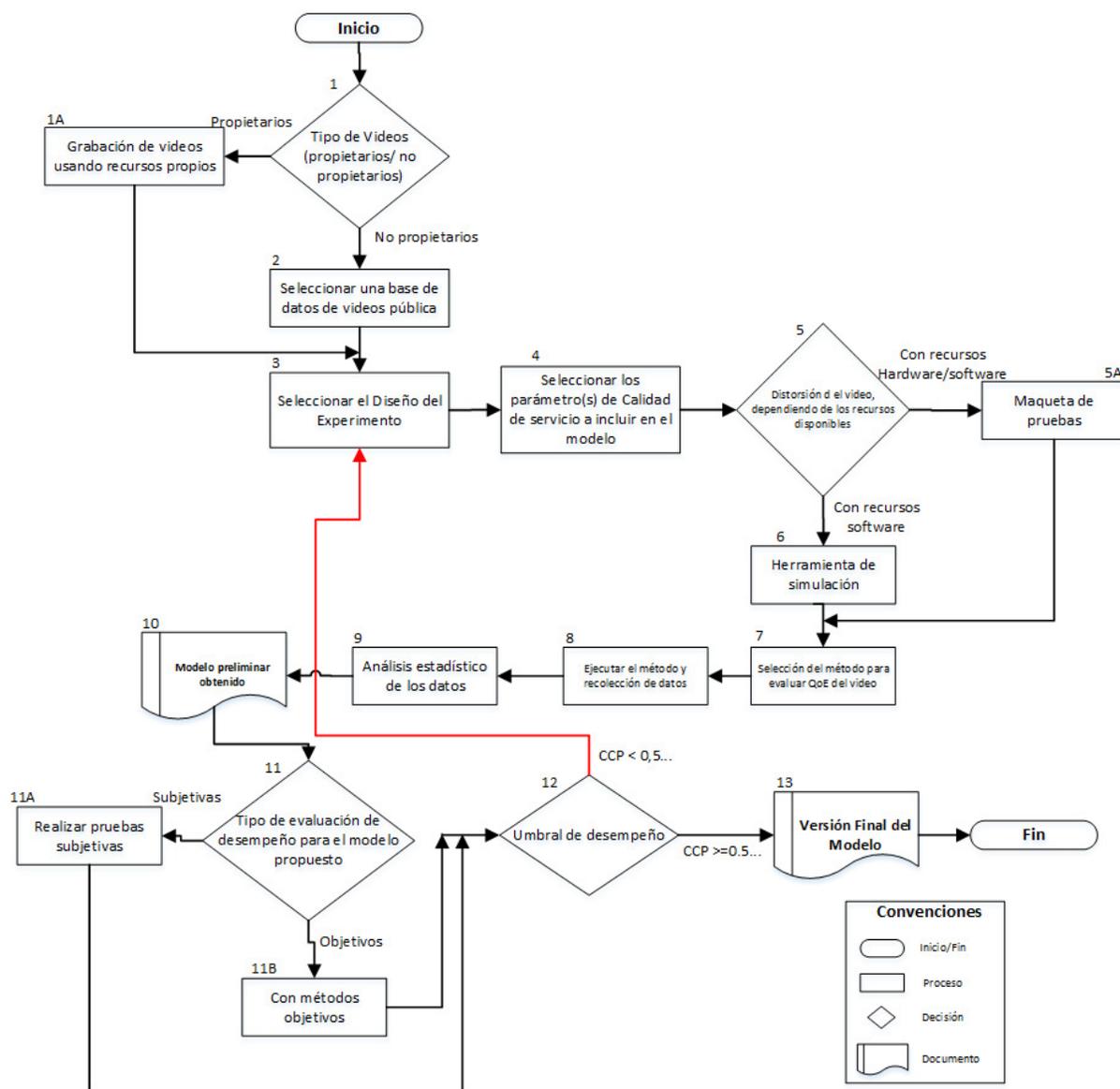


Figura 36. Diagrama de Flujo del algoritmo para obtener el modelo para medir QoE para el servicio de IPTV.

**Algoritmo 1.** Modelo para la medición de QoE para el servicio de IPTV

**Input:** Valor de parámetros de QoS y/u otro tipo de parámetro de red

**Output:** QoE asociada al video

```

1: Procedure (Seleccionar el tipo de videos)
  if usando videos propios then grabar los videos fuente usando recursos propios
  else if usando videos de una base de datos pública then seleccionar una base de datos de video publica
  end if
end procedure
2: Seleccionar el diseño del experimento
3: Seleccionar los parámetros de QoS y/o parámetros de red a incluir en el modelo
4: Procedure (Distorsión de los videos) /*Depende de los recursos disponibles*/
  If se poseen recursos de HW y SW disponibles
  then implementar una maqueta de pruebas
  else usar una herramienta de simulación
5: Seleccionar el método para evaluar la QoE asociada al video /*Método subjetivo o métodos objetivos pueden ser utilizados*/
6: Procedure (Ejecutar el método elegido en el paso anterior y empezar la recolección de datos)
  for i= 1 to N /*N: Número de tratamientos o ejecuciones que se deben realizar */
    for j= 1 to K /*K: número de parámetros de QoS o número de parámetros de red definidos en el paso 3*/
      X(j)= Valor del parámetro K de QoS;
    end for
    Y(N)= medida de QoE obtenida /* esta medida es obtenida con el método seleccionado en el paso 5*/
  end for
end procedure
7: Análisis estadístico de los datos|
8: Modelo preliminar obtenido
9: Tipo de evaluación de desempeño a realizar al modelo preliminar propuesto
  if prueba subjetivas then realizar pruebas subjetivas
  else realizar comparación con métodos objetivos
  end if
10: Procedure (Umbral de desempeño)
  if CCP >= 0.5
  then se obtiene la versión Final del modelo /* No solo se limita a obtener CCP, también se puede utilizar CCS, RECM y PVA */
  else goto paso 2:
  end if
end procedure
11: end

```

Figura 37. Algoritmo para obtener el modelo para medir QoE para el servicio de IPTV.

## 6 Conclusiones y recomendaciones

En este capítulo se presentan las conclusiones obtenidas durante la realización del proyecto y los productos obtenidos durante el desarrollo de la investigación. Al final se plantean los trabajos futuros y lecciones aprendidas. En la Figura 38 se presenta el contenido descriptivo de este capítulo.

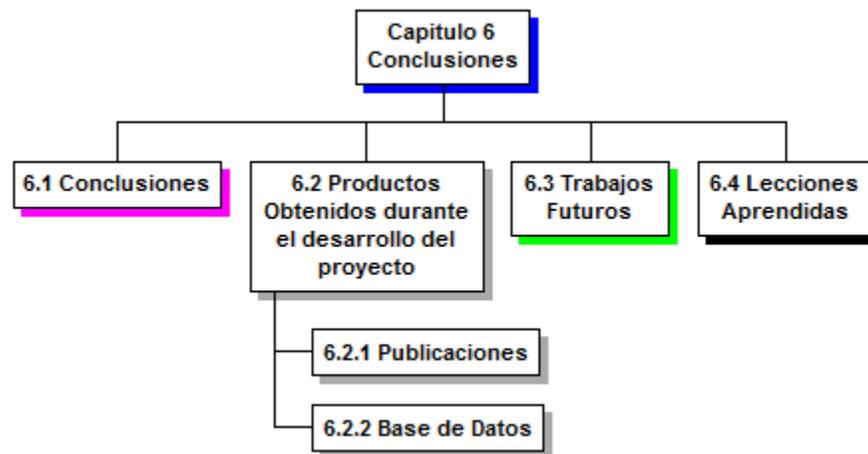


Figura 38. Descripción del contenido del capítulo 6.

### 6.1 Conclusiones

El modelo propuesto en este proyecto fue obtenido basándose en la distorsión realizada a videos al momento de ser transmitidos en un entorno de emulación. Este enfoque presentado en esta propuesta, tiene una serie de ventajas sobre propuestas de otros autores, ya que en dichas propuestas, la distorsión de los videos se realiza al archivo del video sin transmitirlo en un entorno de red. Es decir las distorsiones se hacen al archivo de manera estática y afectando características del archivo elegidas de manera específica.

De acuerdo a lo anterior, este el mayor aporte de este proyecto a la solución del problema, la transmisión en línea del video y su distorsión de manera aleatoria. El enfoque planteado en este trabajo de investigación es novedoso y difiere de otras propuestas que basan sus modelos en pruebas subjetivas y utilizan un número inferior de parámetros de calidad de servicio.

El modelo propuesto puede ser utilizado en un ambiente real de red o lo puede utilizar un proveedor de servicios para estar midiendo de manera continua y **en línea** la calidad de la experiencia del servicio que está prestando. Esto es de suma importancia para el proveedor de servicio ya que puede estar ajustando los procesos de calidad de servicio al interior de sus redes de núcleo.

Ahora bien, para la implementación de este modelo en un proveedor de servicio se recomienda la opción de instalar en los STB del proveedor una aplicación que permita ejecutar el modelo propuesto y obtener las mediciones de QoE asociadas al video. Si esto no es posible, otra opción, es valiéndose del SNMP o una aplicación de gestión propietaria, obtener los valores de los parámetros de QoS y mediante una aplicación web realizar los cálculos de QoE y almacenarlos en la nube para su posterior análisis.

El algoritmo VQM se adapta a los requerimientos definidos para generar el modelo propuesto, debido a que es de los modelos objetivos más robusto y que presenta mejor correlación en pruebas realizadas en diferentes publicaciones. Adicionalmente VQM es usado por un ente estandarizador en mediciones de calidad de video.

Los rangos de los parámetros de QoS se definieron de manera empírica mediante la ejecución del diseño del experimento en repetidas ocasiones y realizando un análisis estadístico de los datos obtenidos para verificar su validez. Adicionalmente un grupo conformado por tres expertos evaluaban las distorsiones causadas en los videos, verificando que estas distorsiones fueran perceptibles para el usuario y que los videos se pudieran decodificar en el lado del cliente.

Para mejorar la fidelidad de los modelos se hace necesario mejorar y automatizar la maqueta de pruebas en dos aspectos: (1) La elección de un número mayor de videos con diferentes características espacio-temporales y en formato YUV. (2) En la maqueta de pruebas utilizar una herramienta de simulación. Estos dos aspectos permitirán realizar mayor número de pruebas con el fin de analizar los datos obtenidos y comparar los resultados obtenidos con esta primera versión del modelo propuesto. Adicionalmente se podrá obtener un modelo genérico que permita medir cualquier tipo de video.

Las pruebas subjetivas se hacen necesarias para verificar la fidelidad y desempeño de nuevos modelos propuestos, pero se debe planear muy bien la ejecución de estas pruebas siguiendo las recomendaciones indicadas por la ITU. Los resultados de las pruebas subjetivas permiten a su vez, la creación de bases de datos de video para futuros estudios o para probar nuevas propuestas. Estas bases de datos pueden ser muy útiles para la comunidad investigativa y para los proveedores de servicio.

Con base en los análisis de varianza de los datos asociados a cada tipo de video se puede apreciar que el parámetro de QoS que influye de manera notoria en la QoE asociada al video, es la pérdida de paquetes. Y la calidad del video se degrada más dependiendo del tipo de cuadro que se pierda, siendo el caso más crítico cuando se pierden marcos tipo I. Los parámetros de QoS como retardo y variación del retardo,

que afectan la medición de la QoE asociada al video se pueden contrarrestar mediante la configuración del tamaño de la memoria intermedia en el cliente.

En cuanto a la fidelidad del modelo, el MVMA-video baloncesto es el que presentó un coeficiente de correlación de Pearson más alto de los tres modelos (82,9%) cuando se comparó con las pruebas subjetivas. Esto debido a que al momento de las personas evaluar la calidad del video, en las pruebas subjetivas, detectaba con mayor facilidad las distorsiones del video. Caso contrario ocurrió con el MVMM – video cafetería donde debido a la complejidad de la imagen para el usuario era más complicado detectar las distorsiones del video debido a que se enfocaba en analizar solo parte de la imagen.

## 6.2 Productos obtenidos durante el desarrollo del proyecto

### 6.2.1 Publicaciones

En esta sección se presentan las publicaciones y asistencia a eventos producto de este trabajo de investigación.

Artículos publicados:

- **J. C. Cuéllar**, Jesus Hamilton Ortiz, and J. L. Arciniegas, “**Clasificación y Análisis de Métodos para medir Calidad de la Experiencia del Servicio de Televisión sobre Protocolo IP (IPTV)**,” *Información Tecnológica*, Chile. vol. 25, no. 5, pp. 121–128, Jan. 2014. [9]

Durante la pasantía realizada en la Universidad Politécnica de Madrid se trabajó de manera conjunta con un grupo de profesores y se logró la publicación del artículo:

- **QoE-Based Analysis of DASH Streaming Parameters over Mobile Broadcast Networks**. IEEE Access. Vol. 5. Págs. 20684-20694. ISSN 2169-3536. 21 de septiembre de 2017. doi: [10.1109/ACCESS.2017.2755438](https://doi.org/10.1109/ACCESS.2017.2755438).

**Autores:** Carlos M. Lentisco, Luis Bellido, **Juan Carlos Cuellar Q.**, Encarna Pastor and Jose L. Arciniegas.

#### **Abstract**

Multimedia services over mobile networks are subject to multiple factors that affect the quality of experience (QoE) perceived by users. In a long term evolution (LTE) multimedia broadcast service, these factors are initial delay, technical quality of multimedia content, and service disruptions. However, there is a lack of QoE models to enable QoE measurement. We propose to use video

mean opinion score as a new and objective QoE model that is in the process of standardization, and we show how it can be applied to LTE multimedia broadcast services for QoE assessment. The results obtained are useful to better understand the impact of several service parameters on perceived quality. By applying this assessment methodology, it is possible to determine the most adequate service parameters for a speci\_c LTE broadcast deployment.

Eventos en que se participó:

- D. Acosta, **J. C. Cuellar**, and J. L. Arciniegas, “**Guidelines to Implement Video Dataset to Assess Video Quality**,” *IEEE Colombian Conference on Communications and Computing. 2018*, May 16th-18th. Medellín, Colombia, ISBN: 978-1-5386-6820-7.
- **Spring School on Networks**  
Santiago de Chile Noviembre 9 y 10 de 2015.  
**Poster: Propuesta Metodológica para Implementar un Modelo para Medir Calidad de la Experiencia para el Servicio de IPTV.** J.C. Cuéllar, Jesus Hamilton Ortiz and J.L. Arciniegas
- **Hybrid Model to Measure QoE in IPTV: Methodological Proposal**  
J.C. Cuéllar, Jesus Hamilton Ortiz and J.L. Arciniegas  
**VII Congreso Iberoamericano de Telemática CITA 2015.**  
ISBN-13: 978-0-9898393-8-9. Popayán Colombia.
- **Congreso Nacional:**  
**J.C. Cuéllar**, Jesus Hamilton Ortiz y J.L. Arciniegas. **Modelo Híbrido para Medir QoE al Servicio de IPTV.** Octavo Congreso Colombiano de Computación. Universidad del Quindío. Agosto 21 al 23 de 2013.

## 6.2.2 Base de Datos

Con los videos utilizados en las pruebas subjetivas, se diseñó una base de datos que está compuesta por 64 videos: Con 4 secuencias originales y 60 videos con distorsiones. El procedimiento como se llevaron a cabo estas pruebas subjetivas se encuentra detallado en el Anexo 2.

Los videos utilizados y las calificaciones de los observadores pueden ser obtenidas en el repositorio: <ftp://exanime.icesi.edu.co>. Para acceder a ella se debe solicitar la contraseña a [jcuellar@icesi.edu.co](mailto:jcuellar@icesi.edu.co).

Esta base de datos es la única en su tipo, ya que las diferentes bases de datos que se encuentran no tienen en cuenta distorsiones causadas por la variación de parámetros de calidad de servicio y videos transmitidos en un ambiente de red emulado.

## 6.3 Trabajos Futuros

Los trabajos futuros producto de este proyecto se pueden enfocar en varias líneas:

1. Grabar videos de pruebas con cámaras profesionales con el objetivo de tener mayor control al momento de transcodificar cada video o lograr tener los videos en formato YUV. Esta transcodificación es una variable no controlada que influye en el resultado de los datos obtenidos para generar los modelos.
2. Implementar una maqueta de experimentación más robusta, que permita realizar pruebas de manera automática y con mayor precisión que las realizadas durante el desarrollo del proyecto. Esto es importante porque el tiempo necesario para hacer las pruebas sería mucho menor al tiempo que se invirtió en este proyecto. Al mejorar las condiciones de la maqueta de experimentación se puede mejorar la fidelidad de los modelos propuestos y lograr obtener un modelo genérico para ser utilizado en videos con diferentes características espacio-temporales. La implementación de una maqueta de experimentación más robusta, se hace referencia a dos aspectos puntualmente que ayudarán a obtener un modelo genérico: (1) utilizar un número mayor de videos con diferentes características espacio-temporales para ser transmitidos en la maqueta de experimentación y (2) al implementar una maqueta donde las pruebas se puedan realizar de manera automática permitirá aumentar el número de pruebas con diferentes tipos de videos con el fin de lograr obtener un modelo genérico que permita analizar cualquier tipo de video.
3. Implementar un protocolo de pruebas a fin de definir con mayor precisión los rangos de las características espacio-temporales para la utilización de cada modelo al momento de evaluar la QoE. Este trabajo podría permitir ampliar el rango de utilización de los modelos propuestos.
4. Realizar pruebas de validación de desempeño del modelo propuesto vs. los modelos presentados en la sección 2.3 de trabajos relacionados, con el fin de analizar la fidelidad y desempeño del modelo propuesto. Durante la parte final del proyecto dichas pruebas de desempeño no se realizaron, debido a que todo el trabajo se concentró en realizar la validación de fidelidad del modelo, utilizando el protocolo de pruebas propuesto por el VQEG y las pruebas de correlación con SSIM y PSNR.  
Para realizar este tipo de pruebas de desempeño, se debe implementar una maqueta de pruebas o mediante el uso de herramientas de simulación. Los resultados obtenidos serán empleados para una nueva publicación que ya está en curso.
5. Seguir alimentando la base de datos diseñada con nuevos videos con distorsiones no solamente causadas por la modificación de parámetros de calidad de servicio. La actualización de la base de datos es fundamental debido

a que algunas de las bases de datos encontradas, ya están desactualizadas o la información técnica que entregan es escasa.

## **6.4 Lecciones aprendidas**

Uno de los aspectos fundamentales en el desarrollo de este proyecto estuvo relacionado con el diseño del experimento para obtener los datos para la generación de los modelos propuestos. Durante el desarrollo del proyecto, y al momento de analizar los datos obtenidos tanto en los videos transmitidos en la maqueta de experimentación, como las pruebas subjetivas realizadas, se observó que el diseño del experimento cobra vital importancia.

Es por esta razón que basándonos en la experiencia adquirida de este proyecto, el diseño de experimento se debe planificar cuidadosamente antes de empezar la toma de datos. Esto con el fin resolver inconvenientes durante el desarrollo del proyecto y tomar acciones correctivas en el momento adecuado.

Otro aspecto a tener en cuenta es la transcodificación que se lleva a cabo en el proceso de transmisión de los videos, al transcodificar los videos van perdiendo resolución, nitidez o calidad, haciendo que las medidas de QoE varíen y por ende las medidas subjetivas realizadas a los videos se puedan ver afectadas. Por ejemplo un video que no haya sufrido distorsiones al momento de su transmisión, puede ser evaluado negativamente por un usuario, como ocurrió en el video Baloncesto donde el video del tratamiento No.0 se tuvo que descartar de los análisis.

Es por esta razón que se deben tener videos de excelente calidad para minimizar el deterioro de los mismos al momento de transcodificarlos para transmitirlos o cambiarlos de formato para poderlos editar.

Lo anterior, va ligado con la cámara utilizada para grabar los videos de prueba, si se logra utilizar una cámara profesional este tipo de situación se puede mejorar, ya que se podrían tener los videos en formato YUV y así mejorar la calidad de los videos.

## Anexo 1: Configuración VLC Y NetEM

En este anexo se describirá el proceso de configuración de las herramientas VLC para el envío de video y el NetEM para la modificación de los parámetros de calidad de servicio utilizados para la toma de datos para generar los modelos propuestos.

### A. CONFIGURACIÓN VLC

Como se explicó en la sección x.x. se utilizó VLC como servidor de video y el proceso de configuración se describirá a continuación.

Para empezar el proceso de envío del video en la barra superior del menú presione **Media>Stream** y debe aparecer una pantalla como se aprecia en la figura A1, elija el video a enviar presionando el botón **Add**, después de elegir el video a enviar presione el botón **Stream** de la parte de abajo.

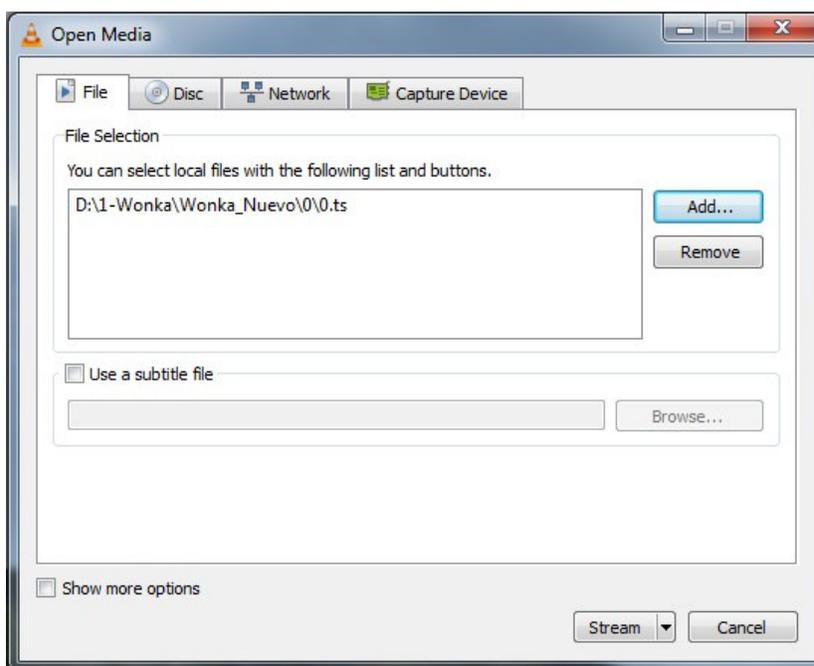


Figura A1. Pantalla para la elección del video a enviar por la red.

Al presionar el botón **Stream** debe aparecer una pantalla como la que se aprecia en la figura A2. Presione **Next**.

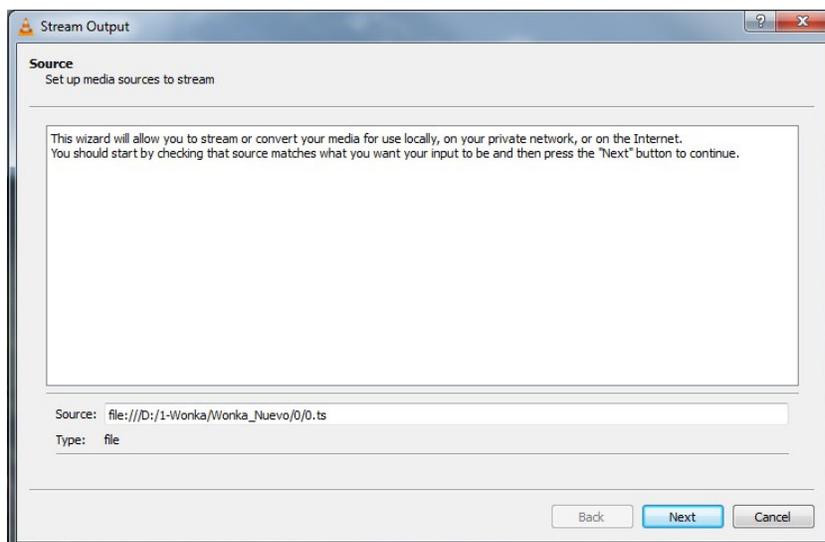


Figura A2. Pantalla de envío del video.

Después de realizado lo anterior debe aparecer una pantalla como la que se aprecia en la figura A3. Presione el botón **File** y elija la opción ***RTP/MPEG Transport Stream***, adicionalmente seleccione la opción que está en la parte izquierda de la pantalla, ***Display locally***, esto con el fin de que el video se pueda ver al momento del envío.

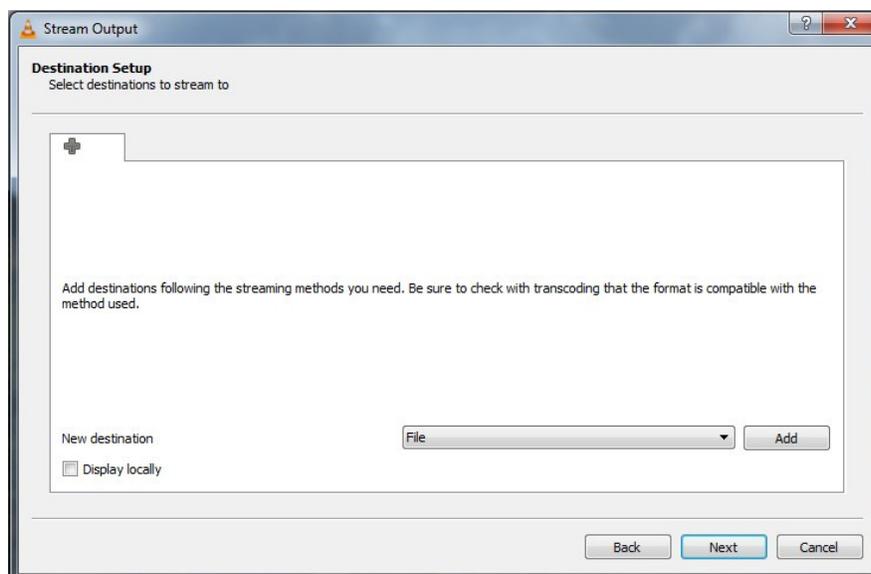


Figura A3. Pantalla *Destination Setup*.

Después de seleccionada la opción ***RTP/MPEG Transport Stream*** de la pantalla de la figura A3, presione el botón **Add** que está al lado derecho y debe aparecer una pantalla como la que se aprecia en la figura A4. En esta pantalla se configura el puerto por donde se enviará el video y la dirección de *broadcast* o *multicast* para el envío del video.

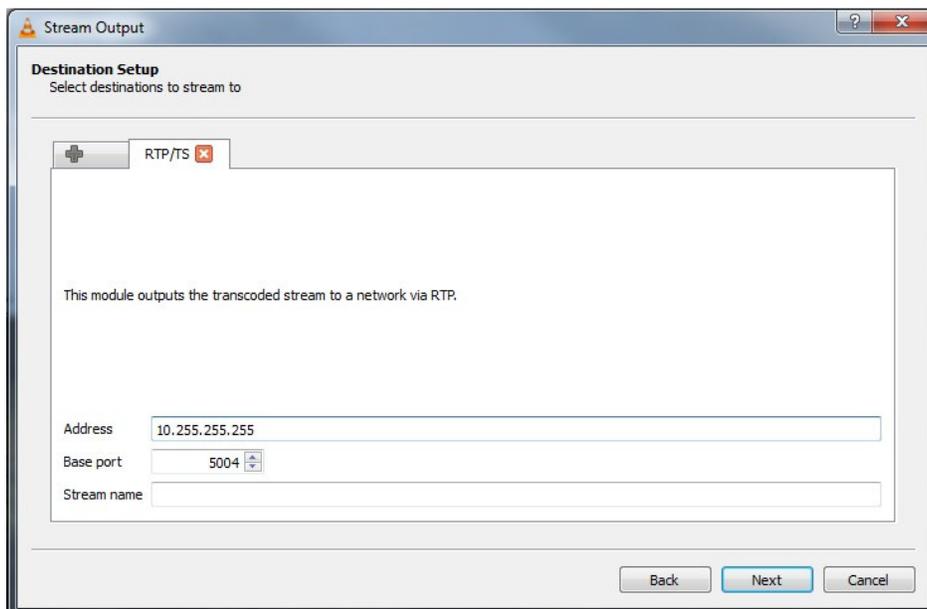


Figura A4. Configuración del puerto y dirección IP de broadcast o multicast de envío.

Después de configurado lo anterior, presione **Next** y debe aparecer una pantalla como se aprecia en la figura A5. En la opción **profile** se elige **Video – H.264 + MP3(MP4)** y la opción **Activate Transcoding** se deja sin habilitar, esto con el fin de reducir consumo de recursos de procesamiento en el servidor. La transcoding es un proceso en el cual se toma el video que se está transmitiendo y se convierte en un formato diferente o se cambia la rata de transmisión. Para este caso no se requiere, ya que se va a analizar solo el stream transmitido.

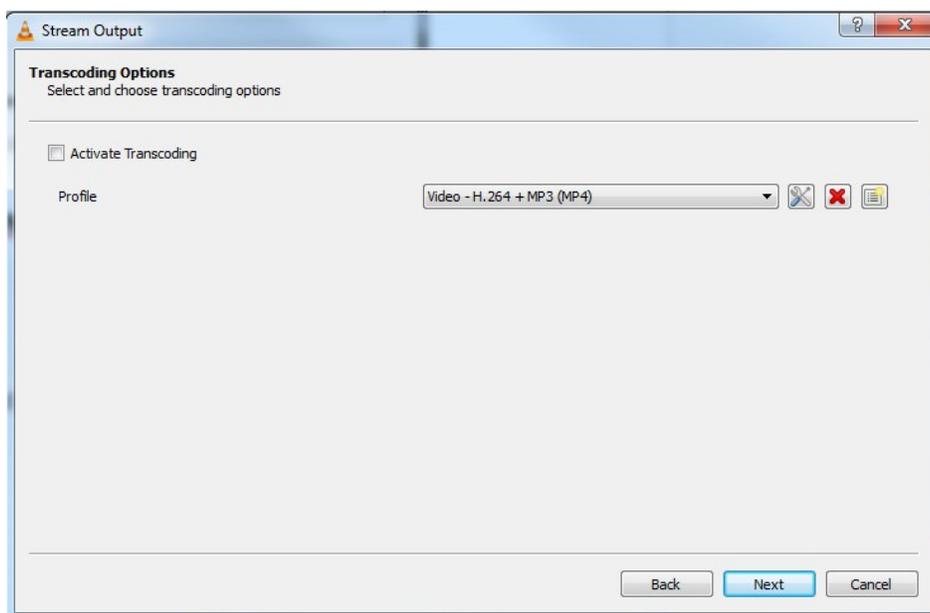
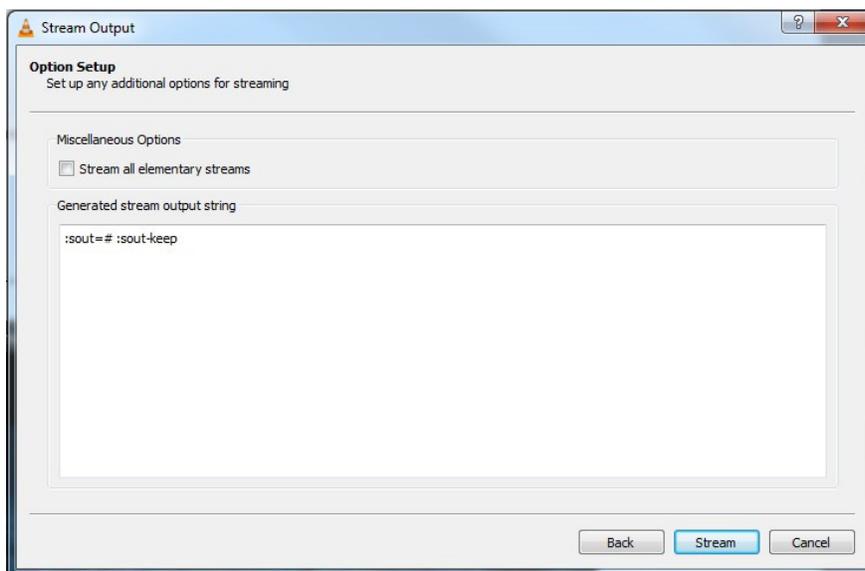


Figura A5. Pantalla de Trascoding Options.

Configurado lo anterior, se presiona **Next** y debe aparecer una pantalla como la que se aprecia en la figura A6. Para empezar la transmisión se presiona el botón de **Stream**.

Figura A6. Pantalla de *Option Setup*.

En el cliente VLC, en una ventana de comandos digitar el siguiente comando:

```
vlc -vvw protocolo://@direccion-ip-stream:puerto --sout file/extension-del-archivo-a-guardar:  
"ubicacion-del-archivo.extension"
```

Con este comando, la grabación del archivo empieza automáticamente al momento de ser recibido.

Posterior al envío y grabación del archivo en el cliente, se procede a la edición de los archivos con el fin de recortarlos a 10 segundos y proceder a la medición de VQM con la herramienta MSU Video Quality Measurement Tool.

## B. CONFIGURACIÓN NetEM

Para utilizar el NetEm se generaron tres de scripts a fin de optimizar el uso de la herramienta al momento de hacer las pruebas. Los scripts se encuentran en la CD entregado con este documento en la carpeta NetEm.

Los archivos que se deben instalar en la raíz de la sesión de administrador son: *EmuRed*, *ConfigPuentes.sh* y *DetenerPuentes.sh*. *EmuRed* es el archivo principal que se encarga de ejecutar *ConfigPuentes.sh* y *DetenerPuentes.sh*, estos dos archivos configuran el servidor en modo puente y al finalizar las pruebas se encargan de detener el modo puente del servidor. Adicionalmente *EmuRed*, se encarga de recibir por parte del usuario los valores de los parámetros de QoS a configurar.

En la figura A7, se puede apreciar el resultado de ejecutar *EmuRed* y configurar el retardo con 325 mseg, la variación del retardo con 0.009 mseg y la pérdida de paquetes con el 0.09%.

```

jcc@jcc-Vostro-230: ~/Netem
jcc@jcc-Vostro-230:~/Netem$ ls
ConfigPuentes.sh  DetenerPuentes.sh~  NetemManualScript
ConfigPuentes.sh~  EmuRed              RegistroTiempo.ods
DetenerPuentes.sh  NetemAutoScript     ScriptPruebaRed.ods
jcc@jcc-Vostro-230:~/Netem$ ./EmuRed
SCRIPT PARA EMULACIÓN DE RED
Seleccione qué Interfaz desea Modificar
0

¿Qué retardo desea introducir a la red?
Valor en ms. Introduzca 0 si no desea utilizar este parámetro.
325

¿Qué desviación temporal desea introducir a la red?
Valor en ms. Introduzca 0 si no desea utilizar este parámetro.
0.009

¿Qué pérdida de paquetes desea introducir a la red?
Valor en porcentaje. Introduzca 0 si no desea utilizar este parámetro.
0.09

Por Favor Confirme los Parámetros Deseados:
INTERFAZ: eth0
Delay: 325ms
Jitter: 0.009ms
Pérdida: 0.09%

[Y/N]
y

OK. Pulse Cualquier tecla para continuar.

sudo tc qdisc add dev eth0 root netem delay 325ms 0.009ms loss 0.09%

Si desea DETENER las modificaciones de la interfaz, escriba 'stop'.
Si desea CAMBIAR las modificaciones de la interfaz, escriba 'reset'.

```

Figura A7. Configuración de parámetros de QoS en el NetEM.



## Anexo 2: Pruebas Subjetivas

En este anexo se describirá el proceso realizado para la implementación de la base de datos de videos que se utilizó para las pruebas de desempeño de los modelos propuestos. Se empieza describiendo el procedimiento llevado a cabo para la evaluación subjetiva de los videos, luego se presenta el sitio donde se llevaron a cabo las pruebas subjetivas, se describen las características espacio-temporales de los videos utilizados, para finalizar con la descripción de los observadores que realizaron la prueba.

### Procedimiento de la evaluación subjetiva de los videos

La prueba de evaluación subjetiva se realizó para determinar el MOS de los primeros 15 tratamientos de cada video utilizado para generar los modelos propuestos en el capítulo 3, más el tratamiento cero que no tenía ningún tipo de distorsión. Y adicionalmente, 16 secuencias de un video con características espacio temporales similares al video de la cafetería. Cada secuencia de video tenía una duración de 10 segundos.

Para las pruebas subjetivas, cada observador debió evaluar 70 videos en total (6 secuencias de prueba, 4 secuencias sin distorsión, 60 secuencias distorsionadas). Para evaluar las secuencias de vídeo se utilizó el método de clasificación absoluta con referencia oculta (ACR-HR de la norma ITU-T P.910) en una escala de calidad de nueve grados.

La escala de evaluación fue diseñada con nueve grados de ponderación, donde 5 era la calificación más alta calidad y con valores intermedios hasta llegar a la menor ponderación que es 1, como se aprecia en la tabla A-1.

Tabla A-1. Escala MOS utilizada en la evaluación de los videos.

Nivel	1	2	3	4	5	6	7	8	9
Calificación	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0

El proceso de evaluación realizado por cada observador se realizó de la siguiente manera:

- Al finalizar la presentación de un video se colocaba una pantalla de color gris que tenía una duración de 5 segundos, esto con el fin de que el observador pudiera evaluar el video visto en un formato impreso que contenía la escala presentada en la tabla A-1.

- En el tercer segundo de visualizada la pantalla gris, sonaba un pitido indicando al observador que la pantalla gris se retiraría en dos segundos. Esto indicaba que el siguiente video se iba a proyectar.

## Descripción del Espacio de Proyección

Las pruebas subjetivas se realizaron en una sala de video que fue adecuada para cumplir con los requisitos de la recomendación UIT-R BT.500-13, como se muestra en la figura A8. Las condiciones de iluminación de la sala y la calibración de la pantalla del televisor se midieron y ajustaron usando un **X-rite Color Munki Display** [103]. El televisor utilizado fue un "Samsung (UN32D6000) de 32 pulgadas con una resolución nativa de 1080p, la distancia del observador a la pantalla era de 3 veces la altura de la imagen.



Figura A8: Espacio de Proyección y realización de las pruebas.

## Descripción de los Videos Utilizados

Los videos fuente fueron grabados utilizando una cámara Nikon D3100 (single-lens reflex), cada video fue grabado con una resolución de 1080x1920, 24 imágenes por segundo o cuadros por segundo (*frame rate* - 23.976 fps) y codificado en H.264/MPEG-4 en formato MOV (Formato de video y audio creado por Apple). Los videos fuente tenían una duración de 30 segundos y después de ser transmitidos en la maqueta de experimentación, fueron editados a una duración de 10 segundos.

Los nombres de los videos fueron codificados de la siguiente manera para realizar las pruebas subjetivas, los videos identificados con 00 indican que son los videos fuente que no tienen ningún tipo de distorsión aplicada:

- B00 - Entrenamiento de baloncesto filmado desde un costado de la cancha.

- C00 - Cafetería con peatones cruzando.
- E00 - Una escena con plantas y un arroyo.
- I00 - Escena de una entrevista.

En la tabla A-2, se pueden apreciar los videos elegidos para la las pruebas subjetivas. Se presenta el nombre del video, una captura de pantalla (*screenshot*) del mismo, su información espacial (IE), su información temporal (IT) y la clase de video.

De acuerdo a la Recomendación UIT-T P.910 se calcularon los valores IE e IT (información espacial y temporal) de cada video para determinar objetivamente su cantidad de movimiento y para identificar el número de objetos y detalles de cada escena.

Tabla A-2. Características espacio temporales de los videos usados en la prueba.

Nombre del video		IE	IT	Clase
Entrevista (I00)		35,9166	5,38516	Movimiento Bajo
Cafetería (C00)		65,9848	6,48074	Movimiento Moderado
Ambiente (E00)		65,6734	6,63325	Movimiento Moderado

Nombre del video		IE	IT	Clase
Baloncesto (B00)		57,4195	17,4069	Movimiento Alto

En la Figura A9 presenta la información de manera gráfica, en donde se puede apreciar que las escenas E00 & C00 poseen valores prácticamente iguales.

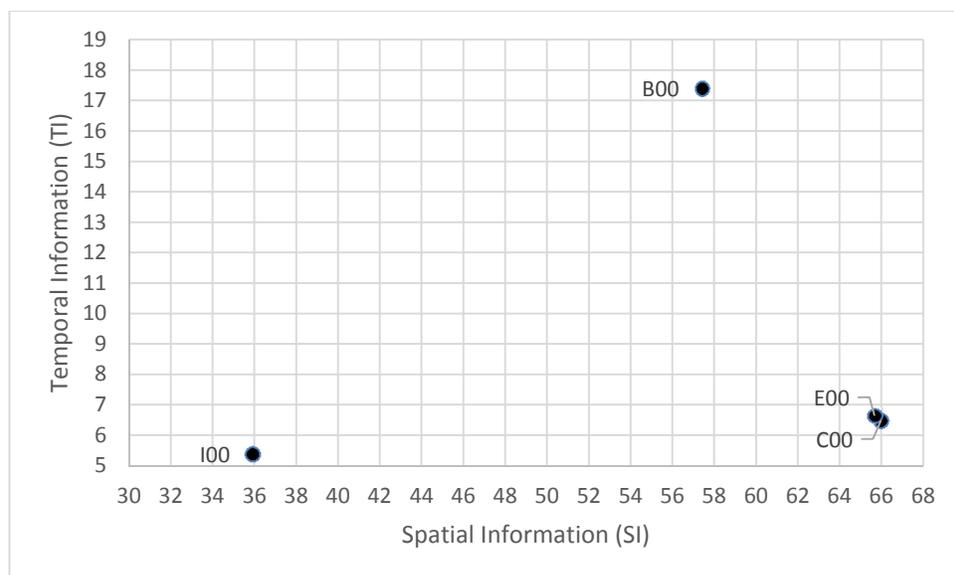


Figura A9. Distribución de los valores SI & TI de los videos originales utilizados en las pruebas subjetivas.

### Descripción de Observadores que realizaron la Prueba

En total, 68 observadores participaron en la prueba subjetiva, de los cuales 57 fueron considerados válidos según el cumplimiento de las instrucciones impartidas durante el experimento, el correcto uso de los formatos establecidos y también se realizaron pruebas siguiendo los pasos de la norma ITU 500-13. En la tabla A-3 se aprecia la descripción de los observadores que participaron de la prueba.

Tabla A-3. Participantes por categoría.

TOTAL SUJETOS	CATEGORÍA	Cantidad	Rango Edad	Cruce de Categorías	
	FEMENINO			30	17-39 años
	MASCULINO	27	17-53 años	19 Estudiantes	8 Colaboradores
57	COLABORADORES	23	24-53 años	8 Hombres	15 Mujeres
	ESTUDIANTES	34	17-26 años	19 Hombres	15 Mujeres

La calificación de cada video fue el promedio de las calificaciones de los sujetos válidos. De igual manera se sacaron calificaciones promedio de los videos separando los observadores diferentes categorías para comparar los distintos grupos. Las 6 secuencias ficticias utilizadas para entrenar a los observadores no fueron tomadas en cuenta para la base de datos ni el post-procesamiento de las puntuaciones subjetivas.

Cada video fue identificado de manera consecuente con el tratamiento del cual fue obtenido, por ejemplo el video C01, indica que el video es cafetería asociado al tratamiento No.1 que fue obtenido al configurar el retardo con 325 mseg, la variación del retardo con 0.0045 mseg y la perdida de paquetes con 0,045%.

En las tablas A-3, A-4, A-5 y A-6 se presentan los resultados obtenidos en la evaluación de cada video.

Analizando el MOS promedio de cada tabla se aprecia que el video Cafetería fue el video que obtuvo menor MOS promedio, esto se pudo dar debido a que los evaluadores detectaban de manera más fácil las distorsiones que se presentaban en el video y para su criterio el daño era evidente. Contrario a lo ocurrido en el video entrevista donde el MOS promedio fue mucho más elevado que el MOS promedio obtenido en el video de Cafetería. Esto pudo ocurrir debido a que para el usuario distorsiones en el fondo del entrevistado eran imperceptibles y para el usuario primaba la calidad del video en la zona de la cara.

En el caso del video de baloncesto, los hombres lo evaluaron de manera más crítica por las mujeres (MOS promedio masculino: 3,39 – MOS promedio femenino:3,58) esto se debe a que los hombres por la costumbre de ver deportes analizan mejor la imagen que las mujeres.

Tabla A-3. Resultados Pruebas subjetivas Video Baloncesto.

VIDEOS BALONCESTO					
VIDEO	FEMENINO	MASCULINO	COLABORADORES	ESTUDIANTES	MOS Promedio
B00	3,47	3,83	3,50	3,74	3,64
B01	4,75	4,72	4,72	4,75	4,89
B02	4,41	4,31	4,34	4,38	4,43
B03	3,17	2,96	3,15	3,01	3,07
B04	3,15	3,09	3,15	3,10	3,12
B05	3,13	2,94	3,15	2,97	3,04
B06	3,45	2,98	3,41	3,10	3,22
B07	3,08	2,48	3,02	2,65	2,79
B08	3,37	3,13	3,37	3,18	3,25
B09	3,58	3,52	3,65	3,49	3,55
B10	3,28	3,22	3,37	3,18	3,25
B11	3,48	2,94	3,59	2,99	3,22
B12	3,35	3,06	3,46	3,04	3,21
B13	4,66	4,55	4,64	4,59	4,74
B14	3,35	3,26	3,39	3,25	3,30
B15	3,62	3,31	3,48	3,47	3,47

Tabla A-4. Resultados Pruebas subjetivas Video Cafetería.

VIDEO CAFETERÍA					
VIDEO	FEMENINO	MASCULINO	COLABORADORES	ESTUDIANTES	MOS Promedio
C00	4,22	4,20	4,15	4,25	4,2
C01	3,96	4,13	3,98	4,07	4,1
C02	3,42	3,59	3,40	3,57	3,5
C03	2,60	2,69	2,85	2,50	2,6
C04	2,98	3,11	3,15	2,97	3,0
C05	3,47	3,63	3,48	3,59	3,5
C06	3,78	4,11	4,02	3,88	3,9
C07	2,82	3,07	3,07	2,85	2,9
C08	3,22	3,38	3,48	3,16	3,3
C09	2,72	2,67	2,83	2,60	2,7
C10	2,17	2,22	2,37	2,07	2,2
C11	3,53	3,76	3,70	3,60	3,6
C12	2,88	3,15	2,96	3,04	3,0
C13	3,75	3,76	3,71	3,79	3,8
C14	2,88	3,02	2,91	2,97	2,9
C15	3,03	3,10	3,36	2,87	3,1

Tabla A-5. Resultados Pruebas subjetivas Video Entrevista.

VIDEO ENTREVISTA					
VIDEO	FEMENINO	MASCULINO	COLABORADORES	ESTUDIANTES	MOS Promedio
I00	4,40	4,61	4,46	4,53	4,500
I01	4,00	4,05	3,96	4,06	4,061
I02	4,09	4,05	4,21	3,98	4,096
I03	3,45	3,44	3,37	3,50	3,474
I04	3,55	3,69	3,33	3,81	3,632
I05	4,22	4,17	4,41	4,05	4,228
I06	2,68	2,65	2,72	2,63	2,667
I07	3,77	3,86	3,89	3,76	3,825
I08	3,55	3,43	3,74	3,32	3,491
I09	4,29	4,27	4,29	4,27	4,333
I10	3,43	3,55	3,37	3,57	3,509
I11	4,19	4,25	4,17	4,25	4,237
I12	4,22	4,32	4,19	4,32	4,342
I13	3,54	3,83	3,64	3,71	3,693
I14	3,08	3,00	3,07	3,03	3,044
I15	4,08	4,29	4,23	4,14	4,211

Tabla A-6. Resultados Pruebas subjetivas Video Ambiente.

VIDEO AMBIENTE					
VIDEO	FEMENINO	MASCULINO	COLABORADORES	ESTUDIANTES	MOS Promedio
E00	4,48	4,70	4,63	4,56	4,59
E01	3,81	3,91	3,85	3,86	3,86
E02	3,81	3,75	3,85	3,74	3,78
E03	4,61	4,45	4,36	4,65	4,54
E04	4,43	4,25	4,22	4,43	4,34
E05	4,30	4,33	4,37	4,28	4,31
E06	3,65	3,44	3,80	3,39	3,55
E07	3,42	3,20	3,35	3,29	3,32
E08	3,71	3,75	3,74	3,72	3,73
E09	2,65	2,72	2,83	2,58	2,68
E10	1,66	1,46	1,50	1,61	1,57
E11	3,75	3,96	3,70	3,95	3,85
E12	3,82	3,82	3,86	3,79	3,82
E13	4,08	4,08	4,04	4,11	4,08
E14	2,33	1,89	2,21	2,06	2,12
E15	3,81	3,92	3,93	3,81	3,86



## 7 Bibliografía

- [1] A. A. García, J. . Cuellar, y (primero), «Calidad de servicio en proveedores de servicios IPTV», *Ingenium*, vol. 6, n.º 13, pp. 11-25, sep. 2012.
- [2] «Cisco Visual Networking Index: Forecast and Methodology, 2015-2020», *Cisco*, 18-ene-2017. [En línea]. Disponible en: <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html>. [Accedido: 18-ene-2017].
- [3] International Telecommunications Union (ITU), «ITU-T Recommendation G.1080: Quality of Experience requirements for IPTV Services.», *Ginebra Suiza*, dic. 2008.
- [4] A. Takahashi, D. Hands, y V. Barriac, «Standardization activities in the ITU for a QoE assessment of IPTV», *IEEE Commun. Mag.*, vol. 46, n.º 2, pp. 78-84, 2008.
- [5] W. Robitza *et al.*, «Challenges of future multimedia QoE monitoring for internet service providers», *Multimed. Tools Appl.*, pp. 1-24, jun. 2017.
- [6] A. R. Reibman, V. A. Vaishampayan, y Y. Sermadevi, «Quality monitoring of video over a packet network», *IEEE Trans. Multimed.*, vol. 6, n.º 2, pp. 327-334, 2004.
- [7] N. Liao y Z. Chen, «A packet-layer video quality assessment model with spatiotemporal complexity estimation», *EURASIP J. Image Video Process.*, vol. 2011, n.º 1, p. 5, ago. 2011.
- [8] C. J. Pardo, C., «Revisión Sistemática de la Armonización de Marcos en la Mejora de los Procesos de Software.» Universidad Castilla - La Mancha, ene-2009.
- [9] J. C. Cuéllar, J. H. Ortiz, y J. L. Arciniegas, «Clasificación y Análisis de Métodos para medir Calidad de la Experiencia del Servicio de Televisión sobre Protocolo IP (IPTV)», *Inf. Tecnológica*, vol. 25, n.º 5, pp. 121-128, 2014.
- [10] T. Rahrer, R. Fiandra, y S. Wright, «Technical Report. TR-126. Triple-play Services Quality of Experience (QoE) Requirements.», *DSL Forum*, 13-dic-2006. [En línea]. Disponible en: <https://www.broadband-forum.org/technical/download/TR-126.pdf>. [Accedido: 18-ene-2017].
- [11] International Telecommunication Union (ITU), «ITU-R Recommendation ITU-R BS.775-1, Multichannel stereophonic sound system with and without accompanying picture.», *Ginebra Suiza*, jul. 1994.
- [12] International Telecommunication Union (ITU), «ITU-R Recommendation BS.1286. Methods for the subjective assessment of audio systems with accompanying picture.», *Ginebra Suiza*, 1997.
- [13] International Telecommunication Union (ITU), «ITU-R Recommendation BT.500-13, Methodology for subjective assesment of the quality of television pictures.», *Ginebra Suiza*, 2002.
- [14] International Telecommunication Union (ITU), «ITU-T Recommendation P.910, Subjective video quality assesment methods for multimedia applications.», *Ginebra Suiza*, 2008.

- [15] Y. Chen, K. Wu, y Q. Zhang, «From QoS to QoE: A Tutorial on Video Quality Assessment», *IEEE Commun. Surv. Tutor.*, vol. 17, n.º 2, pp. 1126-1165, Secondquarter 2015.
- [16] «UIT-T Recommendation P.800 Métodos de determinación subjetiva de la calidad de transmisión», *Ginebra, Suiza*, ago. 1996.
- [17] H. Rifai, S. Mohammed, y A. Mellouk, «A brief synthesis of QoS-QoE methodologies», presentado en 2011 10th International Symposium on Programming and Systems (ISPS), Piscataway, NJ, 2011, pp. 32-38.
- [18] S. Winkler, «Video quality measurement standards - Current status and trends», presentado en 7th International Conference on Information, Communications and Signal Processing, 2009. ICICS 2009, Piscataway, NJ, 2009, pp. 1-5.
- [19] S. Winkler y P. Mohandas, «The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics», *IEEE Trans. Broadcast.*, vol. 54, n.º 3, pp. 660-668, 2008.
- [20] R. Serral-Gracià, E. Cerqueira, M. Curado, M. Yannuzzi, E. Monteiro, y X. Masip-Bruin, «An overview of quality of experience measurement challenges for video applications in IP networks», presentado en WWIC'10 Proceedings of the 8th international conference on Wired/Wireless Internet Communications. Springer-Verlag, Berlin, Heidelberg, 2010, pp. 252–263.
- [21] J. You, U. Reiter, M. M. Hannuksela, M. Gabbouj, y A. Perkis, «Perceptual-based Quality Assessment for Audio-visual Services: A Survey», *Image Commun.*, vol. 25, n.º 7, pp. 482–501, ago. 2010.
- [22] S. Daly, «The visible differences predictor: An algorithm for the assessment of image fidelity», en *Digital images and human vision*, A. B. Watson, Ed. Cambridge, MA, USA: MIT Press, 1993, pp. 179–206.
- [23] J. Lubin y D. Fibush, «Sarnoff JND vision model». TIA 1.5 Working Group Document #97-612, ANSI TI Standards Committee, 1997.
- [24] C. J. Van den Branden Lambrecht y O. Verscheure, «Perceptual quality measure using a spatiotemporal model of the human visual system», 1996, vol. 2668, pp. 450-461.
- [25] S. Winkler, «A perceptual distortion metric for digital color video», presentado en Proceedings of SPIE, Whashington, D.C., 1999, pp. 175–184.
- [26] J. You, U. Reiter, M. M. Hannuksela, M. Gabbouj, y A. Perkis, «Perceptual-based quality assessment for audio-visual services: A survey», *Signal Process. Image Commun.*, vol. 25, n.º 7, pp. 482-501, ago. 2010.
- [27] Z. Wang, L. Lu, y A. C. Bovik, «Video Quality Assessment Based on Structural Distortion Measurement», *Signal Process. Image Commun.*, vol. 19, n.º 1, pp. 1-9, ene. 2004.
- [28] H. R. Sheikh, M. F. Sabir, y A. C. Bovik, «A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms», *IEEE Trans. Image Process.*, vol. 15, n.º 11, pp. 3440-3451, nov. 2006.
- [29] Z. Wang y L. Qiang, «Video Quality Assessment Using a Statistical Model of Human Visual Speed Perception», *J. Opt. Soc. Am. A*, vol. 24, n.º 12, pp. B61-B69, dic. 2007.
- [30] K. Seshadrinathan y A. C. Bovik, «Motion-based Perceptual Quality Assessment of Video», *Proc. SPIE - Human Vision and Electronic Imaging*, 2009.

- [31] D. M. Chandler y S. S. Hemami, «VSNR: A Wavelet-Based Visual Signal-to-Noise Ratio for Natural Images», *IEEE Trans. Image Process.*, vol. 16, n.º 9, pp. 2284-2298, 2007.
- [32] M. H. Pinson y S. Wolf, «A new standardized method for objectively measuring video quality», *IEEE Trans. Broadcast.*, vol. 50, n.º 3, pp. 312-322, 2004.
- [33] A. Catellier y M. Pinson, «Characterization of the HEVC Coding Efficiency Advance Using 20 Scenes, ITU-T Rec. P.913 Compliant Subjective Methods, VQM, and PSNR», en *2015 IEEE International Symposium on Multimedia (ISM)*, 2015, pp. 282-288.
- [34] T. Wang, A. Pervez, y H. Zou, «VQM-based QoS/QoE mapping for streaming video», en *2010 3rd IEEE International Conference on Broadband Network and Multimedia Technology (IC-BNMT)*, 2010, pp. 807-812.
- [35] «ITU-T Recommendation J.144 Objective Perceptual Video Quality Measurement Techniques for Digital Cable Television in the presence of a full reference.», *Ginebra Suiza*, mar. 2004.
- [36] «ITU-T Recommendation J.148 Requirements for an Objective Perceptual Multimedia Quality Model.», *Ginebra Suiza*, may 2003.
- [37] «ITU-T Recommendation G.1070 Opinion model for video-telephony applications», *Ginebra Suiza*, abr. 2007.
- [38] «ITU-T Recommendation P.1202 Parametric non-intrusive bitstream assessment of video media streaming quality (P.NBAMS)», *Ginebra Suiza*, oct. 2012.
- [39] A. G. Davis, D. Bayart, y D. S. Hands, «Quality assurance for IPTV», presentado en *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, 2009. BMSB '09, Piscataway, NJ, 2009, pp. 1-7.
- [40] Tobias Hoßfeld, David Hock, Phuoc Tran-Gia, Kurt Tutschku, y Markus Fiedler, «Testing the IQX Hypothesis for Exponential Interdependency between QoS and QoE of Voice Codecs iLBC and G.711», *Wurzburg Alem. Univesity Wurzburg*, 2008.
- [41] M. Fiedler, T. Hoßfeld, y P. Tran-Gia, «A generic quantitative relationship between quality of experience and quality of service», *IEEE Netw.*, vol. 24, n.º 2, pp. 36-41, 2010.
- [42] M. Fiedler y T. Hoßfeld, «Quality of Experience-Related Differential and Provisioning-Delivery Hysteresis.», *21st Int. Teletraffic Congr. Spec. Semin. Multimed. Appl. Perform. QoE*, vol. Tokio, Japón., mar. 2010.
- [43] T. Hoßfeld, M. Fiedler, y T. Zinner, «The QoE provisioning-delivery-hysteresis and its importance for service provisioning in the Future Internet», presentado en *2011 7th EURO-NGI Conference on Next Generation Internet (NGI)*, Piscataway, NJ, 2011, pp. 1-8.
- [44] H.-J. Kim, D. H. Lee, J. M. Lee, K.-H. Lee, W. Lyu, y S.-G. Choi, «The QoE Evaluation Method through the QoS-QoE Correlation Model», presentado en *Fourth International Conference on Networked Computing and Advanced Information Management*, 2008. NCM '08, 2008, vol. 2, pp. 719-725.
- [45] H.-J. Kim y S.-G. Choi, «A study on a QoS/QoE correlation model for QoE evaluation on IPTV service», presentado en *2010 The 12th International*

- Conference on Advanced Communication Technology (ICACT), Seúl, Corea., 2010, vol. 2, pp. 1377-1382.
- [46] H.-J. Kim, D.-G. Yun, H.-S. Kim, K.-S. Cho, y S.-G. Choi, «QoE assessment model for video streaming service using QoS parameters in wired-wireless network», presentado en 2012 14th International Conference on Advanced Communication Technology (ICACT), Seúl, Corea., 2012, pp. 459-464.
- [47] L. Li-yuan, Z. Wen-an, y S. Jun-de, «The Research of Quality of Experience Evaluation Method in Pervasive Computing Environment», presentado en 2006 1st International Symposium on Pervasive Computing and Applications, 2006, pp. 178-182.
- [48] K. Yamagishi, T. Kawano, y T. Hayashi, «Hybrid Video-Quality-Estimation Model for IPTV Services», presentado en IEEE Global Telecommunications Conference, 2009. GLOBECOM 2009, 2009, pp. 1-5.
- [49] A. Riker, E. Aguiar, D. Cardoso, P. De Oliveira, E. Cerqueira, y A. Abelém, «A Hybrid Prediction and Assessment Quality of Experience Approach for Videostreaming Applications over Wireless Mesh Networks», *XXIX Simpósio Bras. Redes Comput. E Sist. Distrib.*, vol. Campo Grande, Brasil., 2011.
- [50] I. Abdeljaouad, G. Kandavanam, y A. Karmouch, «A Loss-Based Utility Function for Predicting IPTV Quality of Experience over an Overlay Network», presentado en 2011 IEEE Global Telecommunications Conference (GLOBECOM 2011), 2011, pp. 1-6.
- [51] «Sigmoid Function». [En línea]. Disponible en: [http://research.omicsgroup.org/index.php/Sigmoid\\_function](http://research.omicsgroup.org/index.php/Sigmoid_function). [Accedido: 17-feb-2017].
- [52] Z. Wang, A. C. Bovik, H. R. Sheikh, y E. P. Simoncelli, «Image quality assessment: from error visibility to structural similarity», *IEEE Trans. Image Process.*, vol. 13, n.º 4, pp. 600-612, abr. 2004.
- [53] S. Aroussi, T. Bouabana-Tebibel, y A. Mellouk, «Empirical QoE/QoS correlation model based on multiple parameters for VoD flows», presentado en 2012 IEEE Global Communications Conference (GLOBECOM), 2012, pp. 1963-1968.
- [54] M. Elhoseny, A. Shehab, y L. Osman, «An Empirical Analysis of User Behavior for P2P IPTV Workloads. In: In: Hassanien A., Tolba M., Elhoseny M., Mostafa M. (eds) The International Conference on Advanced Machine Learning Technologies and Applications (AMLTA2018). AMLTA 2018.», *Adv. Intell. Syst. Comput. Vol 723 Springer Cham*.
- [55] A. Shehab, M. El Aziz, y A. Hassanien, «Efficient Schemes for Playout Latency Reduction in P2P-VoD Systems. In: Hassanien A., Oliva D. (eds) Advances in Soft Computing and Machine Learning in Image Processing», *Stud. Comput. Intell. 2018 Vol 730 Springer Cham*.
- [56] A. Shehab, M. Elhoseny, y A. Hassanien, «An Efficient Scheme for Video Delivery in Wireless Networks. In: Hassanien A., Elhoseny M., Kacprzyk J. (eds) Quantum Computing», *Environ. Intell. Large Scale Real Appl. Stud. Big Data 2018 Vol 33 Springer Cham*.
- [57] C. M. Lentisco, L. Bellido, J. C. C. Q, E. Pastor, y J. L. A. H, «QoE-Based Analysis of DASH Streaming Parameters Over Mobile Broadcast Networks», *IEEE Access*, vol. 5, pp. 20684-20694, 2017.

- [58] «ITU-T Recommendation Y.1540. IP Packet Transfer and Availability Performance Parameters.», dic. 2002.
- [59] «ITU-T Recommendation Y.1541. Network Performance Objectives for IP-Based Services.», feb. 2006.
- [60] G. Baltoglou, E. Karapistoli, y P. Chatzimisios, «IPTV QoS and QoE measurements in wired and wireless networks», presentado en 2012 IEEE Global Communications Conference (GLOBECOM), 2012, pp. 1757-1762.
- [61] «ITU-T Recommendation P.1202.2 Parametric non-intrusive bitstream assessment of video media streaming quality – higher resolution application area», *Ginebra Suiza*, may 2013.
- [62] «Video LAN Organization». [En línea]. Disponible en: <https://www.videolan.org/vlc/index.es.html>.
- [63] «Analizador de Tráfico Wireshark». [En línea]. Disponible en: <https://www.wireshark.org/>.
- [64] «H.264/AVC Context Adaptive Binary Arithmetic Coding (CABAC)», *Vcodex*. [En línea]. Disponible en: <http://www.vcodex.com/h264avc-context-adaptive-binary-arithmetic-coding-cabac/>. [Accedido: 08-mar-2017].
- [65] «H.264/AVC Context Adaptive Variable Length Coding», *Vcodex*. [En línea]. Disponible en: <http://www.vcodex.com/h264avc-context-adaptive-variable-length-coding/>. [Accedido: 08-mar-2017].
- [66] «Introducción a la compresión: Huffman y Entropía.» [En línea]. Disponible en: <http://www.grc.upv.es/docencia/tadm/practicasp1/P1.pdf>. [Accedido: 08-mar-2017].
- [67] A. Mittal, M. A. Saad, y A. C. Bovik, «A Completely Blind Video Integrity Oracle», *IEEE Trans. Image Process.*, vol. 25, n.º 1, pp. 289-300, ene. 2016.
- [68] K. Seshadrinathan, R. Soundararajan, A. C. Bovik, y L. K. Cormack, «Study of Subjective and Objective Quality Assessment of Video», *IEEE Trans. Image Process.*, vol. 19, n.º 6, pp. 1427-1441, 2010.
- [69] S. Chikkerur, V. Sundaram, M. Reisslein, y L. J. Karam, «Objective Video Quality Assessment Methods: A Classification, Review, and Performance Comparison», *IEEE Trans. Broadcast.*, vol. 57, n.º 2, pp. 165-182, 2011.
- [70] M. H. Pinson y S. Wolf, «A new standardized method for objectively measuring video quality», *IEEE Trans. Broadcast.*, vol. 50, n.º 3, pp. 312-322, sep. 2004.
- [71] «The Linux Foundation». [En línea]. Disponible en: <https://wiki.linuxfoundation.org/networking/netem>.
- [72] L. Lu y X. Lu, «Quality assessing of video over a packet network», en *Second Workshop on Digital Media and its Application in Museum Heritages (DMAMH 2007)*, 2007, pp. 365-369.
- [73] L. Cruvinel y T. Vazao, «A Simple Metric for Predicting Video Quality of Experience.», *10º Conf. Sobre Redes Comput. CRC 2010*, vol. Universidade do Minho, Braga, n.º Portugal., pp. 77-82.
- [74] H. J. Kim y S. G. Choi, «QoE assessment model for multimedia streaming services using QoS parameters», *Multimed. Tools Appl.*, vol. 72, n.º 3, pp. 2163-2175, oct. 2014.
- [75] R. Lübke, P. Büschel, D. Schuster, y A. Schill, «Measuring accuracy and performance of network emulators», en *2014 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, 2014, pp. 63-65.

- [76] I. Sobel y G. Feldman, *A 3x3 Isotropic Gradient Operator for Image Processing*, Artificial Intelligence Project (SAIL). Stanford -USA, 1968.
- [77] W. Song y D. W. Tjondronegoro, «Acceptability-Based QoE Models for Mobile Video», *IEEE Trans. Multimed.*, vol. 16, n.º 3, pp. 738-750, abr. 2014.
- [78] A. K. Moorthy, L. K. Choi, A. C. Bovik, y G. de Veciana, «Video Quality Assessment on Mobile Devices: Subjective, Behavioral and Objective Studies», *IEEE J. Sel. Top. Signal Process.*, vol. 6, n.º 6, pp. 652-671, oct. 2012.
- [79] M. Leszczuk, M. Hanusiak, M. C. Q. Farias, E. Wyckens, y G. Heston, «Recent developments in visual quality monitoring by key performance indicators», *Multimed. Tools Appl.*, vol. 75, n.º 17, pp. 10745-10767, sep. 2016.
- [80] P. Paudyal, F. Battisti, y M. Carli, «Impact of video content and transmission impairments on quality of experience», *Multimed. Tools Appl.*, vol. 75, n.º 23, pp. 16461-16485, dic. 2016.
- [81] C. Wang, X. Jiang, y Y. Wang, «Video Quality Assessment Models for IPTV Services», *Int. J. Digit. Content Technol. Its Appl.*, vol. 7, n.º 7, pp. 733-741, abr. 2013.
- [82] E. Danish, A. Fernando, M. Alreshoodi, y J. Woods, «A hybrid prediction model for video quality by QoS/QoE mapping in wireless streaming», en *2015 IEEE International Conference on Communication Workshop (ICCW)*, 2015, pp. 1723-1728.
- [83] Montgomery, *Diseño y análisis de experimentos*, Edición: 2 Tra. México, D.F.: Limusa, 2013.
- [84] «MSU Video Quality Measurement Tool», dic-2016. [En línea]. Disponible en: [http://compression.ru/video/quality\\_measure/video\\_measurement\\_tool.html](http://compression.ru/video/quality_measure/video_measurement_tool.html).
- [85] H. Gutierrez y R. De la Vara, *Análisis y Diseño de Experimentos*, 3.ª ed. McGraw-Hill Interamericana de España S.L, 2012.
- [86] Douglas A. Lind, William G. Marchal, y Samuel Wathen, *Estadística Aplicada a los Negocios y a la Economía*. McGrawHill., 2012.
- [87] «Minitab». [En línea]. Disponible en: <http://www.minitab.com/es-mx/>.
- [88] «VQEG Objective Video Quality Model Test Plan.» may-1998.
- [89] A. N. Moldovan, I. Ghergulescu, y C. H. Muntean, «VQAMap: A Novel Mechanism for Mapping Objective Video Quality Metrics to Subjective MOS Scale», *IEEE Trans. Broadcast.*, vol. 62, n.º 3, pp. 610-627, sep. 2016.
- [90] «An Introduction to Statistical Learning - with Applications in R | Gareth James | Springer». [En línea]. Disponible en: <https://www.springer.com/us/book/9781461471370>. [Accedido: 26-jun-2018].
- [91] «Reduced-reference metric design for objective perceptual quality assessment in wireless imaging - ScienceDirect». [En línea]. Disponible en: <http://www.sciencedirect.com/science/article/pii/S0923596509000836>. [Accedido: 15-jun-2017].
- [92] Video Quality Experts Group, «VQEG Final Report of FR-TV Phase II Validation Test.. "Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment, phase II".» Video Qual. Experts Group, Boulder, CO, USA, 2003.
- [93] R. F. Nau. (2015)., «What's a Good Value for R-Squared?» .

- [94] «Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment, phase II (FR-TV2). VQEG.» sep-2003.
- [95] N. Goran, B. Bečiragić, y M. Hadžialić, «An example of mapping the degradation of network parameters with video QoE parameters in case of IPTV service», en *2016 XI International Symposium on Telecommunications (BIHTEL)*, 2016, pp. 1-5.
- [96] F. Tommasi, V. De Luca, y C. Melle, «Packet losses and objective video quality metrics in H.264 video streaming», *J. Vis. Commun. Image Represent.*, vol. 27, pp. 7-27, feb. 2015.
- [97] O. B. Maia, H. C. Yehia, y L. de Errico, «A concise review of the quality of experience assessment for video streaming», *Comput. Commun.*, vol. 57, pp. 1-12, feb. 2015.
- [98] Y. A. Syahbana, Herman, A. A. Rahman, y K. A. Bakar, «Aligned-PSNR (APSNR) for Objective Video Quality Measurement (VQM) in video stream over wireless and mobile network», en *2011 World Congress on Information and Communication Technologies*, 2011, pp. 330-335.
- [99] T. Zinner, O. Hohlfeld, O. Abboud, y T. Hoßfeld, «Impact of frame rate and resolution on objective QoE metrics», en *2010 Second International Workshop on Quality of Multimedia Experience (QoMEX)*, 2010, pp. 29-34.
- [100] Z. Wang y A. C. Bovik, «A universal image quality index», *IEEE Signal Process. Lett.*, vol. 9, n.º 3, pp. 81-84, mar. 2002.
- [101] S. Winkler, *Digital Video Quality: Vision Models and Metrics*. Hoboken, NJ, USA: Wiley, 2005.
- [102] J. Gross, J. Klaue, H. Karl, y A. Wolisz, «Cross-layer optimization of OFDM transmission systems for MPEG-4 video streaming», *Comput. Commun.*, vol. 27, n.º 11, pp. 1044-1055, jul. 2004.
- [103] «x-rite pantone: <http://www.xrite.com/categories/calibration-profiling/colormunki-display>».