

Modelo para la medición de QoE en IPTV

Juan Carlos Cuéllar Q.
José Luis Arciniegas H.
Jesús Hamilton Ortiz



Modelo para la medición de QoE en IPTV

Modelo para la medición de QoE en IPTV

**Juan Carlos Cuéllar Q.
José Luis Arciniegas
Jesús Hamilton Ortiz**

Modelo para medir QoE en IPTV

© Juan Carlos Cuéllar Q., José Luis Arciniegas y Jesús Hamilton Ortiz.

1 ed. Cali, Colombia. Universidad Icesi, 2018

118 p., 19x24 cm

Incluye referencias bibliográficas

ISBN: 978-958-8936-56-7

<https://doi.org/10.18046/EUI/ee.3.2018>

1. IPTV 2. Quality assurance 3. Quality management I.Tit
621.38 – dc22

© Universidad Icesi, 2018

Facultad de Ingeniería

Rector: Francisco Piedrahita Plata

Decano Facultad de Ingeniería: Gonzalo Ulloa Villegas

Coordinador editorial: Adolfo A. Abadía



Producción y diseño: Claros Editores SAS.

Editor: José Ignacio Claros V.

Impresión: Carvajal Soluciones de Comunicación.

Impreso en Colombia / *Printed in Colombia.*

Este libro presenta los resultados de la investigación realizada por Juan Carlos Cuéllar Q., profesor investigador de la Facultad de Ingeniería de la Universidad Icesi, bajo la dirección de José Luis Arciniegas, Ph.D., profesor investigador de la Universidad del Cauca, co-dirigido por Jesús Hamilton Ortiz, Ph.D., para la obtención de su título de Doctor en Ingeniería Telemática en la Universidad del Cauca. La publicación de este libro se aprobó luego de superar un proceso de evaluación doble ciego por dos pares expertos. El contenido de esta obra no compromete el pensamiento institucional de la Universidad Icesi o de la Universidad del Cauca, ni les genera responsabilidades legales, civiles, penales o de cualquier otra índole, frente a terceros.



Calle 18 #122-135 (Pance), Cali-Colombia

editorial@icesi.edu.co

www.icesi.edu.co/editorial

Teléfono: +57(2) 555 2334

Juan Carlos Cuéllar Quiñonez

Profesor de tiempo completo del Departamento de Tecnologías de la Información y las Comunicaciones de la Facultad de Ingeniería de la Universidad Icesi (Cali, Colombia). Es Doctor en Ingeniería Telemática de la Universidad del Cauca (Popayán, Colombia), Máster en Ingeniería, con énfasis en Telecomunicaciones de la Universidad Pontificia Bolivariana, (Medellín, Colombia), Especialista en Redes y Telecomunicaciones de la Universidad Icesi, Especialista en Redes y Servicios Telemáticos de la Universidad del Cauca e Ingeniero Electricista de la Universidad del Valle. jcuellar@icesi.edu.co

José Luis Arciniegas Herrera

Profesor de tiempo completo del Departamento de Telemática de la Facultad de Ingeniería Electrónica y Telecomunicaciones de la Universidad del Cauca (Popayán, Colombia). Es Doctor en Ingeniería de la Universidad Politécnica de Madrid (España) e Ingeniero en Electrónica y Telecomunicaciones y Especialista en Redes y Servicios Telemáticos de la Universidad del Cauca. Sus áreas de interés son: servicios y aplicaciones usando televisión digital interactiva, multimedia, arquitecturas de software y calidad de software. jlarci@unicauca.edu.co

Jesús Hamilton Ortiz

Doctor en Tecnología Informática Avanzada de la Universidad de Castilla-La Mancha (España), con un DEA en Sistemas Telemáticos de la Universidad Politécnica de Madrid (España). Es candidato a Doctor en Informática y Telecomunicaciones de la Universidad Autónoma de Madrid. Es fundador de Closemobile R&D, empresa enfocada en telecomunicaciones y temas aeroespaciales. Sus áreas de interés son: las redes Ad hoc, 4G-5G, Internet de las Cosas, VANET, vehículos autónomos e IPv6. jesushamilton.ortiz@closemobile.com

Tabla de contenido

Resumen	15
Introducción	17
Estado del arte	21
Metodología	21
Antecedentes	21
Modelos subjetivos	22
Modelos objetivos	22
Modelos indirectos	25
Trabajos relacionado	26
Componentes para generar los datos y obtener el modelo	31
Parámetros de calidad de servicio seleccionados	31
Modelo objetivo seleccionado para medir la QoE	32
Maqueta de experimentación	37
Diseño y ejecución del experimento	45
Diseño del experimento	45
Identificación de los factores y de sus respectivos rangos	45
Selección de variable de respuesta	46
Elección del diseño experimental	47
Ejecución del experimento	54
Obtención de los rangos de los factores del experimento	55
Análisis estadístico de los datos	57
Acerca de los resultados	70

Análisis de la evaluación del desempeño de los modelos propuestos	73
Ajuste de datos para el análisis comparativo	73
Desempeño: modelos propuestos vs modelos objetivos	79
Desempeño: modelos propuestos vs pruebas subjetivas	85
Pruebas estáticas	93
Acercas de los resultados	95
Conclusiones, lecciones aprendidas y trabajo futuro	99
Conclusiones	99
Lecciones aprendidas	101
Trabajo futuro	101
Referencias	103
Anexo: Configuración de VLC Y NetEM	113

Índice de Tablas

Tabla 1. Comparativo de trabajos relacionados que utilizan modelos indirectos	26
Tabla 2. Parámetros de QoS seleccionados	37
Tabla 3. Especificaciones técnicas de las máquinas de la maqueta de experimentación	38
Tabla 4. Características espacio temporales de los videos usados en la maqueta de experimentación	41
Tabla 5. Rango de los parámetros de QoS configurados en el experimento	46
Tabla 6. Tratamientos para el experimento con los valores codificados y valores reales asociados a cada factor	49
Tabla 7. Datos del modelo que se presenta en la Ecuación 4	51
Tabla 8. Medición de QoE obtenida para el tratamiento No.1	55
Tabla 9. Datos obtenidos para generar el modelo del video entrevista	59
Tabla 10. Análisis de varianza para los datos obtenidos del video entrevista	60
Tabla 11. Datos obtenidos para generar el modelo del video cafetería	63
Tabla 12. Análisis de varianza para los datos obtenidos del video Cafetería	64
Tabla 13. Datos obtenidos para generar el modelo del video Baloncesto	67
Tabla 14. Análisis de varianza para los datos obtenidos del video Baloncesto	68
Tabla 15. Coeficientes de los tres modelos obtenidos	70
Tabla 16. Resultados de la evaluación de los modelos	74

Tabla 17. Valores de las pruebas subjetivas realizadas a los tres videos con su correspondiente valor en escala VQM	76
Tabla 18. Constantes de a,b,c,d y R_cuadrado de la regresión no lineal para cada modelo propuesto	76
Tabla 19. Escalas para mapear MOS vs VQM para cada video	78
Tabla 20. Medidas obtenidas de cada modelo en escala MOS	79
Tabla 21. Resultados de la evaluación con PSNR	80
Tabla 22. Resultados de la evaluación con SSIM	80
Tabla 23. Mapeo de escala MOS a PSNR y SSIM	81
Tabla 24. Datos obtenidos para evaluar el MVMB	82
Tabla 25. Resultados estadísticos para comparar el desempeño del MVMB	83
Tabla 26. Datos obtenidos para evaluar el MVMM	83
Tabla 27. Resultados estadísticos para comparar el desempeño del MVMM	84
Tabla 28. Datos obtenidos para evaluar el MVMA	84
Tabla 29. Resultados estadísticos para comparar el desempeño del MVMA	85
Tabla 30.MVMB: resultados en escala de MOS y de pruebas subjetivas ...	87
Tabla 31. Datos estadísticos de los resultados del MVMB y de las pruebas subjetivas	87
Tabla 32. MVMB: resultados de pruebas de desempeño	88
Tabla 33.MVMM: resultados en escala de MOS y de pruebas subjetivas ..	89
Tabla 34. Datos estadísticos de los resultados del MVMM y de las pruebas subjetivas	90
Tabla 35. MVMM: resultados de pruebas de desempeño	91
Tabla 36.MVMB: resultados en escala de MOS y de pruebas subjetivas ...	91
Tabla 37. Datos estadísticos de los resultados del MVMA y de las pruebas subjetivas	92
Tabla 38. MVMA: resultados de pruebas de desempeño	93
Tabla 39. Resumen de las pruebas de desempeño de los modelos propuestos	95

Índice de Figuras

Figura 1. Encapsulamiento de H.264	33
Figura 2. Maqueta de experimentación implementada	37
Figura 3. Ejemplos de videos de prueba con sus características espacio-temporales	40
Figura 4. Distribución de los valores IE & IT de los videos originales utilizados en el experimento	42
Figura 5. Codificación y su representación geométrica de la porción factorial del DCC (tres factores)	48
Figura 6. Puntos Axiales y factoriales del DCC / Valores asociados a la codificación del eje X, parámetro retardo	48
Figura 7. Proceso para lo obtener los datos para la generación de los tres modelos	57
Figura 8. QoE obtenida para el video Entrevista en escala VQM	58
Figura 9. QoE_Entrevista vs Pérdida de paquetes, Variación del Retardo	61
Figura 10. QoE_Entrevista vs Pérdida de Paquetes, Retardo	61
Figura 11. QoE_Entrevista vs Retardo, Variación del Retardo	61
Figura 12. QoE obtenida para el video Cafetería en escala VQM	62
Figura 13. QoE_Cafetería vs Retardo, Perdida de Paquetes	65
Figura 14. QoE_Cafetería vs Retardo, Variación del Retardo	65
Figura 15. QoE_Cafetería vs Perdida de Paquetes, Variación del Retardo	65

Figura 16. QoE obtenida para el video Entrevista en escala VQM	66
Figura 17. QoE_Baloncesto vs Pérdida de Paquetes, Retardo.	69
Figura 18. QoE_Baloncesto vs Pérdida de paquetes, Variación del retardo	69
Figura 19. QoE_Baloncesto vs Retardo, Variación del retardo.	69
Figura 20. Regresión no lineal correspondiente al MVMB - video Entrevista	77
Figura 21. Regresión no lineal correspondiente al MVMM - video Cafetería	77
Figura 22. Regresión no lineal correspondiente al MVMA - video Baloncesto	77
Figura 23. Interpretación de los valores que entrega el coeficiente de correlación de Pearson	81
Figura 24. Medidas (escala MOS) obtenidas al evaluar los parámetros de QoS del MVMB / Medidas obtenidas al realizar las PS a los 16 tratamientos del video Entrevista	88
Figura 25. Medidas (escala MOS) obtenidas al evaluar los parámetros de QoS del MVMM / Medidas obtenidas al realizar las PS a los 16 tratamientos del video Cafetería	90
Figura 26. Medidas (escala MOS) obtenidas al evaluar los parámetros de QoS del MVMA / Medidas obtenidas al realizar las PS a los 15 tratamientos del video Baloncesto	92
Figura 27. MOS vs PP y VR para el MVMB	94
Figura 28. MOS vs PP y R para el MVMM	94
Figura 29. MOS vs PP y VR para el MVMA	94
Figura 30. Algoritmo usado para obtener el modelo de medición de QoE en IPTV	98

Acrónimos

ANSI	American National Standard Institute
CABAC	Context Adaptative Binary Arithmetic Coder
CAVLC	Context-Adaptative Variable Length Code
CCP	Coefficiente de Correlación de Pearson
CCS	Coefficiente de Correlación de Spearman
DCC	Diseño Central Compuesto
GoP	Group of Pictures
HyQoE	Hybrid Quality of Experience
IPTV	Televisión sobre Protocolo IP
IQX	Exponential Interdependency of Quality of Experience and Quality of Service
JND	Just Noticeable Differences
MOS	Mean Opinion Score
Movie	Motion-based Video Integrity Evaluation
MPQM	Moving Quality Metric
MSE	Mean Squared Error
MSR	Metodología de Superficies de Respuesta
MSSIM	Mean SSIM
MS-SSIM	Multi-scale SSIM

MVMA	Modelo Video Movimiento Alto
MVMB	Modelo Video Movimiento Bajo
MVMM	Modelo Video Movimiento Moderado
NAL	Network Abstraction Layer
NetEm	Network Emulator
P2P	Point to Point
PDM	Perceptual Distortion Metric
PSNR	Peak Signal to Noise Ratio
PVA	Porcentaje de Valores Atípicos
QoE	Quality of Experience
QoS	Quality of Service
RECM	Raíz del Error Cuadrático Medio
RMSE	Root Mean Square Error
RNN	Random Neural Network
SI	Spatial Information
SNMP	Simple Network Management Protocol
SSIM	Structural Similarity
STB	Set-Top-Box
TI	Temporal Information
UIT	Unión Internacional de Comunicaciones
VDP	Visual Differences Predictor
VIIDEO	Video Intrinsic Integrity and Distortion Evaluation Oracle
VQEG	Video Quality Expert Group
VQM	Video Quality Metric
VSNR	Visual Signal to Noise Ratio

Resumen

Con el crecimiento de la demanda de servicios de video en Internet, en especial el servicio de televisión por IP (IPTV), es necesario que sus proveedores implementen esquemas de calidad de servicio (QoS) en sus redes de núcleo para garantizar el funcionamiento adecuado del servicio de que ofrecen a sus usuarios. 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 (QoE) del video transmitido. Generalmente, el proceso que se ha seguido para obtener estos modelos se basa en la realización de pruebas subjetivas que utilizan un número limitado de parámetros de QoS, esta investigación propone algo diferente, un modelo para medir la QoE basado en parámetros de QoS y el uso de un modelo objetivo para la evaluación de la calidad del video.

El objetivo de esta investigación fue proponer un modelo que permita medir la calidad de la experiencia para el servicio de IPTV en tiempo real, en un entorno residencial, y así permitir a los proveedores de servicio el monitoreo continuo de la calidad de la QoE del video y realizar ajustes en tiempo real en la red, para garantizarla.

En su desarrollo, con base en la revisión del estado del arte se identificaron los parámetros de QoS y el método objetivo con el cual se generaría el modelo. Se grabaron tres videos de prueba y utilizando una maqueta de experimentación diseñada para el proyecto, ellos se transmitieron modificando algunos parámetros de QoS (retardo, variación del retardo y pérdida de paquetes. Una vez los videos fueron transmitidos y procesados en el destino, se evaluó su QoE utilizando el método objetivo elegido. Con los valores de los parámetros de QoS y las medidas obtenidas, después de procesar los datos y utilizando

Resumen

metodología de superficies de respuesta, se generó el modelo propuesto para cada video de prueba.

El modelo propuesto permite calcular la QoE asociada al video conociendo los valores de los citados parámetros de QoS del video analizado. Para verificar el desempeño de cada modelo, sus resultados se compararon con los de dos métodos objetivos estandarizados y con los de pruebas subjetivas, la comparación muestra una correlación positiva moderada entre los resultados que entregan los modelos propuestos, las pruebas subjetivas y los métodos objetivos estandarizados.

El modelo propuesto, tiene ventajas respecto de otras propuestas en la medida en que se puede utilizar para la medición de la QoE en línea, sin decodificar el video y sin recurrir a pruebas subjetivas, y no requiere de la alta complejidad computacional propia de otros modelos basados en métodos objetivos.

Introducción

Dado el crecimiento de los servicios de video sobre IP –incluyendo el servicio de Televisión sobre Protocolo IP (IPTV) [1]–, se espera que en 2020 el tráfico del video sobre IP corresponda al 82% del tráfico global de Internet, una cifra incluso conservadora, si se considera que en 2015 llegó a 70% [2]. Éste crecimiento acelerado de tráfico implica que los proveedores de estos servicios deben implementar esquemas de calidad de servicio (*Quality of Service, QoS*) en sus redes de núcleo, para poder garantizar la calidad del producto que ofrecen a sus usuarios.

La implementación de esquemas de QoS no garantiza que la percepción del usuario acerca del servicio sea siempre positiva, ya que existen múltiples factores que pueden afectarla, tales como: el tamaño de la pantalla, el contenido de la programación y tipo de servicio contratado. Por ello, esta percepción ha evolucionado al concepto de Calidad de la Experiencia (*Quality of Experience, QoE*), el mismo que la Unión Internacional de Comunicaciones (UIT), en su Recomendación UIT-T G.1080 [3] define como “la aceptabilidad global de una aplicación o servicio, tal como la percibe subjetivamente el usuario final”. La QoE combina parámetros no técnicos, como la percepción, la experiencia y la expectativa del usuario, y parámetros técnicos de QoS a nivel de red (e.g., retardo, variación del retardo, pérdida de paquetes). Para los proveedores de servicio es importante evaluar esta percepción constantemente, pues les permite mejorar continuamente y mantener satisfechos a sus usuarios [4], [5], por lo tanto, es necesario establecer y analizar la relación que existe entre la QoE percibida por el usuario y los parámetros de QoS a nivel de red, de tal manera que sea posible, en cualquier momento, estimar con base en ellos la QoE percibida por el usuario.

Introducción

En concordancia con lo dicho, la investigación cuyos resultados se reportan en este documento estuvo dirigida a encontrar una respuesta a la pregunta ¿Cómo generar un modelo para medir QoE en línea para el servicio de IPTV? Y para responderla se propuso como objetivo general: “Proponer un modelo que permita medir la QoE en línea –esto es, en tiempo real–, para el servicio de IPTV en un entorno residencial” y como objetivos específicos: identificar y seleccionar los parámetros y las métricas asociados a QoS que permitan evaluar la calidad del servicio de IPTV en un entorno residencial; seleccionar y adaptar un algoritmo sin referencia (*No Reference*) para la medición de la calidad de video que permita obtener una medida de la QoE percibida por el usuario e identificar e implementar una arquitectura de pruebas en laboratorio para obtener los datos necesarios para la generación de un modelo; y generar dicho modelo empleando análisis de regresión multivariable con base en los datos obtenidos de la variación de los parámetros de QoS y la respuesta entregada por el algoritmo sin referencia.

El producto obtenido contribuye a la solución del problema en la medida en que propone un modelo utilizando parámetros de QoS y un algoritmo sin referencia para evaluar la QoE del servicio de IPTV. Tiene algunas ventajas, entre ellas: no requiere de pruebas subjetivas para obtener datos para generar el modelo; no requiere de la decodificación del video; necesita menor ancho de banda para su monitoreo [6]; y permite evaluar la calidad del video en tiempo real, donde usualmente no se tiene la señal de video original como referencia [7].

El modelo propuesto en este documento permitirá a los proveedores de servicios evaluar directamente la QoE en el cliente IPTV residencial y con ello ajustar automáticamente los parámetros de QoS que le permitan ofrecer un servicio adecuados al usuario final. Sus ventajas contrastan con las limitaciones de otros modelos propuestos (ver el estado del arte en el capítulo 2), en unos casos debido a la cantidad de parámetros de QoS incluido en el análisis, en otros, por el uso de pruebas subjetivas.

Luego de esta introducción el texto se divide en cinco capítulos, en ellos se presenta: el estado del arte y la metodología y fuentes de datos usadas para su realización (capítulo 2); los componentes utilizados para generar los datos necesarios en el modelo –incluyendo la identificación de los parámetros de QoS seleccionados, la explicación, tanto del modelo objetivo utilizado para evaluar la QoE, como de las especificaciones técnicas de la maqueta de experimentación, y la descripción de los videos usados en el experimento– (capítulo 3); el diseño y

la ejecución del experimento realizado para la obtención del modelo, incluyendo el análisis estadístico de los datos obtenidos de los modelos generados para cada tipo de video (capítulo 4); los resultados de las pruebas de evaluación de los modelos propuestos, con énfasis en la estandarización de las medidas que entregan los modelos –para que sean comparables– y en los resultados, tanto de las pruebas comparativas de desempeño con modelos objetivos y subjetivos, como de las pruebas estáticas realizadas a cada modelo dirigidas a analizar su comportamiento y estabilidad (capítulo 5); y las conclusiones, trabajo futuro y lecciones aprendidas en este proceso (capítulo 6).

Con base en información recopilada durante la investigación que dió origen a este libro, se publicó el artículo “Clasificación y análisis de métodos para medir calidad de la experiencia del servicio de televisión sobre protocolo IP (IPTV) [9]; se presentaron las ponencias “Modelo híbrido para medir QoE al servicio de IPTV”, en el Octavo Congreso Colombiano de Computación e “Hybrid model to measure QoE in IPTV: Methodological proposal”, en el VII Congreso Iberoamericano de Telemática CITA 2015; y se presentó el poster “Propuesta metodológica para implementar un modelo para medir calidad de la experiencia para el servicio de IPTV”, en la Spring School on Networks.

Estado del arte

Metodología

El estado del arte se construyó a partir de la revisión y el análisis de trabajos relacionados con base en lo presentado en [8] y enfocando los siguientes temas y palabras clave: QoS para IPTV, QoE 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: IEEE Explore, ACM Digital Library, Springer, Elsevier, Science Direct, Scopus, Wiley online Library y el portal web de la Unión Internacional de Telecomunicaciones (UIT).

La revisión se realizó analizando en las publicaciones: el proceso utilizado para obtener los datos para generar los modelos, el tipo de modelo empleado para evaluar la QoE y los parámetros de QoS utilizados para generar el modelo. Para lograr identificar estas tres características se revisaron sesenta artículos, de los cuales veinte abordaban la temática de QoE; doce de ellos fueron 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.

La revisión del estado del arte permitió identificar la metodología utilizada en los diferentes modelos de evaluación la QoE para el servicio de IPTV y otros aspectos útiles para diseñar una propuesta metodológica que permitiera obtener el modelo buscado en la investigación que da origen a este texto.

Antecedentes

Con el fin de monitorear la QoE y garantizar una percepción positiva del usuario sobre el servicio prestado, se ha fortalecido la línea de investigación que busca

medir la QoE al servicio de IPTV. Rahrer, Fiandra y Wright [10] proponen tres tipos de modelos para evaluar la QoE: subjetivos, objetivos e indirectos.

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

MODELOS SUBJETIVOS

En este tipo de modelos se evalúa la QoE utilizando personas que califican la calidad de una secuencia de videos mediante el diligenciamiento de encuestas. Este tipo de experimentos se diseñan siguiendo los lineamientos planteados por la UIT en [11]-[14] y utilizan como escala de medición *Mean Opinion Score* (MOS), la escala de medida más popular para analizar y calificar la calidad del video [15]. Cuando el usuario ve el video lo califica utilizando una escala discreta de cinco niveles: 1-malo (*bad*), 2-pobre o mediocre (*poor*), 3-justo o regular (*fair*), 4-bueno (*good*) y 5-excelente (*excellent*) [16].

Aunque los métodos de este tipo requieren tiempo y son costosos –por la logística necesaria para su implementación–, son de gran utilidad y representan un estándar de facto para la validación de modelos propuestos. Sin embargo, no es fácil automatizar esta medida, ya que se deben tener en cuenta: la influencia de los factores psicológicos humanos y el sesgo del usuario al momento de calificar cada video. En [15] se presenta la metodología para realizar este tipo de pruebas, incluyendo la elección del ambiente de pruebas, la selección de videos y el análisis de los datos obtenidos.

Este tipo de modelos no se puede utilizar para analizar la QoE en contenidos de video en plataformas en Internet, debido a que las pruebas se deben hacer en un espacio diseñado para tal fin y porque la elección de los videos a utilizar en la prueba subjetiva es de vital importancia.

MODELOS OBJETIVOS

Con el fin de obtener una medida confiable de QoE –y así evitar las pruebas subjetivas–, se han desarrollado los modelos objetivos, los cuales estiman una medida de QoE en función de algunos parámetros de QoS y de características asociadas al video. Su enfoque no permite que la medida se pueda realizar en tiempo real, pues requieren de una señal de referencia para comparar los

resultados del video transmitido con el video recibido. Solo en los métodos basados en paquetes se pueden hacer mediciones en línea, aunque en la actualidad este tipo de modelos basados en métodos indirectos está en desarrollo. Los modelos objetivos se clasifican de acuerdo con el tipo de análisis o de la extracción de información realizada a la señal de video en: métodos con señal de referencia, métodos basado en datos, métodos basados en imágenes y métodos basados en paquetes.

Métodos con señal de referencia

Su base es la utilización o no de la señal original del video como referencia para ejecutar las mediciones de QoE y se clasifican en: modelo de referencia completa (*Full Reference*), modelo de referencia reducida (*Reduced Reference*) y modelo sin referencia (*No Reference*). En el modelo referencia completa, la señal de video original está disponible en el receptor para compararla con la señal del video recibida y así determinar la calidad del video enviado al usuario; su limitación o desventaja es que no evalúa la calidad absoluta y es costoso en tiempo, porque requiere un elevado nivel de procesamiento para ejecutar algoritmos complejos [17], [18]. En el modelo de referencia reducida la información parcial de la señal de video original transmitida está disponible en el receptor, con el fin de determinar la calidad del video recibido; este modelo es más económico en tiempo y procesamiento que el de referencia completa [17]. Por último, en el modelo sin referencia solo se posee la señal de video recibida para determinar la calidad del video, por lo que la calidad se mide con base en parámetros que reflejan la degradación de la señal (e.g., pérdida de cuadros I, P, B, cambio del bitrate). Es muy útil cuando se desean hacer mediciones en tiempo real [17].

Métodos basados en datos

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

Métodos basados en imagen

Estos métodos analizan la información visual que contiene el video y presentan dos enfoques para evaluarlo: el de modelado de la visión y el de ingeniería. El

Estado del arte

enfoque de modelado de la visión se conoce también como enfoque psicológico [21], debido a que se basa en el análisis de varios componentes del sistema visual humano (i.e., contraste, percepción de color y enmascaramiento temporal y espacial). Los algoritmos basados en este enfoque presentan buen rendimiento ya que los modelos imitan el proceso de formación de la visión humana, sin embargo, debido a la complejidad de sus algoritmos, su procesamiento computacional es elevado [22], entre los algoritmos que trabajan bajo este esquema se encuentran: *Visual Differences Predictor* (VDP) [23], Sarnoff JND (*Just Noticeable Differences*) [24], *Moving Quality Metric* (MPQM) [25] y *Perceptual Distortion Metric* (PDM) [26]. El enfoque de ingeniería, por su parte, se basa en extraer los patrones de distorsión del video debidos a compresión y codificación. Entre los algoritmos más utilizados en este enfoque 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] –uno de los modelos más utilizados para evaluar la QoE del video, adoptado por el American National Standard Institute (ANSI TI.801.03-2003) e incluido en las recomendaciones ITU-TJ.144 e ITU-R BT.1883 [34].

Los métodos basados en imagen presentan una desventaja: requieren para su análisis tanto de la señal fuente, como de la señal recibida en el lado del cliente, lo que hace necesaria la 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 [22], aunque debido a que están estandarizados ya se encuentran implementados en algunas herramientas comerciales.

Métodos basados en paquetes

Trabajan con la información que se obtiene de los encabezados de nivel 3 y 4 de los paquetes y con parámetros como: retardo y variación del retardo. Estos métodos se clasifican en modelos de nivel de red, modelos paramétricos de nivel de paquete, modelos paramétricos de planeación, modelos de nivel de flujo de bits y modelos híbridos [4].

Los modelos de nivel de red permiten entregar una medición de QoE del video analizando información obtenida del video a nivel de red [35], [36]; los modelos paramétricos de nivel paquete entregan la medición de QoE del

video mediante información que extraen de los encabezados de los paquetes del video [7]; los modelos paramétricos de planeación requieren de parámetros de planeación para el cálculo de la QoE asociada al video [37]; los 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 [38]; y los modelos híbridos basan su funcionamiento en combinar cualquiera de los modelos revisados anteriormente para obtener una medición de QoE del video [39].

MODELOS INDIRECTOS

Los modelos indirectos buscan evaluar la QoE del video mediante el uso de un modelo matemático que 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, sin necesidad de recurrir a pruebas subjetivas.

Con el análisis de los tres tipos de modelos presentado en las secciones anteriores, es claro que cada uno 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 su realización con video en vivo no es posible. Los modelos objetivos, por su parte, no reflejan la percepción del usuario, debido a que utilizan un número limitado de variables, además, algunos son computacionalmente muy complejos en su cálculo.

Ante las desventajas que presentan los modelos subjetivos y los objetivos, una solución para calcular la QoE percibida por un usuario es la utilización de modelos indirectos. Por esta razón múltiples autores –los que se presentan en la sección “Trabajos relacionados” (capítulo 2)–, han planteado una serie de modelos matemáticos para su resolución. 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 QoE.

Con base en los modelos analizados, en la siguiente sección se presenta el resultado de la revisión sistemática realizada a 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 precisar el proceso que siguió cada autor para obtenerlo.

Trabajos relacionados

En esta sección se analizan varias propuestas basadas en modelos indirectos, identificando el procedimiento que cada autor siguió para obtener un nuevo modelo para evaluar la QoE. Los métodos indirectos relacionan parámetros de QoS, tales como: retardo, variación del retardo y pérdida de paquetes, con un modelo subjetivo u objetivo que mida la QoE asociada al video. Esta relación del valor que pueden tomar los parámetros de QoS versus el valor obtenido de la QoE se expresa mediante un modelo matemático o un modelo de correlación.

Después analizar los trabajos relacionados se identificaron tres características que permiten su comparación: el proceso utilizado para obtener los datos que permiten generar el modelo, el tipo de modelo utilizado para evaluar la QoE y los parámetros de QoS utilizados para generar el modelo. La primera, identifica el tipo de experimento que se realizó para obtener dichos datos, es decir, si se utilizaron pruebas subjetivas, una herramienta de simulación o una maqueta de experimentación en laboratorio; la segunda muestra si para evaluar la QoE se utilizó un modelo subjetivo u objetivo, una referencia completa o

Tabla 1. Comparativo de trabajos relacionados que utilizan modelos indirectos

Autores	Proceso para la obtención de datos	Modelo de evaluación de la QoE	Parámetros de QoS
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 definido	Ambiente de computación invasiva	Pérdida de paquetes / nivel de ráfaga / variación del retardo / retardo de paquetes / ancho de banda
Yamagishi et al. [48]	Plataforma de experimentación en laboratorio	Pruebas subjetivas	Pérdida de paquetes
Riker et al. [49]	Red neuronal aleatoria	Red entrenada con videos evaluados subjetivamente	Ninguno
Abdeljao 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 / pérdida de paquetes

sin referencia, etc.; y la tercera precisa cuáles fueron los parámetros de QoS utilizados para generar el modelo.

Hoßfeld et al. [40] y Fiedler et al. [41] proponen que la relación entre la QoS y la QoE está dada por la hipótesis IQX (*Exponential Interdependency of Quality of Experience and Quality of Service*). Esta hipótesis, mediante pruebas subjetivas, analiza cómo se afecta la calificación del MOS ante la variación del parámetro pérdida de paquetes. En los modelos que presentan Fiedler y Hoßfeld [42] y Hoßfeld et al. [43] se involucran tres parámetros de QoS –medidas de falla, medidas exitosas y medidas de recursos–, para analizar con mayor precisión la sensibilidad del usuario, con base en pruebas subjetivas realizadas con un conjunto de usuarios. La limitación de este modelo, con su posterior mejora, radica en la necesidad de realizar pruebas subjetivas para obtener los datos que permiten generarlo, ya que, como se ha mencionado, estas pruebas son costosas en tiempo y logística, debido a la necesidad de realizar un número proporcional de pruebas a los usuarios si se desea obtener una muestra representativa.

Kim et al. [44]-[46] evaluaron cinco parámetros de QoS –pérdida de paquetes, nivel de ráfaga, variación del retardo, retardo y ancho de banda– y a cada uno de ellos se le asigna un peso relacionado con su grado de importancia relativa, el cual se calcula mediante los experimentos presentados en [47]. Los autores emplean el valor de los parámetros de QoS en una ecuación lineal para obtener el valor de QoS asociado al video, el cual posteriormente es evaluado en una ecuación exponencial con el fin de obtener un valor de QoE que involucra tres parámetros adicionales: el tamaño de la pantalla, la 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 Pictures*, GoP). Kim et al. [44]-[46] no especifican cómo se obtiene la ecuación exponencial y utilizan pruebas subjetivas para asignar a cada parámetro de QoS el peso relacionado con su grado de importancia.

Yamagishi, Kawano y Hayashi [48] 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 ocasionada por 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 video; el modelo se basa en pruebas subjetivas realizadas a un conjunto de videos tomados como ejemplo, y analiza

la QoE con un único parámetro de QoS: la pérdida de paquetes. Al igual que los modelos presentados en los párrafos anteriores, este modelo requiere de pruebas subjetivas para obtener los datos.

Riker et al. [49] proponen un esquema de predicción híbrido (*Hybrid Quality of Experience*, HyQoE) que evalúa el nivel de calidad del video en tiempo real con base 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 usando datos obtenidos mediante pruebas subjetivas.

Abdeljaouad, Kandavanam, y Karmouch [50] presentan un modelo basado 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 del criterio descendente de la función sigmoide, donde la respuesta de la función disminuye mientras el valor de entrada utilizado en dicha función 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 la 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 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 QoE percibida por el usuario.

Aroussi, Bouabana-Tebibel y Mellouk [53] proponen cuantificar la QoE utilizando varios parámetros de QoS, con base en la hipótesis IQX propuesta en [40] y [41] y se utiliza una regresión lineal entre las diferentes ecuaciones generadas por cada parámetro de QoS para obtener el 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 tienen como gran inconveniente la necesidad de realizar pruebas subjetivas, las que, como se mencionó, son costosas y requieren de una logística compleja. Por otra parte, las propuestas que se presentan en [40]-[43] 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 analizar lo anterior –y ante el costo en tiempo y logística que implica realizar pruebas subjetivas–, nuestra propuesta se basa en la elección de un modelo objetivo para evaluar la QoE y la selección de tres parámetros de QoS para obtener los datos su generación. La justificación sobre la elección del modelo objetivo y los parámetros de QoS elegidos se presenta en el capítulo 3.

Las propuestas de Elhoseny, Shehab y Osman [54] Shehab, El Aziz y Hassanien [55] y Shehab, Elhoseny y Hassanien [56] plantean esquemas eficientes de entrega de video P2P (*Point to Point*) con base en el análisis del comportamiento del usuario. Estas propuestas son novedosas porque permiten el análisis del comportamiento del usuario –y con ello la mejora de su 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 las propuestas de modelos de QoE se enfocan en tecnologías móviles, como Wi-Fi y 5G, debido al alto crecimiento (a una tasa esperada de 46% entre 2015 y 2020) de este tipo de usuarios [2]. Al enfocarse en este tipo de tecnologías, los parámetros para analizar la QoE asociada al video cambian, como presentan Lentisco et al. [57], quienes tienen en cuenta aspectos como: el retardo de reproducción del video, el congelamiento del video después de su reproducción y el cambio de resolución del video en su reproducción.

Es importante mencionar que el planteamiento de Wang, Pervez y Zouen [34] es similar al que se presentará en este documento, con una gran diferencia, ellos plantean un modelo utilizando solo dos parámetros de QoS –variación del retardo y pérdida de paquetes–; dejan fijo el retardo a 100mseg; utilizan videos de calidad estándar con resolución 480x270, para la pruebas; presentan un único modelo para un video con movimiento medio; y no presentan pruebas de validación estadística del modelo propuesto ni comparan sus resultados con los de modelos objetivos estandarizados.

El modelo que se propone en este libro mejora la propuesta de Wang et al. [34] y las de los autores incluidos en la TABLA 1 [40], [41], [44]-[46], [48]-[50], [53] en los siguientes aspectos: se utilizan videos en alta definición, con resolución 1080x1920, para obtener los datos para generar el modelo; se trabaja con tres parámetros de QoS simultáneamente –retardo, variación del retardo y pérdida de paquetes–; y al modelo obtenido se le realizan pruebas estadísticas y sus resultados se comparan con modelos objetivos estandarizados.

Componentes usados para generar los datos y obtener el modelo

Parámetros de calidad de servicio seleccionados

Después de analizar los modelos propuestos en el capítulo anterior, se eligieron como parámetros de QoS: retardo (*delay*), variación del retardo (*jitter*) y pérdida de paquetes (*packet loss*), por ser los de mayor influencia en la calidad del video [17].

Retardo

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 superiores a lo previsto en la Recomendación UIT-T Y.1541 –que establece un rango de 0 a 400 mseg– [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.

Variación del retardo

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 Y.1541 –que establece un rango de 0 a 50mseg– [59], hace que el video se congele y pierda definición [60]. Una solución a este problema es configurar un tamaño de memoria intermedia (*buffer*) adecuado, sin embargo, esto puede afectar el retardo.

Componentes usados para generar los datos y obtener el modelo

Pérdida de paquetes

Definido como *IP Packet Loss Ratio*, hace referencia a la tasa de pérdida de paquetes y se obtiene de la relación entre el total de paquetes perdidos y el total de paquetes transmitidos [58]. Es el parámetro de mayor influencia en la calidad del video [60], cuando supera el rango de 0 a 1×10^{-3} causa congelamiento y saltos en el video, debido a que si hay pérdida de cuadros (*frames*), ellos se reemplazan con cuadros previamente decodificados, lo que deteriora la calidad del video.

De acuerdo con 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 que llegue al usuario se congelará y generará un salto en la imagen, lo que se puede compensar retransmitiendo el cuadro perdido, pero ello hará que se incremente el retardo y la variación del retardo; la variación del retardo, por su parte, afecta la reproducción del video al congelar la imagen durante la espera de la llegada de cuadro tardío, eligiendo un tamaño adecuado de la memoria intermedia, 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].

Modelo objetivo seleccionado para medir la QoE

Como se presentó en la sección “Modelos objetivos” (capítulo 2), un modelo sin referencia es muy útil para realizar mediciones en tiempo real, ya que no requiere de la señal de video origen para el cálculo de la QoE.

En la Recomendación UIT-T P.1202.2 [61] se presenta un modelo sin referencia para evaluar la QoE del video y se especifican los procedimientos que permiten monitorear la calidad del video sobre IP, con base en la información de los encabezados de los paquetes del flujo de bits (*bitstream*) del video. Esta Recomendación puede aplicarse para servicios como IPTV y tiene 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 cuenta con cuatro bloques funcionales: Extracción de parámetros básicos; Agregación de parámetros básicos a nivel interno de

cuadro; Agregación de parámetros dentro del modelo de parámetro; y Modelo de estimación de la calidad.

Con base en el análisis de la Recomendación P.1202.2 se decidió su implementación, ya que se ajusta a lo previsto en el objetivo específico: “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”. Adicionalmente, se tuvo en cuenta que la información que requiere cada bloque funcional de dicha recomendación se puede obtener analizando los encabezados de los paquetes del flujo de bits del video, como se explica en ella.

Durante la investigación se inició la implementación del primer bloque “Extracción de parámetros básicos”, para lo cual, en la maqueta de experimentación que se diseñó e implementó, se transmitió una serie de videos utilizando VLC [62] y se capturó 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 1 se aprecia la carga útil de un paquete H.264 hasta el nivel de macrobloque, realizar esta decodificación hasta ese nivel permite obtener información útil para la implementación del bloque de extracción de parámetros de la Recomendación UIT-T P.1202.2.

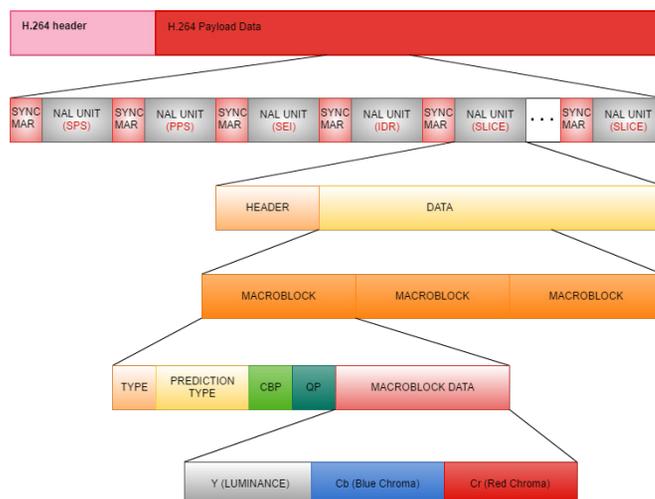


Figura 1. Encapsulamiento de H.264

Componentes usados para generar los datos y obtener el modelo

A nivel de abstracción de red (*Network Abstraction Layer*, NAL), se encontró que existen parámetros requeridos por la Recomendación que estaban codificados entrópicamente, según CABAC (*Context Adaptive Binary Arithmetic Coder*) [64] y CAVLC (*Context-Adaptive 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 están 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.

Al llegar a este nivel, para extraer los parámetros que implica usar la Recomendación UIT-T P.1202.2, se debía identificar cómo el codificador genera la codificación entrópica. Para ello, fue necesario analizar el funcionamiento del codificador a nivel de bit y así 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, dificultad técnica que hacía que el cálculo de la QoE no fuera confiable, por no tener certeza de la información obtenida en la implementación de este módulo.

Por no lograr decodificar esta información, no fue posible implementar la estructura de datos que requería la Recomendación P1202.2 a nivel de cuadro, la que luego se usaría como entrada para el segundo bloque funcional "Agregación de parámetros básicos a nivel interno de cuadro".

Además de no ser viable la implementación de la Recomendación UIT-T P.1202.2, no era factible técnicamente diseñar e implementar un codificador propio, debido a que los datos para proponer el modelo se iban a realizar utilizando videos codificados en H.264, un estándar de amplio uso para la codificación de video.

Ante la serie de dificultades técnicas encontradas en la implementación de la Recomendación UIT-T P.1202.2, se realizó una nueva revisión del estado del arte con el fin de encontrar referentes que estuvieran trabajando específicamente en métodos sin referencia. Producto de esta revisión sistemática se encontró el modelo VIIDEO (*Video Intrinsic Integrity and Distortion Evaluation Oracle*), propuesto por Mittal, Saad y Bovik [67], el cual incorpora modelos de estadísticas intrínsecas que permiten cuantificar las distorsiones de los videos (e.g., congelamiento, baja definición) y 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] se utilizó LIVE Video Quality Assessment Database [68], base en la que los videos ya presentaban: distorsiones debido al proceso de codificación con MPEG y H.264; distorsiones reflejadas en errores causados en la transmisión del video sobre IP; y distorsiones producto de la transmisión del video sobre redes inalámbricas, pero no tenía documentado el proceso que se realizó para distorsionar los videos.
- Se intentó modificar el algoritmo para adecuarlo a la resolución que se utilizaría en las pruebas (1920x1080), 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, el cual presenta dos inconvenientes técnicos: primero, el tamaño del archivo a analizar es del orden de 1 a 2 GB –lo que hace que el tiempo de procesamiento sea muy elevado, de alrededor de 40 minutos por video–, y que el procesamiento en los experimentos tome mucho tiempo; y segundo, que la herramienta empleada (ffmpeg) no puede realizar la conversión en algunas ocasiones –cuando un video distorsionado se intenta convertir a YUV–, porque al perderse cuadros del video no puede codificar de nuevo en otro formato y entonces considera que el archivo origen esta defectuoso y no lo procesa.
- VIIDEO [67] no documenta los rangos de las medidas obtenidas de QoE ni se presenta un mapeo con la escala MOS, lo cual es fundamental para analizar el desempeño de VIIDEO y compararlo con modelos objetivos estandarizados.

Por las consideraciones anteriores y dado que los modelos sin referencia están en fase de desarrollo y hasta el momento no hay modelos que se utilicen de manera estandarizada, se decidió identificar un modelo objetivo para realizar las medidas de QoE asociadas al video. Para ello se identificaron los algoritmos: *Peak-Signal-to-Noise-Ratio* (PSNR), *Structural Similarity Index* (SSIM), *Multi-Scale*

Componentes usados para generar los datos y obtener el modelo

SSIM (MS-SSIM), *Video Integrity Evaluation* (MOVIE), *Visual Signal to Noise Ratio* (VSNR) y *Video Quality Metric* (VQM), con el fin de elegir de ellos el que mejor se ajuste a las necesidades de la investigación.

Teniendo en cuenta las cuatro métricas de comparación definidas por el Grupo de Expertos de Calidad del Video (*Video Quality Expert Group*, VQEG), esto es: Coeficiente de Correlación de Pearson (CCP); Raíz del Error Cuadrático Medio (RECM); Coeficiente de Correlación de Spearman (CCS); y Porcentaje de Valores Atípicos (PVA) y luego de analizar lo presentado en [69] y [15], se eligió VQM como modelo objetivo para realizar las mediciones de QoE asociado al video, porque muestra los valores más elevados de CCP y CCS y los valores RECM y PVA más bajos con respecto a los otros algoritmos analizados [69]. Las métricas de comparación citadas se explican en la sección “Pruebas para comparar el desempeño de los modelos propuestos” (capítulo 5).

Adicionalmente VQM trabaja con más parámetros de medición que los otros algoritmos –siete, como se muestra en la TABLA 2–; es un modelo estandarizado que utiliza bajo ancho de banda; extrae datos de los videos origen y destino; y fue adoptado por el American National Standard Institute (ANSI TI.801.03-2003) e incluido en las Recomendaciones UIT-T J.144 y UIT-R BT.1883 [34]. Adicionalmente, VQM es un método utilizado ampliamente en la comunidad científica para la comparación y validación de nuevos modelos propuestos.

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

Como se indicó, VQM maneja siete parámetros independientes: cuatro extraen gradientes espaciales del componente de luminancia Y; dos están basados en características del vector formado por dos componentes de crominancia (CB y CR); y uno se obtiene del producto de características que miden contraste y movimiento que son extraídas del componente de luminancia Y. La descripción de los siete parámetros se presenta en la TABLA 2. Para profundizar en el funcionamiento de VQM se recomienda consultar [32], [70]. Estos siete parámetros se combinan en la ECUACIÓN 1, ella permite obtener la medida de la calidad del video. En su resultado, una medida cercana a cero indica un video de mejor calidad, mientras que una medida cercana o mayor a uno, lo contrario.

Tabla 2. Parámetros de QoS seleccionados

Parámetro	Descripción
Si_{loss}	Detecta un la disminución o pérdida de la información especial, como por ejemplo la definición defectuosa (<i>blurring</i>)
Hv_{loss}	Detecta un desplazamiento de los bordes de orientación horizontal y vertical hacia la orientación vertical, lo que podría ocurrir si los bordes horizontales y verticales sufren una definición defectuosa mayor que la de los bordes diagonales
Hv_{gain}	Detecta un desplazamiento de los bordes diagonales hacia bordes horizontales y verticales, como cuando la imagen presenta falta de definición en sus bordes y se ve defectuosa
$Chroma_{spread}$	Detecta cambios en la dispersión de la distribución de muestras de color de dos dimensiones
Si_{gain}	Determina el aumento en la calidad del video, midiendo la ganancia de calidad del video mediante el análisis de los bordes de la imagen
$Ci_{ati\ gain}$	Es el producto de la característica de contraste, se calcula midiendo la cantidad de detalle espacial y la característica de información temporal
$Chroma_{extreme}$	Detecta errores severos en el color de la imagen causados por errores en la transmisión.

$$VQM = -0.2097 * Si_{loss} + 0.5969 * Hv_{loss} + 0.2483 * Hv_{gain} + 0.0192 * Chroma_{spread} - 2.3416 * Si_{gain} + 0.0431 * Ci_{ati\ gain} + 0.0076 * Chroma_{extreme} \quad (1)$$

Maqueta de experimentación

En la FIGURA 2 se presenta la maqueta de experimentación implementada, con ella se cumple con la segunda parte del segundo objetivo específico: “identificar e implementar una arquitectura de pruebas en laboratorio, con el fin de obtener los datos necesarios para la generación del modelo”.

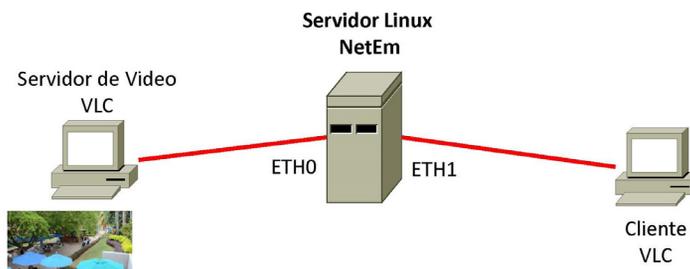


Figura 2. Maqueta de experimentación implementada

Componentes usados para generar los datos y obtener el modelo

La maqueta 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 (ver configuraciones en la TABLA 3).

Tabla 3. Especificaciones técnicas de las máquinas de la maqueta de experimentación

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

La selección de esta maqueta toma como base, además de las lecciones aprendidas de los trabajos relacionados en la revisión de estado del arte (ver capítulo 2), las investigaciones de Lu y Lu [72], Cruvinel y Vazao [73] y Kim y Choi [74]. Se decidió además utilizar todo el ancho de banda de los enlaces Ethernet, para así minimizar las variables no controladas en el experimento (e.g., carga del enlace, congestión, bitrate del video), por esta razón no hay aplicaciones adicionales compartiendo el ancho de banda en la maqueta.

El servicio de IPTV emulado es un servicio no gestionado que va sobre protocolo UDP, sin configurar QoS y sin asignar ancho de banda. Para emular un enlace de red de área amplia entre cliente y servidor se eligió la herramienta NetEm [71], emulador que permite variar los parámetros de QoS escogidos y adicionalmente, como se plantea en [75], 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. Los rangos de los parámetros de QoS utilizados fueron los definidos en la Recomendación UIT-T Y.1541 [59], esto es: retardo, 0 a 400mseg; variación del retardo, 0 a 50mseg; y pérdida de paquetes, 0 a 1×10^{-3} . Un detalle de estos parámetros se puede apreciar en la sección “Videos fuente” (capítulo 3).

Para la selección de los videos fuente se analizaron cuidadosamente dos aspectos: su contenido y sus características espacio-temporales. Respecto del contenido, cuidando que no fuera ofensivo, controversial o capaz de causar distracción en el observador [75]; en el aspecto técnico, se debían elegir las secuencias de video analizando sus características espacio-temporales, las que permiten identificar la cantidad de movimiento y el nivel de detalle que presenta el video en los cuadros. La Recomendación UIT-T P.910 [14] define dichas características, así:

- IE o Información perceptual Espacial (*Spatial Information*, SI), medida que indica la cantidad de detalle espacial de la imagen, cuyo valor, por lo general es más alto para escenas complejas, es decir, para aquellas con mayor cantidad de objetos o bordes en los cuadros (*frames*) de una secuencia; e
- IT o Información perceptual Temporal (*Temporal Information*, TI), medida que indica la cantidad de cambios temporales de una secuencia de video, por lo general su valor es mayor para las secuencias de alto movimiento.

Para calcular el valor de la 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, se mide la desviación estándar del cuadro resultante y luego se selecciona la desviación estándar máxima. En la ECUACIÓN 2 se puede apreciar la expresión que permite su cálculo.

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

El operador de Sobel calcula el gradiente de cada píxel en función de su intensidad y 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 mayor contenido.

El cálculo de la 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

Componentes usados para generar los datos y obtener el modelo

sucesivos. Por lo general indica la cantidad de cambios temporales de una secuencia de video y por ello, suele tener valores más altos para secuencias que representan movimiento elevado. Al igual que con la IE, 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 aprecia la expresión para calcular IT.

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

En las pruebas presentes en la Recomendación UIT-T P.910 se utilizó una serie de videos identificados con letras de la “a” a la “x” (ver FIGURA 3), que tenían diferentes características espacio-temporales. Dependiendo de estas características se puede definir qué tanto movimiento presenta el video.

A manera de ejemplo, en dicha recomendación se presenta la ubicación espacio-temporal para cada uno de estos videos de prueba. En la FIGURA 3, en el eje unidad de IT, en $IT < 20$ se encuentran videos con escenas con movimiento limitado, como l, f y a, mientras en la parte de arriba del eje, $IT > 50$, se encuentran videos con alto movimiento, como p, q e i. Por su parte, en el eje unidades de IE, en $IE < 50$ se encuentran videos con escenas con mínimo detalle espacial, como l, k, x, u y f, mientras en la parte derecha,

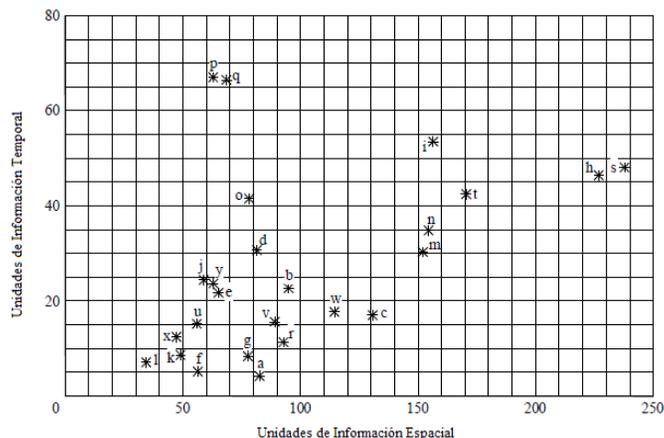


Figura 3. Ejemplos de videos de prueba con sus características espacio-temporales [14]

con $IE > 200$, se encuentran videos con escenas que tienen bastante detalle espacial, como h y s.

Este tipo de representación de los videos mediante una gráfica IT vs IE (FIGURA 3) permite, de manera eficaz y rápida, deducir la cantidad de movimiento y la definición de un video.

Una vez definidas las características espacio-temporales, se procedió a obtener las secuencias de video que se utilizarían. Para empezar, se realizó una exploración en las diferentes bases de datos de videos, en ellas, se encontraron las siguientes dificultades:

- la gran mayoría solo tenía secuencias en definición estándar y se requerían de alta definición;
- las secuencias de video tenían una duración promedio de 10 a 12 segundos, mientras 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; y
- en algunos casos, la calidad de los videos ya había sido afectada, tenían distorsiones de diferente tipo y sus autores no describían cómo se había realizado ese proceso.

Ante esto, se decidió generar 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 –o cuadros– por segundo (*frame*

Tabla 4. Características espacio-temporales de los videos utilizados en la maqueta de experimentación

Video	Imagen	IE	IT	Clase
Entrevista (I00)		35,9166	5,38516	Movimiento bajo
Cafetería (C00)		65,9848	6,48074	Movimiento moderado
Baloncesto (B00)		57,4195	17,4069	Movimiento alto

Componentes usados para generar los datos y obtener el modelo

rate - 23.976 fps) y codificado en H.264/MPEG-4 en formato MOV, y cada toma se realizó por más de treinta segundos. En la TABLA 4 se pueden apreciar los tipos de videos elegidos para la recolección de datos.

El video “Entrevista”, como su nombre lo indica, corresponde a una entrevista que se realiza a una persona. Aquí el movimiento es mínimo ya que la persona solo contesta las preguntas y mira a la cámara, mientras que sus manos y cuerpo se mantienen sin movimiento. El video “Cafetería” muestra a personas que van caminando por debajo de los árboles, algunas se detienen a conversar o están usando las mesas, la cafetería de la imagen está ubicada en la primera planta y el video se tomó desde la planta superior. El vídeo “Baloncesto” corresponde a una sesión de entrenamiento de este deporte, en él, se grabaron jugadores corriendo a lo largo de la cancha.

Estos tres tipos de videos corresponden a la descripción que se presenta en el ANEXO 1, cuadro A.1 de la Recomendación UIT-T P.910, así: Entrevista, categoría A; Cafetería, categorías B y C; y Baloncesto, categorías D y E. De esta manera, los videos definidos para el experimento se ajustan a las características espacio-temporales de cada clase descrita en dicha recomendación.

En la FIGURA 4 se presenta la distribución espacio-temporal de los tres videos utilizados como videos fuente en el experimento, los tres se utilizaron sin ningún tipo de distorsión. En el video “Entrevista”, 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, por su parte, indica

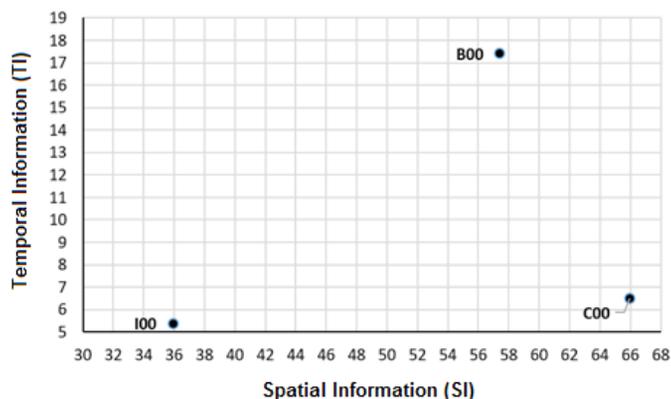


Figura 4. Distribución de los valores IE e IT de los videos originales utilizados en el experimento

un cambio temporal mínimo, es decir que el movimiento en cada cuadro es mínimo o muy bajo. Para el video “Cafetería”, la característica $IE=65,98$ indica un nivel de detalle espacial más elevado que para el video anterior, corresponde a una imagen compleja porque hay más objetos en cada cuadro –árboles, mesas, sombrillas, personas, sillas, quiosco y plantas–; por su parte, 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”, la característica $IE=57,4195$ indica un detalle espacial más bajo que para el video “Cafetería”, es decir que es una imagen con menor complejidad, lo que se explica en el menor número de objetos presentes en cada cuadro; la característica $IT=17,4069$, en cambio, indica un cambio temporal elevado, es decir que en cada cuadro hay movimiento, lo que se explica con la presencia de personas corriendo a lo largo de la cancha.

Los videos fueron elegidos para la toma de datos en los experimentos debido a que su información espacio-temporal difiere y permite entonces identificar cada video dependiendo de la cantidad de movimiento presente –bajo, moderado y alto–. La elección de este tipo de videos se realizó con base en el análisis y las sugerencias procedimentales planteadas en [68], [77]-[82].

Diseño y ejecución del experimento

La obtención de los datos necesarios para la generación del modelo inició con la planeación y el diseño del experimento, según Montgomery [83], los pasos a seguir para el diseño son:

- identificación de los factores con sus respectivos rangos;
- selección de variable de respuesta;
- elección del diseño experimental;
- ejecución del experimento;
- análisis estadístico de los datos; y
- conclusiones.

Para efectos de la presentación del capítulo, los primeros tres se incluyen en el apartado “Diseño del experimento” y los demás en “Ejecución del experimento”.

Diseño del experimento

IDENTIFICACIÓN DE LOS FACTORES Y DE SUS RESPECTIVOS RANGOS

Los factores del diseño de un experimento hacen referencia a las variables independientes que puede variar dentro de un rango que se define previamente [83]. Para este caso, los factores son los parámetros de QoS seleccionados y los rangos establecidos para su variación son los que se presentan en la TABLA 5.

La sección “Obtención de los rangos de los factores del experimento” (capítulo 4) explica ese proceso y precisa las dificultades que se presentaron en

Tabla 5. Rango de los parámetros de QoS configurados en el experimento

Parámetro de QoS	Rango
Retardo (R)	0 -1300 mseg
Variación del retardo (V_R)	0 – 0.018 mseg
Pérdida de paquetes (P_p)	0 – 0.18 %

él. 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 coordina el experimento [83]. El factor no controlable fue el proceso de pérdida de paquetes que utiliza la herramienta NetEm, pues ella trabaja la pérdida de paquetes de manera aleatoria y los paquetes solo se descartan cuando se llena la memoria intermedia (*buffer*) que usa la herramienta. 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, por lo que todo el ancho de banda está dedicado a la transmisión del video, con ello que se descarta que algún tipo de aplicación o tráfico consuma el ancho de banda y pueda afectar la transmisión del video, lo cual sería un factor adicional a tener en cuenta; y
- el ancho de banda entre cliente y servidor es de 100 Mbps, buscando acotar al máximo los factores que pudieran afectar los parámetros de QoS definidos, dado que el modelo en construcción es para el servicio de IPTV residencial, un escenario donde los anchos de banda son menores y hay más aplicaciones o servicios que el usuario final consume, el ancho de banda fijo permite analizar cómo la variación de los parámetros de QoS afecta directamente la calidad del video.

SELECCIÓN DE VARIABLE DE RESPUESTA

La variable de respuesta es la QoE asociada al video y en el experimento será medida con el algoritmo VQM utilizando la herramienta MSU Video Quality Measurement Tool, versión 4.4 Professional [84]. En la sección “Ejecución del experimento” se explica el proceso realizado para utilizar esta herramienta y lograr la medida obtenida mediante VQM.

ELECCIÓN DEL DISEÑO EXPERIMENTAL

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

Identificar el diseño experimental depende del objetivo que se busca, en este caso: “proponer un modelo que evalúe la calidad de la experiencia en línea para el servicio de IPTV en un entorno residencial”. Dado este objetivo, deben existir valores de los factores –los parámetros de QoS seleccionados– que permiten obtener la mejor respuesta –la QoE asociada al video–. Es decir, 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 que se encuentren dentro del rango definido para los factores utilizados en el diseño experimental.

De acuerdo con lo anterior, el tipo de diseño experimental que se ajusta a esos requerimientos es la Metodología de Superficies de Respuesta (MSR), la cual se define como un conjunto de técnicas matemáticas y estadísticas utilizadas en el tratamiento de problemas en los que una respuesta de interés –la QoE del video, en este caso– está influenciada por varios factores de carácter cuantitativo –los parámetros de QoS–, y donde el objetivo es optimizar la respuesta de interés [83]. El propósito inicial de esta metodología es diseñar un experimento bajo condiciones controladas que proporcione valores de la variable respuesta, y a continuación determinar el modelo matemático que mejor se ajuste a los datos obtenidos. Para este caso, se debe elegir una distribución satisfactoria de los puntos experimentales sobre la región experimental.

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 DCC de tres factores, con un alfa (α) igual a 2. Este valor se eligió para facilitar la configuración de los parámetros de QoS, de tal manera que quedaran en un número entero, y homogenizar el número de decimales a utilizar.

Se define como tratamiento o ejecución en desarrollo del experimento al 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. Con base en lo anterior, un DCC está conformado por tres tipos de puntos, los que posteriormente se convertirán en los tratamientos del experimento [85], así:

Diseño y ejecución del experimento

- Porción factorial: es la réplica de un diseño factorial, para este caso 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 5, donde la región experimental es un cubo regular centrado en el origen $(0,0,0)$ cuyos vértices son los ocho tratamientos. En esa misma figura, se presenta, al lado izquierdo, la tabla con los factores codificados [85].
- Puntos axiales: son dos puntos sobre cada eje a una distancia α del origen. La elección del α está definido por $\alpha = (F)^{1/4}$, donde F es el número de puntos de la parte factorial. Por lo tanto, $\alpha = (8)^{1/4} = 1.681$. Sin embargo, para este caso se eligió un $\alpha = 2$, para facilitar que los valores de cada factor fueran números enteros y con ello, a su vez, facilitar los procesos de configuración en el emulador (NetEm) (ver TABLA 6). En la FIGURA 6 se puede apreciar la ubicación de los puntos axiales del DCC.

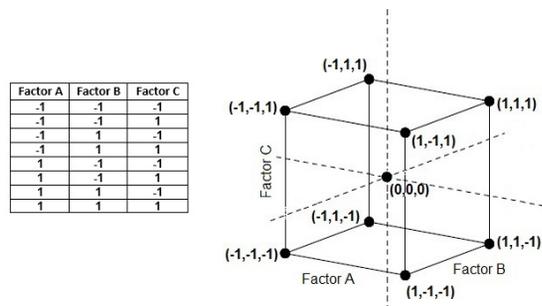


Figura 5. Codificación y representación geométrica de la porción factorial del DCC (tres factores)

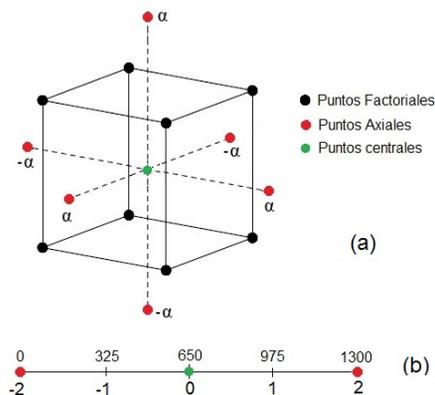


Figura 6. Puntos axiales y factoriales del DCC (a); valores asociados a la codificación del eje X, parámetro retardo (b)

- Puntos centrales: corresponden al número de repeticiones al centro del diseño (n_0), lo que, según se define en [83], para tres factores es $n_0 = 6$.

Tomando en consideración lo anterior, el número de tratamientos para el DCC en cuestión está definido por $N = 2^k + 2k + n_0$ [85], es decir, $N = 2^3 + 2 \times 3 + 6 = 20$ tratamientos.

Para determinar el valor de los factores para cada tratamiento se parte de los rangos presentados en la TABLA 5, los cuales determinan el área de experimentación donde se ubicará el cubo de la FIGURA 6, por lo tanto, el rango de cada factor se debe dividir en cuatro, por el valor del alfa definido (puntos axiales) y los puntos factoriales que están a una distancia de uno del alfa definido.

Tabla 6. Tratamientos para el experimento con los valores codificados y reales asociados a cada factor

Tratamiento	Valores codificados			Valores reales		
	R	Vr	Pp	R	Vr	Pp
1 Puntos factoriales	-1	-1	-1	325	0,0045	0,045
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 Puntos axiales	-2	0	0	0	0,009	0,09
10	2	0	0	1300	0,009	0,09
11	0	-2	0	650	0	0,09
12	0	2	0	650	0,018	0,09
13	0	0	-2	650	0,009	0
14	0	0	2	650	0,009	0,18
15 Puntos centrales	0	0	0	650	0,009	0,09
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

Diseño y ejecución del experimento

Para entender mejor cómo se generan los valores que se presentan en la TABLA 5, se toma como ejemplo el eje X de la sección (a) de la FIGURA 6, que corresponde al retardo (R), cómo el alfa es 2, dicho eje irá desde -2 hasta 2, y será intersectado por los otros ejes en -1, 0 y 1, hasta llegar a 2, como se aprecia en la sección (b) de la FIGURA 6.

El rango definido para el retardo va de 0 hasta 1300 msec. Al dividir dicho rango entre 4, se obtienen los valores asociados a cada punto en la codificación, como se aprecia en la sección (b) de la FIGURA 6. De igual manera, se calculan los valores para cada tratamiento para la variación del retardo y la pérdida de paquetes. En la TABLA 6 se pueden 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), el cual 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 6 se relacionan los puntos factoriales del cubo con sus valores codificados, al igual que los puntos axiales y el punto central presentados en la FIGURA 6. En la sección “Ejecución del experimento” se explica 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 jerárquico de segundo orden para cada tipo de video. Este tipo de modelo tiene un comportamiento más estable, lo que facilita la exploración de la superficie de respuesta que él representa y permite que los términos no significativos permanezcan en el modelo para lograr la jerarquía [85].

El modelo a obtener es de la forma como se presenta en la ECUACIÓN 4, donde: X_i representa a los factores, es decir a los parámetros de QoS; Y a la respuesta de interés, la QoE del video; y ϵ es el error aleatorio:

$$Y = \beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i=1}^k \beta_{ij} X_i^2 + \sum_{i=1}^k \sum_{j=1}^k \beta_{ij} X_i X_j + \epsilon \quad (4)$$

Para estimar los coeficientes (β_i) en la ECUACIÓN 4 se usa el método de mínimos cuadrados, 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 7). Este procedimiento de estimación requiere que el error aleatorio tenga $E(\epsilon)=0$, $V(\epsilon)=\sigma^2$ y que los $\{\epsilon\}$ no estén correlacionados [85].

Tabla 7. Datos del modelo que se presenta 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 de la ECUACIÓN 4 puede ser escrito en términos de los datos presentados en la TABLA 7, como:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \epsilon_i$$

$$Y_i = \beta_0 + \sum_{j=1}^k \beta_j X_{ij} + \epsilon_i \quad i = 1, 2, \dots, n \tag{5}$$

El método de mínimos cuadrados consiste en elegir los β de la ECUACIÓN 5, de tal modo que la suma de los cuadrados de los errores ϵ_i se minimice. Por lo tanto, la función de mínimos cuadrados se define como se indica en la ECUACIÓN 6.

$$S = \sum_{i=1}^n \epsilon_i^2$$

$$S = \sum_{i=1}^n \left[Y_i - \beta_0 - \sum_{j=1}^k \beta_j X_{ij} \right]^2 \tag{6}$$

La función S (ECUACIÓN 6) debe minimizarse con respecto a $\beta_1, \beta_2, \dots, \beta_k$. Los estimadores de mínimos cuadrados, por ejemplo: $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$, deben satisfacer la ECUACIÓN 7.

Diseño y ejecución del experimento

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

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

Al simplificar la ECUACIÓN 7 se obtienen $k+1$ ecuaciones. 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. El modelo que se presentó en la ECUACIÓN 5, puede escribirse en notación matricial como:

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

Donde:

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ \vdots \\ Y_n \end{pmatrix} \quad X = \begin{pmatrix} 1 & X_{11} & X_{12} & \dots & X_{1k} \\ 1 & X_{21} & X_{22} & \dots & X_{2k} \\ 1 & \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot \\ 1 & X_{n1} & X_{n2} & \dots & X_{nk} \end{pmatrix}$$

$$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \vdots \\ \beta_k \end{pmatrix} \quad y \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

El vector de estimadores de mínimos cuadrados β se obtiene minimizando la ECUACIÓN 8, ahora presentada en forma matricial:

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

S puede ser expresada de la siguiente manera:

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

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

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

Lo que implica que:

$$X'X\hat{\beta} = X'y \quad (10)$$

Para resolver la ECUACIÓN 10 en términos de $\hat{\beta}$, se multiplican ambos lados por la inversa $X'X$, así, el estimador de mínimos cuadrados de $\hat{\beta}$ es:

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

Y por lo tanto, el modelo ajustado está dado por:

$$\hat{y} = X'\hat{\beta} \quad (12)$$

Lo que en notación escalar es:

$$\hat{Y}_1 = \hat{\beta}_0 + \sum_{j=1}^k \hat{\beta}_j X_{ij} \quad i = 1, 2, \dots, n \quad (13)$$

Diseño y ejecución del experimento

Partiendo de los valores $\hat{\beta}$ obtenidos en la ECUACIÓN 13, se puede expresar el modelo de la ECUACIÓN 4 de la siguiente forma (ECUACIÓN 14).

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

Por lo tanto, en el modelo propuesto en la ECUACIÓN 14: X_1 = retardo; X_2 = variación del retardo; y X_3 = pérdida de paquetes. Con el fin de hacer el modelo más amigable para su aplicación e interpretación, los factores representados por X_1 , X_2 y X_3 se reemplazan con R , V_R y P_P , respectivamente, como se aprecia en la ECUACIÓN 15.

$$QoE = \beta_0 + \beta_1 R + \beta_2 V_R + \beta_3 P_P + \beta_4 R^2 + \beta_5 V_R^2 + \beta_6 P_P^2 + \beta_7 R V_R + \beta_8 R P_P + \beta_9 V_R P_P \quad (15)$$

Como se puede apreciar, el modelo que se presenta en la ECUACIÓN 15 es 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 “Análisis estadístico de los datos, en este mismo capítulo, se presentan los modelos obtenidos para cada tipo de video y el valor asociado a cada constante β .

Ejecución del experimento

Con base en los tratamientos de la TABLA 6, el experimento se planteó de la siguiente manera:

- para cada tratamiento en el NetEM se configuraron los valores para los factores, de acuerdo con lo planteado en la TABLA 6. Por ejemplo, para el tratamiento 1, los parámetros fueron: retardo, 325 msec; variación del retardo, 0.0045 msec; y pérdida de paquetes, 0.045 %.
- con los factores configurados, se procedió a enviar el video original, que tiene una duración de casi treinta segundos, desde el servidor de video hasta el cliente VLC, donde fue almacenado para su posterior edición y recorte a diez segundos. Esto se hizo con el fin de medir la QoE del video con la herramienta MSU Video Quality Measurement Tool.

- la QoE del video se midió comparando el video original y el video recibido en el cliente VLC, ambos editados a diez segundos. Los valores configurados ($R=325$ mseg, $V_R=0.0045$ y $P_P=0.045$ %) y la medida de QoE realizada se almacenaron –como se aprecia en la TABLA 8 para el tratamiento No.1–, como valores asociados a un tratamiento.

Tabla 8. Medición de QoE obtenida para el tratamiento No.1

Tratamiento	R	V_R	P_P	QoE_Cafetería VQM
1	3,25	0,0045	0,045	4,782823333

Este proceso se repitió para los veinte tratamientos de cada video (ver TABLA 6) y luego se replicó dos veces todo el experimento. Es decir, para cada video se realizó un total de sesenta tratamientos. Además, a los veinte tratamientos realizados se le adicionó uno de control, el Tratamiento No.0, el cual consistió en configurar los tres parámetros de QoS en 0 (cero), con el fin de obtener el video transmitido, sin ningún tipo de distorsión.

OBTENCIÓN DE LOS RANGOS DE LOS FACTORES DEL EXPERIMENTO

Obtener estos rangos requirió de varias iteraciones, realizando réplicas de prueba de todo el experimento explicado en la sección anterior. Las primeras pruebas se realizaron utilizando los rangos definidos en la Recomendación UIT-T Y.1541 [59] para los parámetros elegidos de QoS, esto es: retardo, 0 a 400 mseg; variación del retardo, 0 a 50 mseg; y pérdida de paquetes, 0 a 1×10^{-3} .

Los datos obtenidos en las réplicas iniciales no fueron estadísticamente válidos, pues el análisis de varianza realizado mostró una fidelidad muy baja (cálculo del R_{cuadrado}) en el modelo obtenido. Además, se presentaron los siguientes inconvenientes técnicos:

- los videos obtenidos en las primeras pruebas con los rangos de la Recomendación UIT-TY.1541 [59] presentaban distorsiones generalizadas que no permitían su correcta visualización, por ende, la medición obtenida de VQM era elevada, lo que indicaba una calidad deficiente del video;
- para algunos valores de los parámetros de QoS los videos no se podían almacenar, porque la pérdida de paquetes impedía decodificar el video de manera correcta en el cliente; y

Diseño y ejecución del experimento

- para algunos datos generados se realizaron pruebas estadísticas que no permitían generar un modelo –o solo modelos con niveles de aproximación estadísticamente muy bajos–, debido a que los videos de una prueba no presentaban variaciones, por ejemplo, de los veinte tratamientos, 90% de los videos tenía una distorsión muy alta y por ello un valor de QoE muy bajo.

Por estas razones se realizó una serie de pruebas en la maqueta experimental ajustando los valores de los parámetros de QoS, de tal manera que fuera posible transmitir cada video y apreciar distorsiones significativas mientras se transmitía cada uno de ellos. Después de numerosas pruebas, se obtuvo la serie de valores que se presenta en la TABLA 5. Lo anterior implica que los rangos de los parámetros de QoS se generaron de manera empírica y por lo tanto, 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 5) en las pruebas realizadas se ejecutó el respectivo análisis de varianza a los modelos obtenidos con resultados estadísticamente válidos. Además, después de la modificación de los parámetros de QoS, los videos transmitidos se podían decodificar en el cliente, algo de suma importancia, porque como se dijo, el video debía poder ser almacenado en él, para su posterior edición a diez segundos. En la FIGURA 7 se presenta de manera resumida el proceso general para la obtención de los datos para la generación del modelo.

Cabe mencionar la gran cantidad de pruebas y de archivos de video que se procesaron durante toda la etapa experimental desarrollada para obtener los datos necesarios para generar los modelos. Para cada video, una prueba consistió en ejecutar los veinte tratamientos explicados, pero para obtener un video adecuado en un tratamiento era necesario realizar al menos cinco transmisiones de video para dicho tratamiento, con el fin de trabajar con un video que se pudiera decodificar y almacenar adecuadamente en el cliente VLC. Entonces, para una sola prueba se requería procesar aproximadamente cien videos o realizar cien transmisiones de él; 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 novecientos videos, es decir, para los tres modelos propuestos se procesaron más de 2.700 videos, desde las pruebas iniciales realizadas para definir los rangos de los parámetros de QoS, hasta las pruebas que permitieron obtener datos válidos para generar cada modelo.



Figura 7. Proceso para obtener los datos para la generación de los tres modelos

ANÁLISIS ESTADÍSTICO DE LOS DATOS

A continuación se presenta el modelo obtenido para cada video (Cafetería, Entrevista y Baloncesto), junto con su respectivo análisis estadístico que permite establecer la precisión en la interpretación de los datos. En todos los casos, los datos obtenidos se presentan en una tabla, en ella, para cada tratamiento se relacionan los valores de los factores $-R$, V_R y P_P y la medida de QoE obtenida con el algoritmo VQM. Las TABLAS 9, 11 y 13 corresponden, respectivamente, a la información de los videos: Entrevista, Cafetería y Baloncesto, para su lectura es conveniente tener en cuenta que un valor menor en su última columna (VQM) significa que la calidad del video es mejor, en consecuencia, entre más alto sea ese valor, menor será la calidad del video. En el capítulo 5 se explica cómo se mapeo el valor VQM con la escala MOS en las pruebas comparativas de desempeño.

En las TABLAS 10, 12 y 14 se presenta el análisis de varianza de los datos presentados en las TABLAS 9, 11 y 13, respectivamente, los cuales se obtuvieron utilizando la herramienta Minitab. Para el análisis de los datos de los tres videos se utilizó un nivel de significancia de 0.05 ($\alpha = 0.05$). Cabe recordar que el nivel de significancia, en su definición básica, indica 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; en este

Diseño y ejecución del experimento

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, lo que se hace comparando el valor p con el nivel de significancia (α), donde, si $p < \alpha$ se puede concluir que el parámetro tiene influencia significativa en el modelo. 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) y explica su relación con los factores (parámetros de QoS) [87]. El R-cuadrado presenta valores en un rango de 0 a 100%, cuanto mayor sea ese valor, mejor será el ajuste del modelo a los datos.

Modelo obtenido para el video Entrevista

En la FIGURA 8 se presenta el diagrama de barras de los datos incluidos en la TABLA 9, junto con su coeficiente de variación ($c.v=37,24$), la media de los datos (1,849) y el valor de una desviación estándar por encima y por debajo de la media ($\sigma = 0,689$, para 2,538 y 1,16, respectivamente). En este caso, el valor de VQM del tratamiento 2 (1,071) es el menor de los obtenidos y el del tratamiento 6 (3,405) el más alto. Por tanto, corresponden, respectivamente, a la mayor y menor calidad de video. En la TABLA 10 se presenta el análisis de varianza de los datos presentados en la TABLA 8. Cabe recordar que la

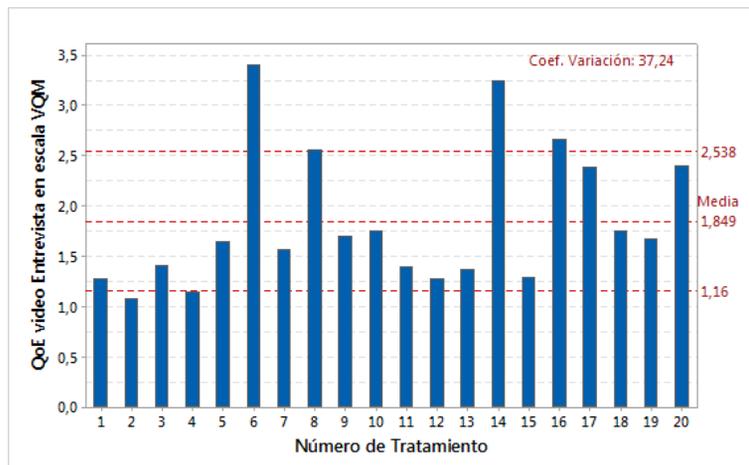


Figura 8. QoE obtenida para el video Entrevista en escala VQM

Tabla 9. Datos obtenidos para generar el modelo del video Entrevista

Tratamiento	R (mseg)	V _R (mseg)	Pp (%)	QoE 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	0	0,009	0,090	1,69293
10	1300	0,009	0,090	1,75235
11	650	0	0,090	1,40056
12	650	0,018	0,090	1,27615
13	650	0,009	0	1,36310
14	650	0,009	0,180	3,24128
15	650	0,009	0,090	1,29295
16	650	0,009	0,090	2,66169
17	650	0,009	0,090	2,38688
18	650	0,009	0,090	1,74842
19	650	0,009	0,090	1,66789
20	650	0,009	0,090	2,40491

comparación de los valores p y α permite establecer si un parámetro tiene influencia significativa en el modelo ($p < \alpha$).

De los datos de la TABLA 10 se concluye que los parámetros con influencia significativa en el modelo son la pérdida de paquetes y la interacción entre los factores retardo y pérdida de paquetes. Esto se obtiene al analizar la última columna (p), cuyo valor debe ser menor a 0,05 para ser considerado influyente en el modelo. Por otra parte, el R-cuadrado obtenido fue de 79,59 %, lo cual indica que el modelo explica en esa proporción la variabilidad de los datos con que se generó. El modelo obtenido usando Minitab para el video Entrevista, que se identifica como Modelo Video Movimiento Bajo (MVMB), se presenta en la ECUACIÓN 16.

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

Diseño y ejecución del experimento

Tabla 10. Análisis de varianza para los datos obtenidos del video Entrevista

Fuente	Grados de libertad	Suma de los cuadrados	Media de los cuadrados	F	p
Modelo	9	7,17118	0,7968	4,33	0,016
..Lineal	3	4,45463	1,48488	8,07	0,005
....R	1	0,3601	0,3601	1,96	0,192
....V _R	1	0,05876	0,05876	0,32	0,584
....P _P	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
....V _R *V _R	1	0,81487	0,81487	4,43	0,062
....P _P *P _P	1	0,09335	0,09335	0,51	0,493
..Interacción de 2 factores	3	1,53299	0,511	2,78	0,096
....R*V _R	1	0,08205	0,08205	0,45	0,519
....R*P _P	1	1,29101	1,29101	7,02	0,024
....V _R *P _P	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	9,0107			

Al modelo que se presenta en la ECUACIÓN 16 se le incluyeron, no solo los factores que tenían influencia significativa, sino también los que no la tenían o cuya influencia era muy baja, porque al retirar estos últimos del análisis de varianza, el R-cuadrado bajaba a 65,29%, es decir que esos factores que se podrían retirar del modelo, de alguna manera aportan para mejorar el R-cuadrado y con ello la precisión del modelo.

Las FIGURAS 9 y 10 corresponden a las gráficas de contorno del parámetro que influye estadísticamente en el modelo (P_P) y de los dos parámetros que no lo hacen (R y V_R). Como se puede apreciar, para R y V_R la variación de la QoE con respecto de la pérdida de paquetes es casi constante, en cada figura hay un sitio donde se obtiene la mejor medición de QoE y ello ocurre cuando la variación del retardo se aproxima a 0,0050 mseg y el retardo es cercano a 1000 mseg. La FIGURA 11 corresponde a los parámetros que no influyen estadísticamente en el modelo (R y V_R), como se puede apreciar, para ciertos intervalos es posible obtener una medición de QoE de hasta 1,6, lo que se debe a que en este tipo de video, por los pocos cambios que tiene, ciertas distorsiones no se reflejan en la medida de VQM.

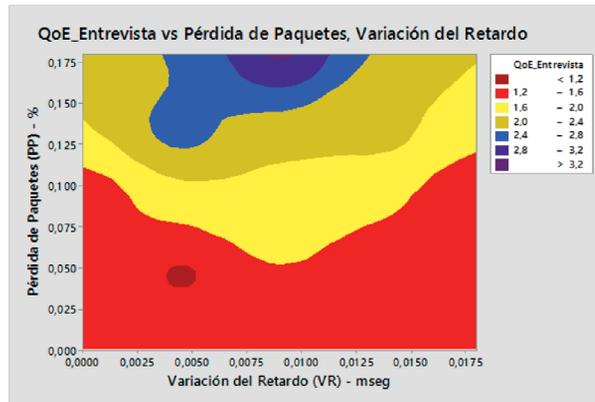


Figura 9. QoE_Entrevista vs P_p y V_R

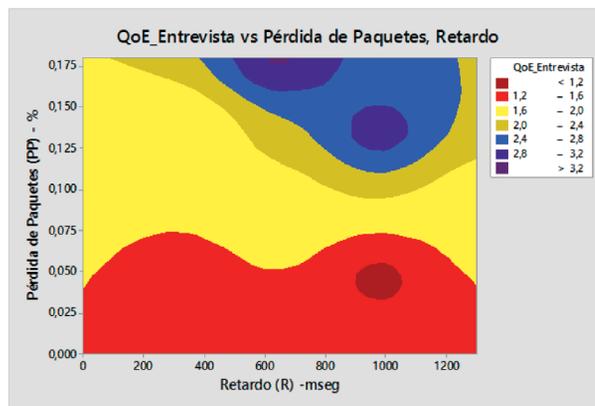


Figura 10. QoE_Entrevista vs P_p y R

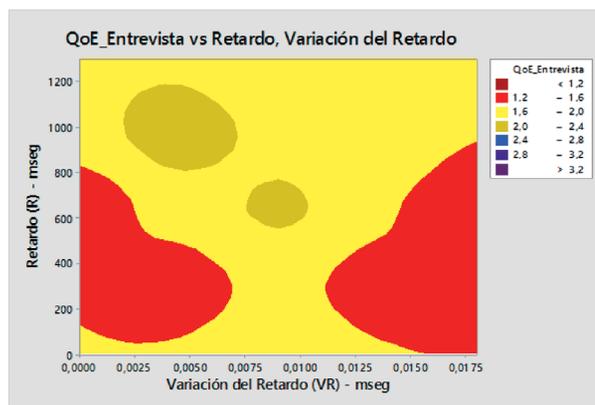


Figura 11. QoE_Entrevista vs R y V_R

Diseño y ejecución del experimento

De las figuras, es claro que el parámetro de mayor influencia es P_p , como se puede ver, después de 0.1%, la QoE del video se empieza a degradar; sin embargo, se debe tener en cuenta que por el tipo de video (sin movimiento) los valores que entrega el modelo propuesto son mucho más bajos, si se compara, por ejemplo, con el modelo obtenido para el video Cafetería, que se explica a continuación.

Modelo obtenido para el video Cafetería

En la TABLA 11 se presentan los resultados de la ejecución del experimento para el video Cafetería, en ella se puede apreciar la diferencia en el valor de VQM obtenido en este caso, en comparación con el obtenido para el video Entrevista, la que surge por las diferencias en sus características espacio-temporales (el video Entrevista es de menor movimiento). En la misma Tabla se puede apreciar cómo para un valor en retardo de 1300 msec (tratamiento 10), la calidad del video se degrada a su valor máximo en escala VQM (11,43), mientras que para un valor de retardo de 325 msec (tratamiento 1) se obtiene el valor mínimo de VQM (3,65), es decir un video de mejor calidad.

En el diagrama de la FIGURA 12 se presentan los datos de la TABLA 11, junto con su coeficiente de variación ($c.v=27,38$), la media de los datos (5,93) y una desviación estándar por encima y por debajo de la media ($\sigma = 1,624$, para 7,55 y 4,31, respectivamente). Un valor de VQM bajo (cercano a 3.65) indica una

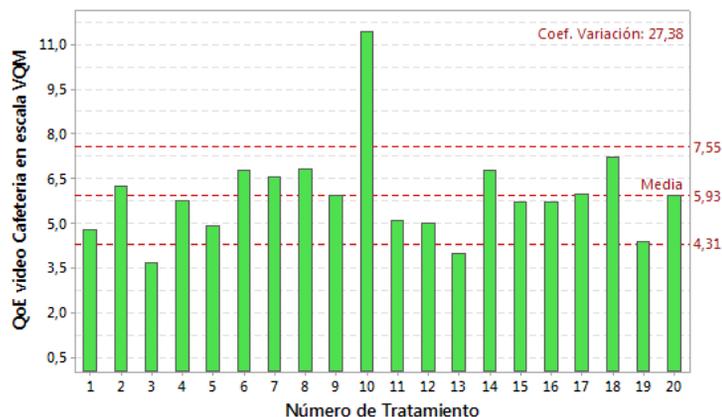


Figura 12. QoE obtenida para el video Cafetería en escala VQM

Tabla 11. Datos obtenidos para generar el modelo del video Cafetería

Tratamiento	R (mseg)	V _R (mseg)	P _p (%)	QoE 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

calidad del video mejor en comparación de uno alto (cercano a 11,43), que indica un video de calidad deficiente.

La TABLA 12 corresponde al análisis de varianza de los datos de la TABLA 11, que fueron obtenidos usando Minitab, y permite concluir que los parámetros con influencia significativa en el modelo son el retardo, la pérdida de paquetes y el retardo al cuadrado. En el análisis de varianza de los datos de la TABLA 12, el R-cuadrado obtenido fue de 84.23%, lo cual indica que el modelo explica en esa proporción la variabilidad de los datos obtenidos. El modelo obtenido usando Minitab para el video Cafetería, identificado como Modelo Video Movimiento Moderado (MVMM), se presenta en la ECUACIÓN 17.

$$QoE = 3,89 - 3,19 \times 10^{-3} R + 85 V_R + 18,1 P_p + 6 \times 10^{-6} R^2 - 0,082 R V_R - 11,937 \times 10^3 V_R^2 + 0,012 R P_p + 20,20 V_R P_p - 79,1 P_p^2 \quad (17)$$

Diseño y ejecución del experimento

Tabla 12. Análisis de varianza para los datos obtenidos del video Cafetería

Fuente	Grados de libertad	Suma de los cuadrados	Media de los cuadrados	F	p
Modelo	9	42,2409	4,6934	5,97	0,005
..Lineal	3	23,8949	7,965	10,14	0,002
....R	1	17,3723	17,3723	22,11	0,001
....V _R	1	0,0007	0,0007	0	0,977
....P _P	1	6,5219	6,5219	8,3	0,016
..Cuadrado	3	16,6433	5,5478	7,06	0,008
....R*R	1	11,231	11,231	14,3	0,004
....V _R *V _R	1	1,4692	1,4692	1,87	0,201
....P _P *P _P	1	0,6446	0,6446	0,82	0,386
..Interacción de 2 factores	3	1,7027	0,5676	0,72	0,561
....R*V _R	1	0,1158	0,1158	0,15	0,709
....R*P _P	1	0,2483	0,2483	0,32	0,586
....V _R *P _P	1	1,3385	1,3385	1,7	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			

Tal como ocurrió con la ecuación del MVMB (Ecuación 16), a la del MVMM se le incluyeron todos los factores, tanto los que tenían influencia significativa, como los que no la tenían o ella era muy baja. Se hizo así porque al retirarlos del análisis de varianza y generar de nuevo el modelo, el R-cuadrado bajaba a un 79.63%, es decir que los factores que se podrían retirar del modelo en algo aportan para la mejora del R-cuadrado y a la precisión del modelo. Para analizar el comportamiento del MVMM, en la FIGURA 13 se presenta la gráfica de contorno de los parámetros que influyen en él (R y P_P), en ella se puede apreciar cómo en dos partes se obtiene un valor de QoE inferior a 4.5, y cómo con valores inferiores en retardo a 1000 mseg y pérdida de paquetes de 0.1% se obtienen logra obtener valores de VQM inferiores a 5.7. Cabe recordar que un menor valor de VQM indica una mejor calidad del video y un mayor valor, lo opuesto.

En las FIGURAS 14 Y 15 se puede apreciar cómo la variación del retardo tiene un aporte mínimo al modelo, pues ante cambios en su valor, la QoE permanece constante (ver el área roja). Además, se puede apreciar que la medida obtenida de QoE está entre 4.5 y 5.7, lo que indica que se trata de video de buena calidad.

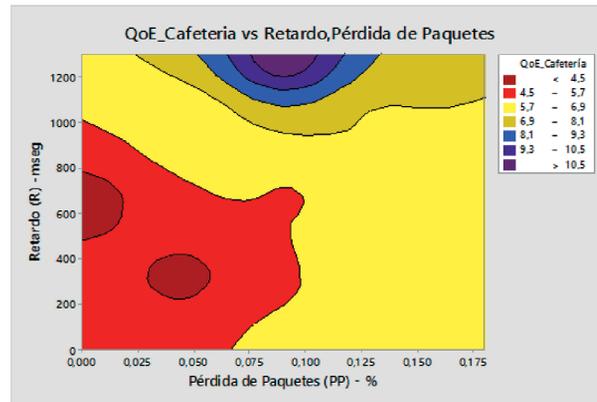


Figura 13. QoE_Cafetería vs R y P_p

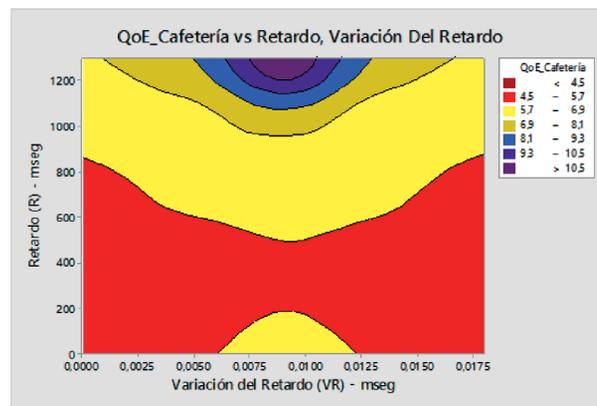


Figura 14. QoE_Cafetería vs R y V_R

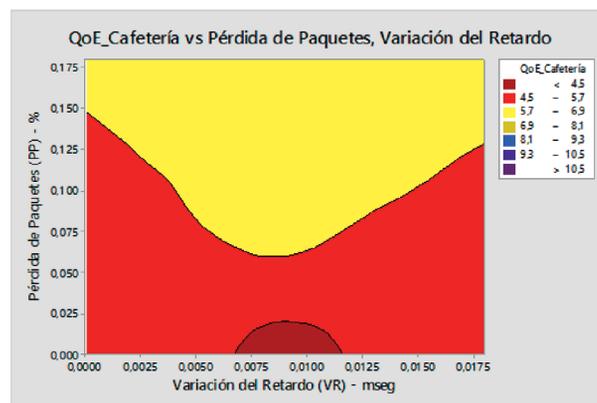


Figura 15. QoE_Cafetería vs P_p y V_R

Diseño y ejecución del experimento

En la FIGURA 13 se puede apreciar además que los parámetros que tienen influencia significativa en el modelo son los que hacen que el valor de QoE mejore o se degrade, en este caso, la P_p influye en el resultado. Asimismo, en las FIGURAS 14 Y 15 se observa que 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.

Modelo obtenido para el video Baloncesto

Los datos obtenidos en la ejecución del experimento para el video Baloncesto se presentan en la TABLA 13. Como consecuencia de la diferencia en sus características espacio-temporales –Baloncesto es un video con alto movimiento.–, sus valores son diferentes de los obtenidos para los videos Entrevista y Cafetería.

En el diagrama de la FIGURA 16 se presentan los datos de la TABLA 13, junto con su coeficiente de variación ($c.v= 13,27$), la media de los datos (6,208) y una desviación estándar por encima y por debajo de la media ($\sigma=0,824$, para 7,032 y 5,384, respectivamente). Comparativamente, el coeficiente de variación obtenido para este video fue bajo, inferior al obtenido en los videos Entrevista y Cafetería.

Los resultados obtenidos en cada tratamiento son similares y esto se refleja en el análisis de varianza que se presentará más adelante. Estos resultados se explican

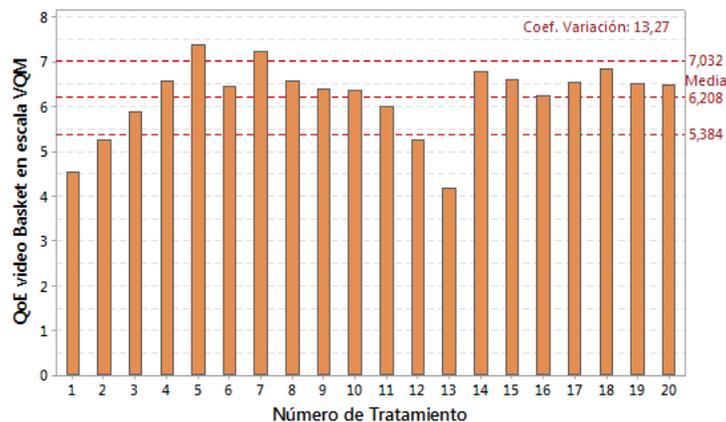


Figura 16. QoE obtenida para el video Baloncesto en escala VQM

Tabla 13. Datos obtenidos para generar el modelo del video Baloncesto

Tratamiento	R (mseg)	V_R (mseg)	Pp (%)	QoE 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

con el tipo de video, pues el mayor movimiento implica una mayor pérdida de paquetes y con ello una pérdida de cuadros que hace que el video se degrade.

El valor de VQM bajo, obtenido en el tratamiento 13 (4,17), indica una calidad del video mejor en comparación con el valor de VQM del tratamiento 5 (7,375). En la TABLA 14 se presenta el análisis de varianza de los datos presentados en la TABLA 13, obtenidos usando la herramienta Minitab. Con estos datos se puede concluir que los parámetros que tiene influencia significativa en el modelo son la pérdida de paquetes, el cuadrado de la variación del retardo, la interacción entre retardo y pérdida de paquetes y la interacción entre variación del retardo y pérdida de paquetes.

El R-cuadrado obtenido del análisis de varianza de los datos de la TABLA 14 fue de 88,56%, lo cual indica que el modelo explica el 88,56% la variabilidad de los datos con que él se generó. El modelo obtenido mediante la herramienta

Diseño y ejecución del experimento

Tabla 14. Análisis de varianza para los datos obtenidos del video Baloncesto

Fuente	Grados de libertad	Suma de los cuadrados	Media de los cuadrados	F	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
....V _R	1	0,0846	0,08464	0,57	0,466
....P _P	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
....V _R *V _R	1	1,0828	1,08278	7,34	0,022
....P _P *P _P	1	1,5319	1,53187	10,39	0,009
..Interacción de 2 factores	3	2,0073	0,66909	4,54	0,030
....R*V _R	1	0,0097	0,00968	0,07	0,803
....R*P _P	1	1,0866	1,08657	7,37	0,022
....V _R *P _P	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			

Minitab para el video Baloncesto se presenta en la ECUACIÓN 18 y se identifica como Modelo Video Movimiento Alto (MVMA).

Al modelo que se presenta en la ECUACIÓN 18 se le incluyeron todos los factores, tanto los que tenían influencia significativa, como los que no la tenían o ella era muy baja. Aunque en este caso al retirarlos el valor del R-cuadrado casi no cambió (bajó a 87,88%), incluirlos facilita su comparación con los otros dos.

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

En las FIGURAS 17 y 18 se grafica el parámetro que estadísticamente influye en el modelo (pérdida de paquetes). En ellas se puede apreciar cómo se logra una medición de QoE entre 4.5 y 5 cuando el retardo está en un rango de 400 a 800 mseg y la variación del retardo se encuentra entre 0.005 y menos de 0.0125.

En la FIGURA 19 se muestra cómo los dos parámetros que no influyen estadísticamente en el modelo (R y V_R) no presentan intervalos donde se logre apreciar una QoE inferior a 5.0, lo que indica su muy baja influencia en él.

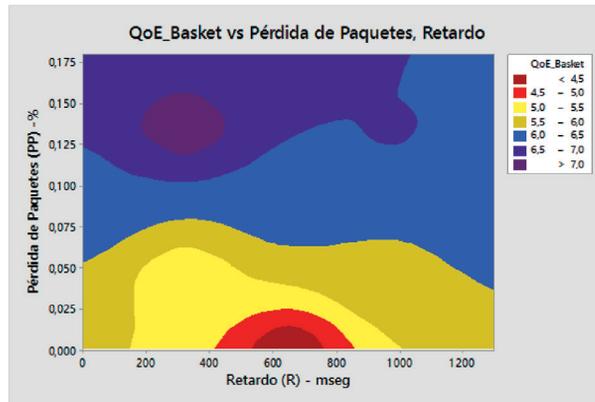


Figura 17. QoE_Baloncesto vs P_p y R

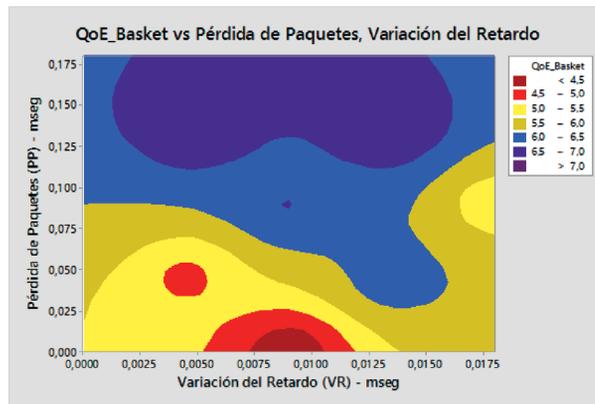


Figura 18. QoE_Baloncesto vs P_p y V_R

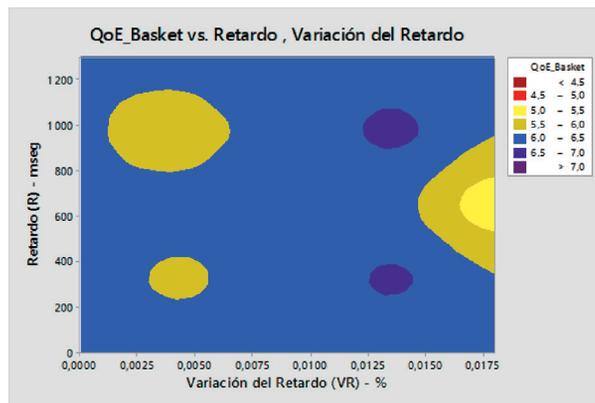


Figura 19. QoE_Baloncesto vs R y V_R

Diseño y ejecución del experimento

Comparativamente, se destaca cómo, al igual que en los videos Cafetería y Entrevista, el parámetro que más influye en la QoE es la pérdida de paquetes, solo que en este caso (MVAM) el rango para obtener una QoE aceptable es mucho menor, debido a sus características espacio – temporales, pues la pérdida de un cuadro afecta la percepción de calidad considerablemente.

Acerca de los resultados

En la TABLA 15 se presenta el resumen del modelo obtenido para cada video, incluyendo los coeficientes asociados a cada término del modelo, los factores que estadísticamente influyen en ellos y su respectivo R-cuadrado.

Se aprecia que el factor que influye estadísticamente en los tres videos es la pérdida de paquetes, un factor crítico al momento de transmitir IPTV, porque al no existir la retransmisión de cuadros, el codificador debe recuperar los cuadros perdidos con base en la información que llega en el cliente, una tarea que en algunos casos no es eficaz y, en consecuencia, degrada la QoE asociada al video produciendo en el video del cliente, congelamiento y pérdida de nitidez.

El efecto de la pérdida de paquetes no es el mismo en todos los casos, sino que depende de las características espacio-temporales de cada video. En el caso de los de alto movimiento, donde el rango para obtener una QoE aceptable es menor al de los otros dos videos, la afectación es mayor.

Tabla 15. Coeficientes de los tres modelos obtenidos

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

Proponer un único modelo que permita evaluar la QoE para el servicio de IPTV hace que el modelo se deba ajustar con la introducción de una serie de constantes –necesarias para mantener un R-cuadrado alto–. 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 descrita se puede apreciar en [45], donde Kim y Choi ponderan el peso de cada parámetro de QoS con una constante que se calcula utilizando un ambiente de computación invasiva, y en [47]. En esos dos casos, adicionalmente se tienen en cuenta otras constantes dentro del modelo, de acuerdo con: el tamaño de la pantalla del usuario; el tipo de servicio que el cliente ha contratado con el proveedor de servicios; y el tipo de cuadro (I, B o P) perdido (la pérdida de un cuadro I degrada la QoE).

Para realizar cálculos y obtener la QoE final, Yamagishi, Kawano y Hayashi [48] se basan en una serie de nueve constantes calculadas a partir de pruebas subjetivas, las mismas que se incorporan al modelo y solo son válidas para los videos utilizados en el experimento.

Con los modelos propuestos no es necesario considerar ningún tipo de variable externa –tamaño del receptor de TV, tipo de servicio, variables calculadas mediante pruebas subjetivas, etc.–, lo que le permite su aplicación en cualquier tipo de video en tiempo real, solo obteniendo los valores de los parámetros de QoS (R , V_R y P_P).

Con base en lo que se presentó en la sección “Videos fuente” (capítulo 3), para la selección de uno de los modelos propuestos se deben conocer las características espacio-temporales del video a evaluar, para así elegir el más adecuado. En la revisión sistemática que se realizó a más de catorce 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 elegir el modelo adecuado a cada tipo de video, de acuerdo con su nivel de movimiento, así: movimiento bajo, IE:30-40 / IT:5-10; movimiento medio, IE:60-70 / IT:5-10; y movimiento alto, IE:50-60 / IT:15-20.

Cabe precisar 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).

Análisis de la evaluación del desempeño de los modelos propuestos

Ajuste de datos para el análisis comparativo

De acuerdo con la revisión sistemática presentada en el capítulo 2 y con lo que plantean Moldovan, Ghergulescu y Munteanen [88], se propone como procedimiento para el análisis de desempeño de los modelos propuestos: seleccionar el conjunto de datos con los cuales se van a desarrollar las comparaciones; estandarizar o normalizar todos los datos a una escala común; y realizar las comparaciones utilizando herramientas estadísticas que permitan deducir el desempeño o precisión del modelo. Adicionalmente, se decidió trabajar con los primeros quince tratamientos realizados a cada video, porque en los tratamientos 16 a 20 la configuración de los parámetros de QoS es igual. Esta decisión va alineada con las pruebas subjetivas que se realizaron a los videos utilizados para generar los datos y obtener los modelos (ver ANEXO 1), pruebas que incluyen el tratamiento cero –video sin distorsión–. En la TABLA 16 se presentan los resultados que arrojó cada modelo propuesto –ECUACIONES 16, 17 y 18– al momento de evaluar los valores en escala VQM de los parámetros de QoS en cada uno de los tratamientos seleccionados.

Como se puede apreciar en la TABLA 16, el valor de la QoE que se obtiene del tratamiento cero es bajo, lo que indica un nivel de calidad superior; asimismo, se observa que el valor de la QoE obtenida para el modelo de video de movimiento bajo es menor que la de los modelos de movimiento medio y alto, lo que se debe a sus características espacio-temporales (bajo movimiento), pues su composición hace que su degradación no sea tan notoria.

Como se indicó, una vez seleccionado el conjunto de datos con el cual se desarrollarían las comparaciones, se estandarizaron las medidas de QoE

Análisis de la evaluación del desempeño de los modelos propuestos

Tabla 16. Resultados de la evaluación de los modelos

Tratamiento	R(mseg)	V _R (mseg)	P _p	MVMB	MVMM	MVMA
0	0	0	0	0,65	3,89	0,55
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	5,21	7,57
6	975	0,0045	0,135	2,98	6,914	6,91
7	325	0,0135	0,135	1,637	6,256	7,15
8	975	0,0135	0,135	2,372	7,48	6,63
9	0	0,009	0,09	1,372	6,313	6,707
10	1300	0,009	0,09	1,638	9,942	7,001
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	3,67	4,536
14	650	0,009	0,18	3,18	6,233	7,197
15	650	0,009	0,09	1,927	5,592	6,854

entregadas por cada modelo, llevándolas a una escala común, para así facilitar su interpretación, es decir, definir la QoE asociada al video en escala de MOS y realizar pruebas de fidelidad con otros modelos objetivos.

Para la estandarización se analizaron y ejecutaron dos técnicas: el análisis de conglomerados (*clusters*) [89] y VQAMap [88]. El análisis de clusters es una técnica que busca agrupar variables de un conjunto de datos tratando de lograr homogeneidad en cada grupo; al analizarla, se encontró que esta técnica no era la adecuada en este caso porque requiere 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 $K=5$, sino para $K=2$ o $K=3$, dependiendo de la dispersión de los valores de QoE que se obtienen para cada tipo de video. Adicionalmente, los resultados de las pruebas de correlación que se realizaron entre las pruebas subjetivas y los resultados que arroja esta técnica fueron inferiores a las obtenidas con el procedimiento VQAMap. Por esta razón, no se utilizó esta técnica. VQAMap [88], por su parte, 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 su uso en la evaluación de video y, aunque se enfoca en modelos basados en referencia completa, también se puede emplear para modelos basados en referencia reducida y sin referencia.

El siguiente procedimiento fue utilizado para obtener el mapeo de los valores que entregan los modelos propuestos a la escala MOS:

1. Realizar una regresión no lineal entre los valores que entrega cada modelo propuesto en escala VQM (ver TABLA 16) y las medidas subjetivas (MOS_k) de los videos utilizados para generar los datos. En las diferentes bases de datos donde ha sido empleado [90], [91], se ha demostrado empíricamente que una función de mapeo no lineal funciona bien. La función es un polinomio cúbico de la forma que se presenta en la ECUACIÓN 19, en ella: $MOS_k^{(p)}$ representa el valor calculado de la calidad percibida por el usuario; VQM_k es la medida que entrega el modelo propuesto para el tratamiento k ; y 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 $MOS_k^{(p)}$ calculado.

$$MOS_k^{(p)} = a + bVQM_k^3 + cVQM_k^2 + dVQM_k \quad (19)$$

2. Calcular el R_{cuadrado} de la bondad de ajuste del modelo obtenido mediante la regresión no lineal, valor útil para indicar el grado por el cual las variaciones en los valores MOS son contabilizadas por el modelo de regresión no lineal [90]. Un valor de R_{cuadrado} cercano a 1 indica un mejor ajuste, sin embargo, dependiendo del número de puntos de datos, incluso un R_{cuadrado} de 0,25 podría considerarse aceptable [92]. El modelo también debe ser estadísticamente significativo (por ejemplo, $p < 0,05$ a un nivel de confianza del 95%), por lo que el software Minitab se debe configurar para realizar la regresión no lineal a ese nivel de confianza.
3. Hacer 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 y el valor para cada nivel de MOS, y con estos valores establecer la escala para mapear los valores VQM a MOS. En la TABLA 17 se presentan los valores obtenidos de las pruebas subjetivas realizadas a cada tipo de video y el valor que entrega cada modelo propuesto en escala VQM. Con base en estos datos se realizó el procedimiento descrito.

Análisis de la evaluación del desempeño de los modelos propuestos

Tabla 17. Valores de las pruebas subjetivas realizadas a los tres videos con su correspondiente valor en escala VQM

Tratamiento	MVMB		MVMM		MVMA	
	MOS	QoE	MOS	QoE	MOS	QoE
0	4,5	0,65	4,211	3,89	3,64	0,55
1	4,061	1,353	4,123	4,396	4,896	4,828
2	4,096	0,884	3,509	6,801	4,43	5,642
3	3,474	1,715	2,64	3,805	3,07	5,757
4	3,632	0,841	3,044	5,731	3,123	6,711
5	4,228	1,841	3,544	5,21	3,044	7,57
6	2,667	2,98	3,939	6,914	3,229	6,91
7	3,825	1,63	2,939	6,256	2,798	7,15
8	3,491	2,372	3,298	7,48	3,254	6,63
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	3,64	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	3,67	4,746	4,536
14	3,044	3,18	2,947	6,233	3,307	7,197
15	4,211	1,927	3,079	5,592	3,474	6,854

En la TABLA 18 se presenta el valor de las constantes a, b, c, d y valor del R_{cuadrado} calculado para cada modelo después de realizar la regresión no lineal para los datos de la TABLA 17. En ella se puede observar que el R_{cuadrado} más bajo corresponde al MVMM (Cafetería), debido a la dispersión de los datos. Esto se corrobora al calcular la varianza de los datos presentes en la TABLA 18, donde la más alta corresponde a las medidas VQM entregadas por el MVMM.

Las gráficas ajustadas en la regresión no lineal realizada a los datos de cada modelo se pueden apreciar en las FIGURAS 20, 21 y 22.

Tabla 18. Constantes de a,b,c,d y R_{cuadrado} de la regresión no lineal para cada modelo propuesto

Modelo	A	B	C	D	R Cuadrado
MVMB	4,35632	-0,0599065	0,142379	-0,339642	0.59
MVMM	10,7662	-0,0289677	0,562472	3,58745	0.32
MVMA	31,5951	-0,0735298	1,58238	-11,5316	0.72

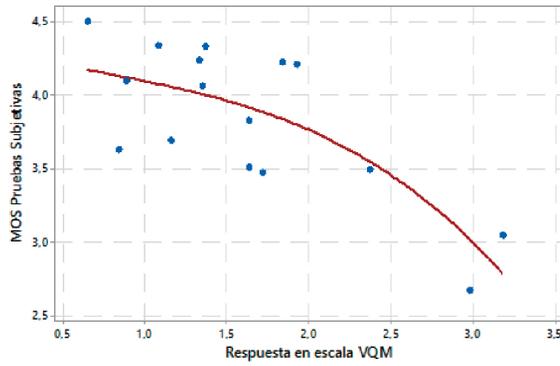


Figura 20. Regresión no lineal: MVMA - Entrevista

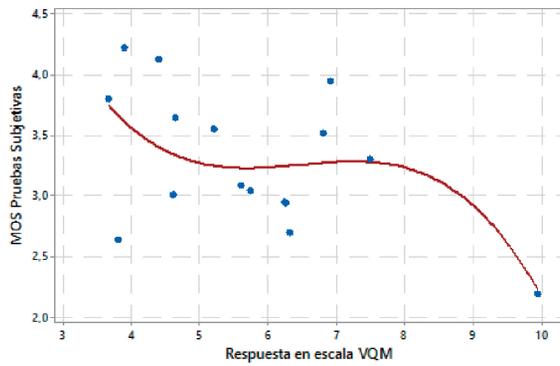


Figura 21. Regresión no lineal: MVMM - Cafetería

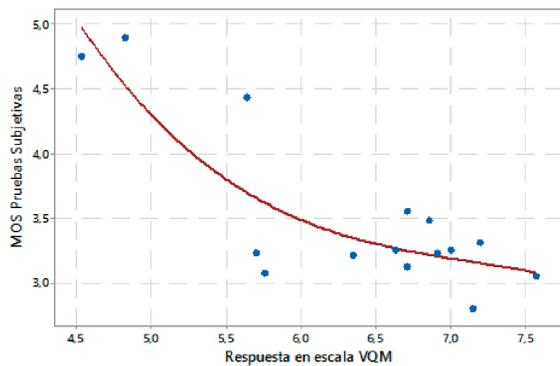


Figura 22. Regresión no lineal: MVMA - Baloncesto

Análisis de la evaluación del desempeño de los modelos propuestos

Como se comentó, el modelo con menor bondad de ajuste es el MVMM –los otros dos presentan una bondad de ajuste superior a 0.5–, lo que indica que se ajusta de manera adecuada a los datos.

Es importante destacar que para la regresión no lineal que se presenta en la FIGURA 20 fue necesario descartar la pareja de datos del tratamiento cero, por ser un dato atípico que hacía que el valor de R_{cuadrado} fuera inferior al 20% –valor con el que el modelo obtenido no sería estadísticamente válido–. Cabe mencionar que este dato atípico es producto de la evaluación subjetiva de dicho video, en la cual, por circunstancias ajenas al experimento, fue calificado como regular.

Después de obtener la regresión lineal para cada modelo (ver TABLA 18), se realizó la interpolación inversa, eligiendo cuatro parejas de datos por cada modelo; luego, a cada grupo de datos se le aplicó el método de interpolación de Newton. El resultado obtenido permite definir los rangos para realizar el mapeo de escala VQM a escala MOS (ver TABLA 19).

Tabla 19. Escalas para mapear MOS vs VQM para cada video

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

La variación de los valores de cada escala depende del tipo de video, por sus características espacio-temporales. Con los intervalos de cada escala definidos, se procedió a mapear el valor de la QoE obtenido en la TABLA 16 a escala MOS (ver TABLA 20).

A manera de ejemplo el mapeo se realizó de la siguiente manera: Al modelo MVMB con tratamiento No.1 y valor 0,65 de QoE, le corresponde una calificación de 5,0 en escala MOS; para el modelo MVMM con tratamiento No.6 y valor 6,914 de QoE, una calificación de 4,0 en escala MOS; y para el modelo MVMA con tratamiento No.15 y valor 6,349 de QoE, una calificación de 4,0 en escala MOS.

Tabla 20. Medidas obtenidas de cada modelo en escala MOS

Tratamiento	MVMB	MVMM	MVMA
0	5	4	Descartado
1	5	4	5
2	5	4	4
3	4	4	4
4	5	4	4
5	4	4	4
6	4	4	4
7	4	4	4
8	4	4	4
9	5	4	4
10	4	3	4
11	5	4	4
12	5	4	4
13	5	4	5
14	3	4	4
15	4	4	4

Con los resultados estandarizados a la escala MOS se procedió a realizar el análisis del desempeño de los modelos propuestos, comparándolos contra dos modelos objetivos y con las pruebas subjetivas, siguiendo las recomendaciones del VQEG [93][94].

Desempeño: modelos propuestos vs modelos objetivos

Los modelos objetivos seleccionados para estas comparaciones de desempeño fueron PSNR (*Peak Signal to Noise Ratio*) y SSIM (*Structural SIMilarity*), ambos de amplio uso en la evaluación de video [15], [49], [68], [69], [75], [77], [78], [95]-[99], y los más usados para comparar el desempeño de nuevos modelos propuestos. PSNR es un modelo de referencia completa que compara pixel a pixel de cada cuadro, evalúa mediante una función simple del error medio cuadrado (*Mean Squared Error*, MSE) entre el video de referencia y el video a evaluar [100]. Su ventaja radica 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,

Análisis de la evaluación del desempeño de los modelos propuestos

lo que permite medir la diferencia de estructuras, entre la imagen original y la imagen distorsionada, teniendo en cuenta tres factores: luminancia, contraste y estructura, los dos primeros afectados principalmente por la iluminación del entorno, el tercero, la característica intrínseca del objeto [15].

Con los modelos objetivos seleccionados, se realizaron las mediciones a los dieciséis videos (tratamiento cero y 15 tratamientos adicionales) utilizando la herramienta MSU Video Quality Measurement Tool. Las mediciones obtenidas para PSNR y SSIM se pueden apreciar, respectivamente, en las TABLAS 21 y 22.

Los valores consignados en estas dos tablas no permiten ningún tipo de comparación, porque no están en una escala normalizada. Para poder comparar el desempeño entre las respuestas que entregan los modelos propuestos y las mediciones obtenidas con PSNR y SSIM, estos valores se deben mapear a la escala MOS. En la TABLA 23 se presenta el mapeo entre la escala MOS y los rangos de medición de PSNR [102] y SSIM [99].

Tabla 21. Resultados de la evaluación con PSNR

Tratamiento	PSNR		
	MVMB	MVMM	MVMA
0	45,116	38,10627	29,706
1	40,847	32,06711	29,706
2	40,326	32,38903	28,18
3	37,079	30,1646	23,946
4	41,268	29,48377	24,584
5	35,97	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	35,7	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	24,56

Tabla 22. Resultados de la evaluación con SSIM

Tratamiento	SSIM		
	MVMB	MVMM	MVMA
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	0,9
6	0,952	0,952	0,911
7	0,979	0,88	0,834
8	0,967	0,876	0,902
9	0,967	0,874	0,923
10	0,97	0,8	0,912
11	0,972	0,909	0,917
12	0,978	0,908	0,93
13	0,986	0,953	0,948
14	0,946	0,891	0,91
15	0,982	0,872	0,904

Tabla 23. Mapeo de escala MOS a PSNR y SSIM

MOS	PSNR (dB)	SIMM
Excelente (5)	> 37	> 0,99
Bueno (4)	31 - 37	$\geq 0,95$ & $< 0,99$
Justo / regular (3)	25 - 31	$\geq 0,88$ & $< 0,95$
Pobre / mediocre (2)	20 - 25	$\geq 0,50$ & $< 0,88$
Malo (1)	< 20	< 0,50

Para proceder con la comparación del desempeño entre los tres modelos propuestos y los dos modelos objetivos seleccionados (PSNR y SSIM), se aplicó el mapeo planteado en la TABLA 23 a las respuestas obtenidas de los dieciséis tratamientos de la TABLA 22. Posteriormente, se realizó un análisis de correlación entre los resultados de cada modelo propuesto y los valores arrojados por PSNR y SSIM, para así verificar el desempeño de cada modelo propuesto [15].

Este análisis de correlación consiste en calcular el coeficiente de correlación de Pearson (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 -1,00 indica una correlación negativa perfecta y +1,00 una correlación positiva perfecta.

En la FIGURA 23 se presentan los valores que puede tomar el coeficiente de correlación y su interpretación.



Figura 23. Interpretación de los valores que entrega el coeficiente de correlación de Pearson [86]

Análisis de la evaluación del desempeño de los modelos propuestos

A continuación se presenta el análisis para el modelo generado para cada tipo de video. En las TABLAS 24, 26 y 28 se incluyen los resultados para los videos Entrevista, Cafetería y Baloncesto, respectivamente, y en las TABLAS 25, 27 y 29, el correspondiente análisis estadístico de las respuestas obtenidas.

Para lograr el mapeo de las escalas PSNR y SSIM a MOS: se obtuvo la medida en PSNR y SSIM para cada cuadro del video analizado; se mapeó la medida obtenida para cada cuadro a escala MOS; y se calculó el promedio de los valores obtenidos en escala MOS para todos los cuadros que componen el video. Por esta razón, las medidas obtenidas no se expresan en números enteros, como se hizo en la TABLA 23.

MVMB - Entrevista

En la TABLA 24 se presentan los resultados para este video y en la TABLA 25 el resumen del análisis estadístico realizado a las respuestas obtenidas. Como se puede apreciar en ellas, los datos que arroja el modelo propuesto indican una correlación positiva fuerte, pues covarían en 71 % y 61 %, respectivamente, con las respuestas obtenidas con SSIM y PSNR.

Tabla 24. Datos obtenidos para evaluar el MVMB

Tratamiento	<i>Mean Opinion Score (MOS)</i>		
	MVMB	PSNR	SIMM
0	5	5	4,876
1	5	4,888	4,606
2	5	4,834	4,544
3	4	4,618	4,502
4	5	4,967	4,423
5	4	4,465	3,979
6	4	4,05	3,481
7	4	4,855	4,025
8	4	4,344	3,905
9	5	4,382	4,004
10	4	4,407	4,187
11	5	4,415	4,315
12	5	4,593	4,349
13	5	4,959	4,224
14	3	4,033	3,382
15	4	4,772	4,241

Tabla 25. Resultados estadísticos para comparar el desempeño del MVMB

Modelo	Promedio	σ	c.v.	CCP - PSNR	CCP - SIMM
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		

Asimismo, se puede notar que el coeficiente de variación del modelo propuesto es significativamente mayor que los coeficientes calculados para las respuestas obtenidas con PSNR y SSIM, lo que se debe a la manera de evaluar la QoE en 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.

MVMM - Cafetería

En la TABLA 26 se presentan los resultados para este video y en la TABLA 27 el resumen del análisis estadístico realizado a las respuestas obtenidas.

Tabla 26. Datos obtenidos para evaluar el MVMM

Tratamiento	<i>Mean Opinion Score (MOS)</i>		
	MVMM	PSNR	SIMM
0	4	4,776	4
1	4	4,154	3,598
2	4	4,029	3,593
3	4	3,216	3,199
4	4	3,311	2,929
5	4	3,286	3,112
6	4	3,855	3,51
7	4	3,133	2,66
8	4	3,083	2,357
9	4	3,183	2,232
10	3	2	2
11	4	3,274	2,909
12	4	3,232	3,004
13	4	3,78	3,456
14	4	3,112	3
15	4	3,332	2,531

Análisis de la evaluación del desempeño de los modelos propuestos

Tabla 27. Resultados estadísticos para comparar el desempeño del MVMM

Modelo	Promedio	σ	c.v.	CCP - PSNR	CCP - SIMM
MVMM	3,9375	0,25	6,35	0,623	0,485
PSNR	3,422	0,609	17,8		
SSIM	3,006	0,553	18,39		

Como se puede apreciar, el valor obtenido en el coeficiente de correlación indica una correlación positiva moderada $-62,3\%$ y $48,5\%$, respectivamente, para PSNR y SSIM– entre los resultados que entrega el modelo y las respuestas obtenidas con PSNR y SSIM.

MVMA - Baloncesto

Tal como se hizo para los videos anteriores, se presentan los resultados obtenidos para este video (TABLA 28) y el resumen del análisis estadístico realizado a las respuestas obtenidas (TABLA 29). Cabe recordar que para este modelo no aplica el tratamiento 0.

Tabla 28. Datos obtenidos para evaluar el MVMA

Tratamiento	<i>Mean Opinion Score (MOS)</i>		
	MVMA	PSNR	SIMM
1	5	3,996	3,548
2	4	3,419	3,232
3	4	2,685	2,639
4	4	2,643	3,062
5	4	2,303	2,776
6	4	2,714	3,116
7	4	2,178	2,232
8	4	2,627	2,851
9	4	3,174	2,975
10	4	2,954	2,971
11	4	2,71	2,913
12	4	3,141	3,299
13	5	3,934	3,539
14	4	2,71	2,942
15	4	2,689	2,763

Tabla 29. Resultados estadísticos para comparar el desempeño del MVMA

Modelo	Promedio	σ	c.v.	CCP - PSNR	CCP - SIMM
MVMA	3.333	0.617	18.52	0.749	0,572
PSNR	2,925	0.527	18.00		
SSIM	29.905	0.3392	11.34		

Como se puede apreciar, el cálculo del coeficiente de correlación indica una correlación positiva fuerte entre los resultados del modelo propuesto y los datos obtenidos por PSNR y SSIM –los datos que entrega el modelo propuesto covarían en un 74.9 % con respecto de los resultados obtenidos con PSNR y en un 57,2 % con respecto de los obtenidos con SSIM. La fuerte correlación positiva que se observa con PSNR se debe a las distorsiones causadas al momento de transmitir el video, lo que está relacionado directamente con la comparación pixel a pixel, que realiza PSNR al efectuar la medición. Adicionalmente, cabe recordar que el tratamiento cero no se tuvo en cuenta para estos análisis, por ser un dato atípico.

Desempeño: modelos propuestos vs pruebas subjetivas

Para verificar el desempeño de los modelos para evaluar QoE, los resultados obtenidos con ellos deben ser correlacionados con resultados obtenidos mediante pruebas subjetivas, para lo cual el grupo de expertos en calidad de video (VQEG) ha planteado un plan de pruebas que ha sido utilizado por sus investigadores [91].

La comparación entre los resultados obtenidos por los modelos propuestos y los obtenidos a partir de las pruebas subjetivas se realiza mediante una regresión no lineal que permite evaluar el desempeño del modelo propuesto teniendo en cuenta tres aspectos: la precisión de la predicción (*Prediction Accuracy*), la predicción de la monotocidad (*Prediction Monotonicity*) y la predicción de la consistencia (*Prediction Consistency*) [15].

La precisión de la predicción se representa por el coeficiente de correlación de Pearson (ver su lectura en la FIGURA 23) y por la raíz del error cuadrático medio (*Root Mean Square Error*, RMSE), definida por la ECUACIÓN 20, donde $X_j = 1, \dots, N$ y \bar{x} es la media del conjunto de datos.

$$RMSE = \sqrt{\frac{1}{N} \sum (X_i - \bar{X})^2} \quad (20)$$

La predicción de la monotocidad está representada por el coeficiente de correlación de Spearman –también conocido como Rho de Spearman–, el cual 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, de manera lineal o logarítmica, etc. El coeficiente de Spearman se calcula con base en parejas de datos X,Y y su valor se da en un rango de -1 a 1, donde -1 indica que X puede ser representada como una función monótona decreciente de Y, y 1 que X puede ser presentada como una función creciente monótona de Y [15].

La predicción de consistencia es representada por el porcentaje de valores atípicos (*Outlier Ratio*, OR), definido como el porcentaje del número de predicciones por fuera del rango de ± 2 desviaciones estándar de los resultados de las pruebas subjetivas [69]. Si N es el número total de datos y N' el número de valores atípicos, el porcentaje de valores atípicos está dado por: $OR = N'/N$.

En el proceso de comparación, para cada modelo propuesto se calculó: el Coeficiente de Correlación de Pearson (CCP); la Raíz del Error Cuadrático Medio (RECM); el Coeficiente de Correlación de Spearman (CCS); y el Porcentaje de Valores Atípicos (PVA). El detalle de los resultados de las pruebas de desempeño para cada modelo, se presenta a continuación.

MVMB - Entrevista

En la TABLA 30 se presentan los resultados obtenidos en escala MOS del MVMB-Entrevista y de los dieciséis tratamientos evaluados mediante pruebas subjetivas, en ella se puede apreciar que para el tratamiento cero el MVMB entregó un valor de MOS de 5,0 y las pruebas subjetivas su valor más alto (4,5). Asimismo, se observa que únicamente en el tratamiento 14 se obtuvo una calificación de 3,0 en el modelo propuesto, un valor similar al obtenido para el

mismo tratamiento con las pruebas subjetivas (3,004). De ello 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 buscaba que los resultados de las pruebas y coeficientes calculados fueran más altos.

Tabla 30. MVMB: resultados en escala de MOS y de Pruebas Subjetivas (PS)

Tratamiento	<i>Mean Opinion Score (MOS)</i>	
	MVMB	PS
0	5	4,500
1	5	4,061
2	5	4,096
3	4	3,474
4	5	3,632
5	4	4,228
6	4	2,667
7	4	3,825
8	4	3,491
9	5	4,333
10	4	3,509
11	5	4,237
12	5	4,342
13	5	3,693
14	3	3,044
15	4	4,211

En la TABLA 31 se presentan los datos estadísticos del modelo propuesto (MVMB) y los resultados de las pruebas subjetivas, como se puede observar, los valores obtenidos para la desviación estándar indican que los resultados no están tan dispersos de la media.

Tabla 31. Datos estadísticos de los resultados del MVMB y de las pruebas subjetivas

Modelo	Media	σ	c.v.
MVMB	4,438	0,629	14,18
PS	3,834	0,512	13,35

Análisis de la evaluación del desempeño de los modelos propuestos

Lo anterior se puede corroborar visualmente en la FIGURA 24, en ella se aprecian los resultados que arroja el MVMB cuando se evalúan los valores de los parámetros de QoS para los dieciséis tratamientos (sección a) y los resultados de las pruebas subjetivas para cada tratamiento (sección b). Las líneas rojas horizontales en la sección (a) de la figura indican el límite de dos desviaciones estándar por encima y por debajo del promedio de los datos obtenidos mediante las pruebas subjetivas –límites con los que se calcula el PVA–; por su parte, la línea horizontal roja en la sección (b) de la figura corresponde a la media de los datos. Aunque los coeficientes de variación del modelo propuesto y las pruebas subjetivas tienen un valor parecido, la variabilidad de las respuestas, en algunos casos, es diferente, como se puede observar en los tratamientos 3, 6 y 14.

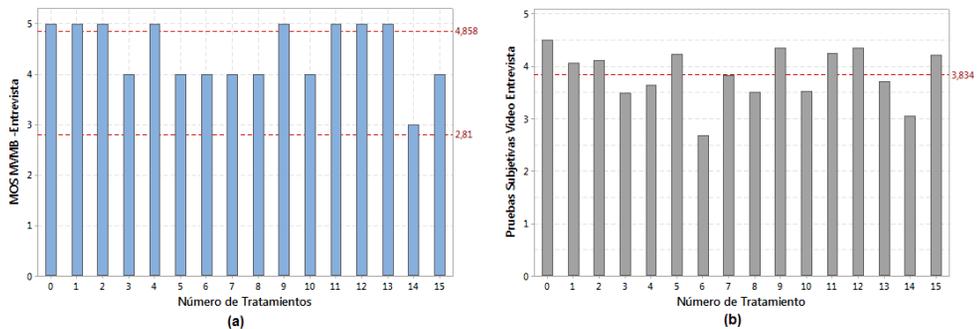


Figura 24. Medidas (escala MOS) obtenidas al evaluar los parámetros de QoS del MVMB (a) Medidas obtenidas al realizar las PS a los 16 tratamientos del video Entrevista (b)

Los resultados de las pruebas de desempeño del modelo (TABLA 32) muestran que la fidelidad del modelo es moderada por los valores obtenidos en el CCP y CCS, que indican que el modelo propuesto explica en alrededor de 63 % la variabilidad de los resultados obtenidos con las PS. Aunque el PVA es muy elevado –este 50 % indica que la mitad de las respuestas está por fuera de dos desviaciones estándar de la media de las pruebas subjetivas–, se debe tener en cuenta que esto puede ocurrir debido a la escala de mapeo diseñada.

Tabla 32. MVMB: resultados de pruebas de desempeño

Aspecto a evaluar	MVMB
CCP	0,624
CCS	0,630
RECM	0,607
PVA	0,5

MVMM - Cafetería

En la TABLA 33 se presentan los resultados en escala MOS del MVMM para el video Cafetería y los resultados de los dieciséis tratamientos evaluados mediante pruebas subjetivas. Respecto de estas últimas, se observa cómo para este tipo de video se obtuvieron calificaciones en un rango que va de 2,1 hasta 4,2, lo que indica que quienes participaron en las pruebas detectaron distorsiones en el video al momento de su evaluación. Respecto de los resultados obtenidos con el modelo, solo el tratamiento 10 obtuvo una calificación de 3,0 –todos los demás obtuvieron un 4.0–, lo que se debe al diseño de la escala que mapea VQM a MOS, donde cerca del 50% de ella correspondía a una calificación de 4,0 en MOS.

Tabla 33. MVMM: resultados en escala de MOS y de Pruebas Subjetivas (PS)

Tratamiento	<i>Mean Opinion Score (MOS)</i>	
	MVMM	PS
0	4	4,211
1	4	4,123
2	4	3,509
3	4	2,640
4	4	3,044
5	4	3,544
6	4	3,939
7	4	2,939
8	4	3,298
9	4	2,693
10	3	2,193
11	4	3,640
12	4	3,009
13	4	3,798
14	4	2,947
15	4	3,079

En la TABLA 34 se presentan los datos estadísticos del modelo propuesto y los resultados de las pruebas subjetivas, donde se puede apreciar que el coeficiente de variación de los resultados es bajo (c.v.=6,35%), lo que se debe a la escala diseñada para este modelo. De manera consecuente, la desviación estándar es mucho más baja en comparación con obtenida en las pruebas subjetivas.

Análisis de la evaluación del desempeño de los modelos propuestos

Tabla 34. Datos estadísticos de los resultados del MVMM y de las pruebas subjetivas

Modelo	Media	σ	c.v.
MVMM	3,9375	0,25	6,35
PS	3,288	0.569	17,32

En la FIGURA 25 se aprecian los resultados que arroja el MVMM cuando se evalúan los valores de los parámetros de QoS para los dieciséis tratamientos (sección a) y los resultados de las pruebas subjetivas para cada tratamiento (sección b). En la figura: las líneas rojas horizontales en la sección (a) indican el límite de dos desviaciones estándar por encima y por debajo del promedio de los datos obtenidos mediante las pruebas subjetivas, con los que se calcula el PVA; y la línea horizontal roja en la sección (b) corresponde a la media de los datos. Además, en la figura se puede apreciar una variabilidad reducida de los resultados obtenidos por el video Cafetería en comparación con los resultados de las pruebas subjetivas, lo que, como se mencionó, se debe al diseño de la escala para mapear y a que las personas que evaluaros este tipo de video fueron más críticas o detectaron mayores distorsiones al momento de calificar el video.

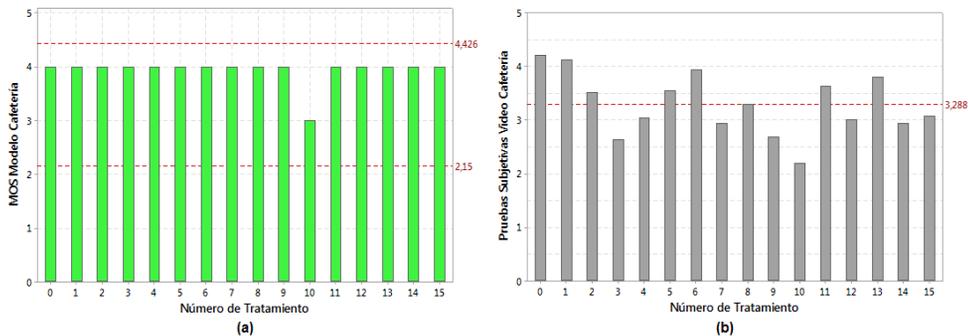


Figura 25. Medidas (escala MOS) obtenidas al evaluar los parámetros de QoS del MVMM (a) Medidas obtenidas al realizar las PS a los 16 tratamientos del video Cafetería (b)

En los resultados de las pruebas de desempeño de este modelo (TABLA 35) se puede apreciar un CCP de 0,513, que indica una correlación positiva moderada, y un PVA de 0, que se debe a la escala propuesta para mapear los resultados del modelo con la escala MOS. La RECM es baja en comparación con la del MVMB, debido a la escala de mapeo propuesta, la cual para este modelo hace que las respuestas que entrega tengan un coeficiente de variación bajo.

Tabla 35. MVMM: resultados de pruebas de desempeño

Aspecto a evaluar	MVMM
CCP	0,513
CCS	0,420
RECM	0,24
PVA	0,0

MVMA - Baloncesto

En la TABLA 36 se presentan los resultados obtenidos en escala MOS del MVMA y los resultados de los quince tratamientos evaluados mediante pruebas subjetivas, en ella se puede observar cómo en el tratamiento 1 se obtiene el valor más alto de calificación en ambas pruebas, y se puede apreciar que un 73% de las calificaciones tiene valores en el rango de 3,0 a 3,9, lo que indica que las personas que evaluaron el video detectaron con mayor facilidad distorsiones. El modelo propuesto entrega una calificación de 5,0 en los tratamientos 1 y 13 –donde la pérdida de paquetes es cero–, lo que corrobora el efecto significativo de la pérdida de paquetes, en la QoE para este tipo de videos.

Tabla 36. MVMA: resultados en escala de MOS y de Pruebas Subjetivas (PS)

Tratamiento	<i>Mean Opinion Score (MOS)</i>	
	MVMM	PS
1	5	4,896
2	4	4,430
3	4	3,070
4	4	3,123
5	4	3,044
6	4	3,229
7	4	2,798
8	4	3,254
9	4	3,553
10	4	3,254
11	4	3,228
12	4	3,211
13	5	4,746
14	4	3,307
15	4	3,474

Análisis de la evaluación del desempeño de los modelos propuestos

En la TABLA 37 se presentan los datos estadísticos del modelo propuesto y los resultados de las pruebas subjetivas, en ella se puede apreciar cómo en los resultados obtenidos para el modelo la desviación estándar es más baja que la obtenida con las pruebas subjetivas, lo que se relaciona con el coeficiente de variación obtenido por los datos arrojados por el modelo (c.v.=8,51%).

Tabla 37. Datos estadísticos de los resultados del MVMA y de las pruebas subjetivas

Modelo	Media	σ	c.v.
MVMA	4,133	0,351	8,51
PS	3,508	0,643	18,33

En la FIGURA 26 se aprecian los resultados que arroja el MVMM cuando se evalúan los valores de los parámetros de QoS para los quince tratamientos (sección a) y los resultados de las pruebas subjetivas para cada tratamiento (sección b). Tal como se indicó para las FIGURAS 24 y 25, las líneas rojas horizontales en la sección (a) de la figura indican el límite de dos desviaciones estándar por encima y por debajo del promedio de los datos obtenidos mediante las pruebas subjetivas –límites con los que se calcula el PVA– y la línea horizontal roja en la sección (b) de la figura corresponde a la media de los datos. Se puede apreciar además la baja variabilidad en las respuestas que entrega el MVMA-Baloncesto, lo que se asocia con el c.v. obtenido y está ligado con el diseño de la escala de mapeo entre VQM y MOS.

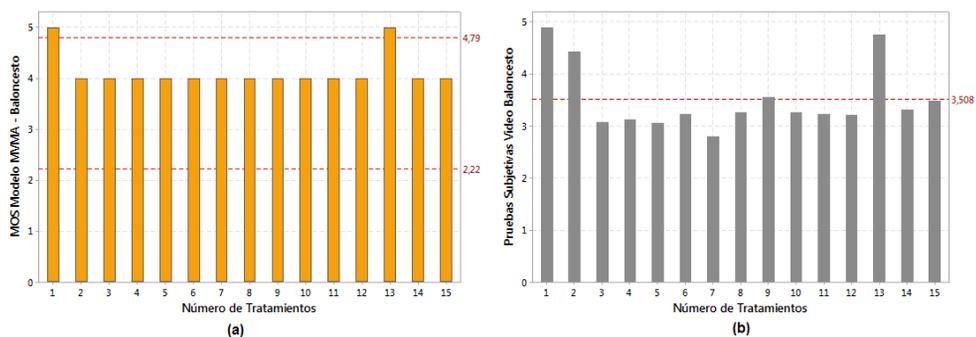


Figura 26. Medidas (escala MOS) obtenidas al evaluar los parámetros de QoS del MVMA (a) Medidas obtenidas al realizar las PS a los 15 tratamientos del video Baloncesto (b)

En la TABLA 38 se presentan los resultados de las pruebas de desempeño del MVMA, donde se puede apreciar que el CCP y el CCS entregan una correlación positiva moderada –más alta la obtenida con el CCP, que llega al 82,9%) y un PVA de 13%, un valor bajo que junto con el CCP obtenido indican que el modelo posee una fidelidad alta (en la siguiente sección se presentarán, de manera conjunta, las pruebas con PSNR y SSIM que se le realizaron al modelo).

Tabla 38. MVMA: resultados de pruebas de desempeño

Aspecto a evaluar	MVMA
CCP	0,829
CCS	0,591
RECM	0,3415
PVA	0,13

Pruebas estáticas

Estas pruebas se realizaron con el fin de analizar el comportamiento de los modelos ante diferente cantidad de tratamientos, para su desarrollo se generaron valores aleatorios para dos parámetros de QoS y se mantuvo un valor fijo para el otro, en mil tratamientos.

Para el MVMB, se dejó constante el retardo en 1300 ms. En la FIGURA 27 se puede apreciar la influencia de la pérdida de paquetes en la QoE que entrega el modelo, pues para valores inferiores a 0.1%, se puede obtener una QoE superior a 4 en escala MOS. Asimismo, se observa que para valores inferiores a 0.005 ms en la variación del retardo, la QoE cae drásticamente, hasta 2 en escala MOS, lo que 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.

Para el MVMM, la variación del retardo se dejó configurada en 0.018 ms. En la FIGURA 28 se ve cómo para valores de pérdida de paquetes cercanos a 0.18 y de retardo superiores a 1100 ms, la QoE del video baja a niveles entre 1 y 2 en la escala MOS, mientras que para valores inferiores a 0.1% en pérdida de paquetes y retardo menores a 600ms, la QoE del video se puede mantener en un rango de 4 a 5 –bueno y excelente, respectivamente– en esa

Análisis de la evaluación del desempeño de los modelos propuestos

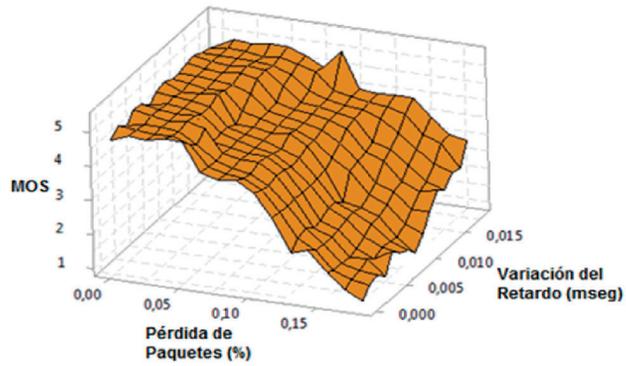


Figura 27. MOS vs P_p y V_p para el MVMB

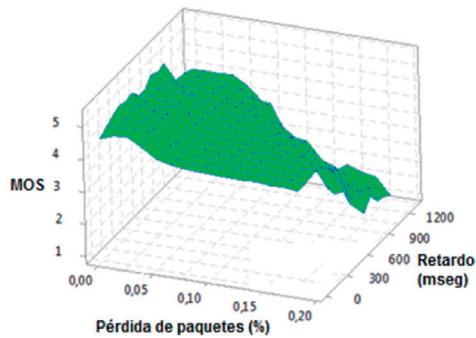


Figura 28. MOS vs P_p y R para el MVMM

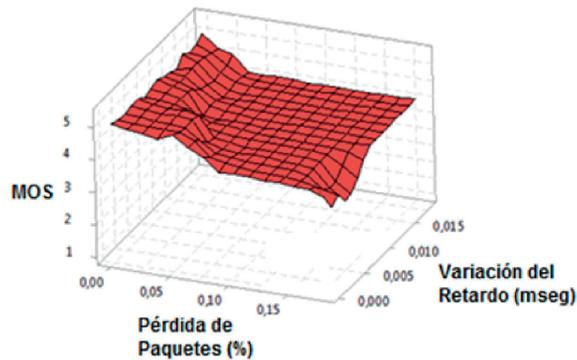


Figura 29. MOS vs P_p y V_R para el MVMA

misma escala. Este resultado es importante para los proveedores del servicio porque les permite ajustar los parámetros de QoS en los rangos descritos y así mantener la QoE del video en un nivel superior.

Para el MVMA, tomando en consideración lo establecido en la Recomendación UIT-T Y.1541 [59], el retardo se dejó constante en 400 ms. En la FIGURA 29 se puede observar la influencia de la pérdida de paquetes en el modelo, a partir de valores superiores a 0.1%, el MOS obtenido empieza a descender hasta un valor de 3. La notable influencia de la pérdida de paquetes en la QoE de este tipo de video se debe a su alto movimiento, por ello, solo con valores muy bajos de pérdida de paquetes –inferiores a 0.05%– se consiguen medidas de MOS superiores a 4.

Acerca de los resultados

En la TABLA 39 se presenta un resumen de las pruebas realizadas. Respecto de los modelos objetivos, el análisis de los resultados en CCP-PSNR y CCP-SSIM permite afirmar que tienen una correlación positiva moderada y que quien mejor explica la variabilidad de los datos es el MVMA; asimismo, que el MVMM muestra una correlación positiva débil.

Tabla 39. Resumen de las pruebas de desempeño de los modelos propuestos

Parámetro / Modelo		MVMB	MVMM	MVMA
Modelos objetivos	c.v.	14,18	6,35	8,51
	CCP-PSNR	0,624	0,623	0,802
	CCP-SSIM	0,607	0,485	0,662
Pruebas subjetivas	CCP	0,624	0,513	0,829
	CCS	0,630	0,420	0,591
	RECM	0,607	0,24	0,341
	PVA	0,5	0	0,13

Sin embargo, al analizar los resultados de desempeño respecto de las pruebas subjetivas, se encuentra que:

- el MVMA es el que mejor explica la variabilidad –lo hace en un 82,9%, como corresponde a su CCP de 0,829– y tiene un PVA bajo (13 %); y

Análisis de la evaluación del desempeño de los modelos propuestos

- el MVMM es el de menor fidelidad, su CCP de 0,513 indica una correlación positiva moderada de tan solo un 51,3%, lo que se puede explicar en la mayor rigurosidad con que los sujetos evaluaron las distorsiones en el video en las pruebas subjetivas –al comparar los resultados que entrega el modelo y las pruebas subjetivas, se notan las diferencias que refleja 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 y el análisis de los resultados obtenidos, ya que un diseño inadecuado de una prueba de este tipo puede generar conclusiones erradas respecto de la precisión y de la fidelidad de los modelos propuestos.

Los resultados de las pruebas estáticas, por su parte, son consistentes con los análisis realizados en los capítulos 3 y 4, no se encontraron comportamientos atípicos en las respuestas que entrega cada modelo y se corroboraron los resultados que entrega el MVMA, con medidas máximas de 4,0 en las mil iteraciones realizadas.

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 permiten que los proveedores de servicio estén continuamente evaluando la QoE asociada al video y tomen acciones correctivas: ofrecer mayor ancho de banda o mejorar los esquemas de QoS que ofrecen al usuario.

Dado que el objetivo es que el modelo propuesto sea utilizado por los proveedores del servicio de IPTV, a continuación se plantea el protocolo de pruebas que les permitirá utilizarlo en su plataforma de servicios:

- se deben utilizar los videos asociados al tratamiento cero I00, C00 y B00 (ver TABLA 4) para realizar las medidas de QoE en la red del proveedor;
- en el *Set-Top-Box* (STB) se debe instalar una aplicación que permita ejecutar los modelos obtenidos asociados a cada tipo de video y verificar que sea posible obtener la medición de los parámetros de QoS en él;
- se envían los videos I00, C00 y B00 hacia el usuario donde se desea realizar la medida de QoE por la red, cuando el proveedor lo estime conveniente;
- se calcula la QoE asociada a cada video y la información se envía al nodo central de la plataforma de gestión para su procesamiento; y,

- si se presentan inconvenientes en la instalación de la aplicación en el STB, debido a que son arquitecturas cerradas y propietarias, una opción es extraer los valores de los parámetros de QoS utilizando *Simple Network Management Protocol* (SNMP) y con ellos calcular el valor de la QoE asociada al video mediante una aplicación web en la nube, utilizando el modelo propuesto.

Con el fin de que otros investigadores puedan reproducir el proceso usando otros parámetros de QoS, a continuación se describen los pasos para la obtención del modelo (ver diagrama de flujo en la FIGURA 30).

- Paso 1: escoger si grabaran los videos con recursos propios (paso 1A) o si seleccionaran de una base de datos pública (paso 2).
- Paso 3: seleccionar el diseño del experimento, esto es fundamental porque permite definir el tipo de modelo a obtener, el número de tratamientos a ejecutar y los rangos de las variables a usar.
- Paso 4: seleccionar el número y tipo de parámetros de QoS –u otro tipo de parámetros de red– que se incluirán.
- Paso 5: definir la distorsión de los videos dependiendo de los recursos disponibles, si se cuenta con recursos de hardware y software, se puede implementar una maqueta de pruebas (paso 5A) y distorsionar el video de diferentes maneras, mientras que si solo se cuenta con recursos de software, las distorsiones se deben realizar con una herramienta de simulación (paso 6).
- Paso 7: seleccionar el método –subjetivo u objetivo– para evaluar la QoE asociada al video.
- Paso 8: ejecutar el método elegido en el paso 7 e iniciar la recolección de los datos.
- Paso 9: analizar los datos utilizando herramientas estadísticas.
- Paso 10: realizar pruebas de desempeño al modelo preliminar obtenido.
- Paso 11: elegir el tipo de evaluación de desempeño que se le realizará al modelo propuesto: pruebas subjetivas (paso 11A) o comparación con métodos objetivos (paso 11B).
- Paso 12: establecer el umbral de desempeño analizando factores tales como CCP y CCS.

Si se obtiene una métrica mayor o igual a 0.5, se ha obtenido la versión final del modelo (paso 13), si no, el procedimiento se debe repetir desde el paso 3.

Análisis de la evaluación del desempeño de los modelos propuestos

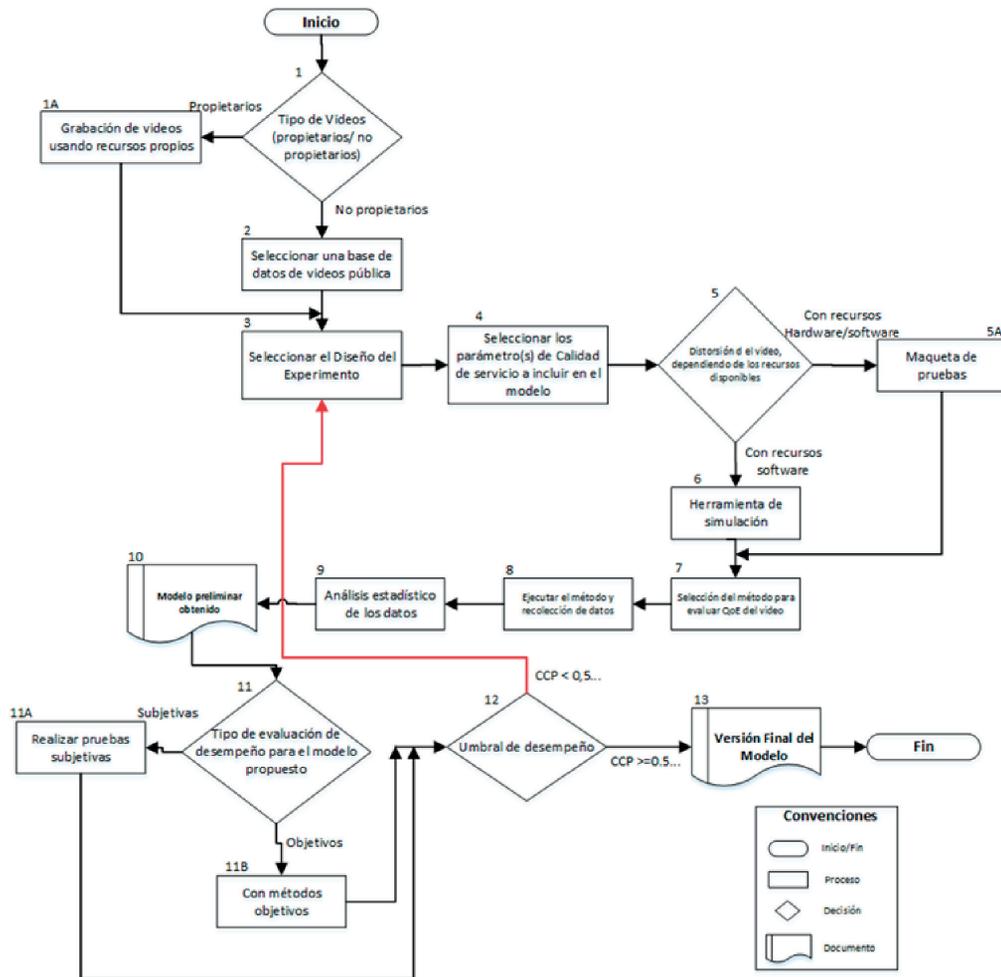


Figura 30. Algoritmo usado para obtener el modelo de medición de QoE en IPTV

Conclusiones, lecciones aprendidas y trabajo futuro

Conclusiones

El modelo propuesto se obtuvo con base en la distorsión realizada a videos al momento de su transmisión en un entorno de emulación, este enfoque tiene una serie de ventajas sobre otras propuestas, en las cuales la distorsión de los videos se realiza al archivo del video sin transmitirlo en un entorno de red, es decir, de manera estática, afectando características del archivo elegidas de manera específica. Este es el mayor aporte de esta investigación a la solución del problema: la transmisión en línea del video y su distorsión de manera aleatoria, su enfoque es novedoso y difiere de otras propuestas que basan sus modelos en pruebas subjetivas y utilizan un número inferior de parámetros de QoS.

El modelo propuesto puede ser utilizado en un ambiente real de red –por ejemplo, por un proveedor de servicios– para medir de manera continua y en línea, la QoE del servicio que está prestando, algo de suma importancia pues permite estar ajustando continuamente los procesos de QoS al interior de sus redes de núcleo.

Para la implementación de este modelo en un proveedor de servicio se recomienda instalar una aplicación que permita ejecutar el modelo propuesto y obtener las mediciones de QoE asociadas al video en los STB del proveedor; sin embargo, si esto no es posible, valiéndose del SNMP o de una aplicación de gestión propietaria, puede 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 robusto, presenta una alta correlación

Conclusiones, lecciones aprendidas y trabajo futuro

en pruebas realizadas en diferentes publicaciones y es usado por un ente de estandarización (ANSI) 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 de tres expertos evaluaba las distorsiones causadas en los videos, verificando que fueran perceptibles para el usuario y que los videos se pudieran decodificar en el lado del cliente.

Para mejorar la fidelidad de los modelos es necesario mejorar y automatizar la maqueta de pruebas en dos aspectos: la elección de un número mayor de videos con diferentes características espacio-temporales y en formato YUV; y el uso de una herramienta de simulación en la maqueta de pruebas. Estos dos aspectos permitirán aumentar el número de pruebas, con el fin de analizar los datos obtenidos y comparar sus resultados con los obtenidos en esta primera versión del modelo propuesto; adicionalmente, se podría obtener un modelo genérico que permita medir cualquier tipo de video.

Las pruebas subjetivas son necesarias para verificar la fidelidad y el desempeño de nuevos modelos propuestos, pero se debe planear muy bien su ejecución y hacerlo siguiendo las recomendaciones de la UIT. Los resultados de las pruebas subjetivas permiten la creación de bases de datos de video para futuros estudios y para la prueba de nuevas propuestas, y pueden ser muy útiles para la comunidad de investigadores y para los proveedores del servicio. Cabe mencionar que con los videos utilizados en las pruebas subjetivas durante esta investigación, se diseñó una base de datos compuesta por 64 videos –cuatro secuencias originales y sesenta videos con distorsiones–, una base de datos única en su tipo, pues tiene en cuenta las distorsiones causadas por la variación de los parámetros de QoS y cuenta con videos transmitidos en un ambiente de red emulado.

Con base en los análisis de varianza de los datos asociados a cada tipo de video, se puede apreciar que la pérdida de paquetes es el parámetro de QoS con mayor influencia en la QoE asociada al video. El nivel de degradación de la calidad del video depende del tipo de cuadro que se pierda, el caso más crítico es la pérdida de marcos tipo I. Por su parte, otros parámetros de QoS, como retardo y variación del retardo, afectan la medición de la QoE asociada al video, pero su efecto se puede contrarrestar mediante la configuración del tamaño de la memoria intermedia en el cliente.

El MVMA presentó la mayor fidelidad, como demuestra su mayor coeficiente de correlación de Pearson (82,9%) cuando se comparó con las pruebas subjetivas. Esto se debe a la mayor facilidad para detectar las distorsiones del video, por parte de las personas encargadas de evaluar su calidad en las pruebas subjetivas. Lo contrario ocurrió con el MVMM, donde la complejidad de la imagen para el usuario hacía más difícil la detección de las distorsiones del video, debido a que se enfocaba en analizar solo parte de la imagen.

Lecciones aprendidas

Uno de los aspectos fundamentales de esta investigación fue el diseño del experimento para obtener los datos necesarios para generar los modelos propuestos, así se evidenció durante su desarrollo y al momento de analizar los datos obtenidos –de los videos transmitidos en la maqueta de experimentación y en las pruebas subjetivas realizadas–. Por ello, el diseño de experimento se debe planificar cuidadosamente antes de empezar la toma de datos, hacerlo evita inconvenientes durante su desarrollo y permite la oportuna toma de acciones correctivas.

Otro aspecto a tener en cuenta es la transcodificación que se lleva a cabo en el proceso de transmisión de los videos, al transcodificarlos, ellos van perdiendo resolución, nitidez o calidad, haciendo que las medidas de QoE varíen, por lo que 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. Se debe contar entonces con videos de excelente calidad, para así minimizar su deterioro al transcodificarlos para transmitirlos o cambiarles el formato para su edición. Lo anterior va ligado con la cámara utilizada para grabar los videos de prueba, si se utiliza una cámara profesional, este tipo de situación se puede mejorar, ya que se podrían tener los videos en formato YUV.

Trabajo futuro

El trabajo a futuro se pueden enfocar en varias líneas:

- Grabar videos de prueba con cámaras profesionales, para así tener un mayor control al momento de transcodificar cada video y contar con

Conclusiones, lecciones aprendidas y trabajo futuro

videos en formato YUV. Esto es importante pues la transcodificación es una variable no controlada que influye en el resultado de los datos obtenidos para generar los modelos.

- Implementar una maqueta de experimentación más robusta, que permita realizar pruebas de manera automática y con mayor precisión que las desarrolladas en esta investigación. Esto es importante porque así se reduciría el tiempo de realización de las pruebas. Además, al mejorar las condiciones de la maqueta de experimentación, es posible mejorar la fidelidad de los modelos propuestos y obtener un modelo genérico para ser utilizado en videos con diferentes características espacio-temporales. La implementación de una maqueta de experimentación más robusta hace referencia a dos aspectos: utilizar un número mayor de videos con diferentes características espacio-temporales para ser transmitidos en la maqueta de experimentación; e implementar una maqueta donde las pruebas se puedan realizar de manera automática.
- Implementar un protocolo de pruebas para definir con mayor precisión los rangos de las características espacio-temporales para la utilización de cada modelo al momento de evaluar la QoE. Hacer esto permitiría ampliar el rango de utilización de los modelos propuestos.
- Realizar pruebas de validación de desempeño del modelo propuesto en comparación con los modelos que se presentaron en la sección "Trabajos relacionados" del capítulo 2, con el fin de analizar la fidelidad y el desempeño del modelo propuesto. Durante la parte final de esta investigación, dichas pruebas de desempeño no se realizaron porque el trabajo se concentró en validar la 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 contar con herramientas de simulación.
- Continuar alimentando la base de datos diseñada, con nuevos videos con diversas distorsiones, no solamente con distorsiones causadas por la modificación de parámetros de QoS. La actualización de esta base es fundamental porque algunas de las bases encontradas están desactualizadas o la información técnica que entregan es escasa.

Referencias

- [1] A. A. García, y J. C. Cuellar, “Calidad de servicio en proveedores de servicios IPTV”, *Ingenium*, vol. 6, no. 13, pp. 11-25, sep. 2012.
- [2] *Cisco visual networking index: Forecast and methodology, 2017-2022*, enero 18, 2017 [en línea]. Disponible en: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html#_Toc484813971
- [3] *ITU-T Recommendation G.1080: Quality of experience requirements for IPTV services*, dic. 2008.
- [4] A. Takahashi, D. Hands, y V. Barriac, “Standardization activities in the ITU for a QoE assessment of IPTV”, *IEEE Communications Magazine*, vol. 46, no. 2, pp. 78-84, mar. 2008.
- [5] W. Robitza, A. Ahmad, P. A. Kara, L. Atzori, M. G. Martini, A. Raake y L.Sun, “Challenges of future multimedia QoE monitoring for internet service providers”, en: *Multimedia Tools and Applications*, vol. 76. no. 21. pp. 22243-22246, 2017.
- [6] A. R. Reibman, V. A. Vaishampayan, y Y. Sermadevi, “Quality monitoring of video over a packet network”, *IEEE Transactions on Multimedia*, vol. 6, no. 2, pp. 327-334, mar. 2004.
- [7] N. Liao y Z. Chen, “A packet-layer video quality assessment model with spatiotemporal complexity estimation”, *EURASIP Journal on Image and Video Processing*, ago. 2011, [en línea]. Disponible en: <https://jivp-urasipjournals.springeropen.com/track/pdf/10.1186/1687-5281-2011-5>

Referencias

- [8] C. J. Pardo, “Revisión Sistemática de la Armonización de Marcos en la Mejora de los Procesos de Software”, tesis de maestría, Universidad Castilla - La Mancha, 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)”, *Información Tecnológica*, vol. 25, no. 5, pp. 121-128, may. 2014.
- [10] T. Rahrer, R. Fiandra, y S. Wright, “Triple-play services quality of experience (QoE) requirements”, TR-126, DSL Forum, 13-dic-2006.
- [11] *ITU-R Recommendation ITU-R BS.775-3, Multichannel stereophonic sound system with and without accompanying picture*, ago. 2012.
- [12] *ITU-R Recommendation BS.1286. Methods for the subjective assessment of audio systems with accompanying picture*, oct. 1997.
- [13] *ITU-R Recommendation BT.500-13, Methodology for subjective assessment of the quality of television pictures*, ene. 2012.
- [14] *ITU-T Recommendation P.910, Subjective video quality assessment methods for multimedia applications*, abril, 2008.
- [15] Y. Chen, K. Wu, y Q. Zhang, “From QoS to QoE: A Tutorial on Video Quality Assessment”, *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 1126-1165, Oct 2015.
- [16] *UIT-T Recomendation P.800 Métodos de determinación subjetiva de la calidad de transmisión*, ago. 1996.
- [17] H. Rifai, S. Mohammed, y A. Mellouk, “A brief synthesis of QoS-QoE methodologies”, en *2011 10th International Symposium on Programming and Systems (ISPS)*, abr. 2011, pp. 32-38.
- [18] S. Winkler, “Video quality measurement standards - Current status and trends”, en *7th International Conference on Information, Communications and Signal Processing, ICICS 2009*, dic. 2009, pp. 1-5.
- [19] S. Winkler y P. Mohandas, “The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics”, *IEEE Transactions on Broadcasting*, vol. 54, no. 3, pp. 660-668, sep. 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”, *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”, *Signal Processing: Image Communication*, vol. 25, no. 7, pp. 482–501, ago. 2010.
- [22] J. C. Cuellar, Ortiz y Arciniegas. “Métodos para medir calidad de experiencia para el Servicio de IPTV: clasificación y análisis”, *Informacion Tecnologica*, vol. 25 no. 5, pp. 121-128. Jan. 2014
- [23] S. Daly, “The visible differences predictor: An algorithm for the assessment of image fidelity», en A. B. Watson [Ed.], *Digital images and human vision*, Cambridge, MA: MIT Press, 1993, pp. 179–206.
- [24] J. Lubin y D. Fibush, “Sarnoff JND vision model”, TIA 1.5 Working Group Document #97-612, ANSI TI Standards Committee, 1997.
- [25] C. J. Van den Branden Lambrecht y O. Verscheure, “Perceptual quality measure using a spatiotemporal model of the human visual system”, *Digital Video Compression: Algorithms and Technologies*, pp. 450-461, ene. 1996.
- [26] S. Winkler, “A perceptual distortion metric for digital color video”, in *Proceedings of SPIE*, Whashington, D.C., 1999, pp. 175–184.
- [27] Z. Wang, L. Lu, y A. C. Bovik, “Video Quality Assessment Based on Structural Distortion Measurement”, *Signal Process: Image Communication*, vol. 19, no. 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 Transactions on Image Processing*, vol. 15, no. 11, pp. 3440-3451, nov. 2006.
- [29] Z. Wang y L. Qiang, “Video Quality Assessment Using a Statistical Model of Human Visual Speed Perception”, *Journal of the Optical Society of America*, vol. 24, no. 12, pp. B61-B69, dic. 2007.
- [30] K. Seshadrinathan y A. C. Bovik, “Motion-based Perceptual Quality Assessment of Video”, en *Proceedings SPIE – Human Vision and Electronic Imaging XIV*, vol. 7240, feb. 2009.
- [31] D. M. Chandler y S. S. Hemami, «VSNR: A Wavelet-Based Visual Signal-to-Noise Ratio for Natural Images», *IEEE Transactions on Image Processing*, vol. 16, no. 9, pp. 2284-2298, sep. 2007.

Referencias

- [32] M. H. Pinson y S. Wolf, “A new standardized method for objectively measuring video quality”, *IEEE Transactions on Broadcasting*, vol. 50, no. 3, pp. 312-322, sep. 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)*, Miami, FL, dic. 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)*, Beijing, China, oct. 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*, mar. 2004.
- [36] *ITU-T Recommendation J.148. Requirements for an Objective Perceptual Multimedia Quality Model*, may. 2003.
- [37] *ITU-T Recommendation G.1070. Opinion model for video-telephony applications*, jun. 2018.
- [38] *ITU-T Recommendation P.1202. Parametric non-intrusive bitstream assessment of video media streaming quality (PNBAMS)*, oct. 2012.
- [39] A. G. Davis, D. Bayart, y D. S. Hands, “Quality assurance for IPTV”, en *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, Bilbao, Spain, may. 2009, pp. 1-7.
- [40] T. Hossfeld, D. Hock, P. Tran-Gia, K. Tutschku, y M. Fiedler, “Testing the IQX hypothesis for exponential interdependency between QoS and QoE of voice codecs iLBC and G.711”, en *18th ITC Specialist Seminar on Quality of Experience*, Karlskrona, Sweden, May 2008.
- [41] M. Fiedler, T. Hossfeld, y P. Tran-Gia, «A generic quantitative relationship between quality of experience and quality of service», *IEEE Network*, vol. 24, no. 2, pp. 36-41, 2010.
- [42] M. Fiedler y T. Hossfeld, “Quality of Experience-Related Differential and Provisioning-Delivery Hysteresis”, en *21st Int. Teletraffic Congr. Spec. Semin. Multimed. Appl. Perform. QoE*, Tokio, Japón, 2010.
- [43] T. Hoßfeld, M. Fiedler, y T. Zinner, “The QoE provisioning-delivery-

- hysteresis and its importance for service provisioning in the Future Internet”, en *2011 7th EURO-NGI Conference on Next Generation Internet Networks*, Kaiserslautern, Germany, jun. 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”, en *2008 Fourth International Conference on Networked Computing and Advanced Information Management*, Gyeongju, South Korea, sep. 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”, en *2010 The 12th International Conference on Advanced Communication Technology (ICACT)*, Phoenix Park, South Korea, feb. 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”, en *2012 14th International Conference on Advanced Communication Technology (ICACT)*, PyeongChang, South Korea, feb. 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”, en *2006 First International Symposium on Pervasive Computing and Applications*, Urumqi, China, ago. 2006, pp. 178-182.
- [48] K. Yamagishi, T. Kawano, y T. Hayashi, “Hybrid Video-Quality-Estimation Model for IPTV Services”, *IEEE Global Telecommunications Conference, 2009 -GLOBECOM 2009*, pp. 1-5, dic. 2009
- [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 video streaming Applications over Wireless Mesh Networks”, en *XXIX Simpósio Bras. Redes Comput. E Sist. Distrib.*, 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”, en *2011 IEEE Global Telecommunications Conference -GLOBECOM 2011*, Kathmandu, Nepal, dic. 2011, pp. 1-6.
- [51] *Sigmoid Function* [en línea], Disponible en: http://research.omicsgroup.org/index.php/Sigmoid_function
- [52] Z. Wang, A. C. Bovik, H. R. Sheikh, y E. P. Simoncelli, “Image quality

Referencias

- assessment: from error visibility to structural similarity”, *IEEE Transactions on Image Processing*, vol. 13, no. 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”, en *2012 IEEE Global Communications Conference (GLOBECOM)*, Anaheim, CA, dic. 2012, pp. 1963-1968.
- [54] M. Elhoseny, A. Shehab, y L. Osman, “An empirical analysis of user behavior for P2P IPTV workloads”, en A. Hassanien, M. Tolba, M. Elhoseny, y M. Mostafa, (Eds), *The International Conference on Advanced Machine Learning Technologies and Applications (AMTLA2018)* [AISC, vol. 723, Cham, Suiza: Springer, 2018, pp. 252-263.
- [55] A. Shehab, M. . El Aziz, y A. . Hassanien, “Efficient schemes for playout latency reduction in P2P-VoD systems”, en A. Hassanien, y D. Oliva, (Eds.), *Advances in soft computing and machine learning in image processing* [SCI, vol 730], Cham, Suiza: Springer, 2018, pp. 447-495.
- [56] A. Shehab, M. Elhoseny, y A. Hassanien, “An efficient scheme for video delivery in wireless networks”, en A. Hassanien, M. Elhoseny, y J. Kacprzyk (Eds.), *Quantum Computing: An environment for intelligent large scale real applications* [SBD, vol. 33], Cham, Suiza: Springer, 2018, pp. 447-495.
- [57] C. M. Lentisco, L. Bellido, J. C. Cuellar, E. Pastor, y J. L. Arciniegas, “QoE-Based Analysis of DASH Streaming Parameters Over Mobile Broadcast Networks”, *IEEE Access*, vol. 5, pp. 20684-20694, sep. 2017.
- [58] *ITU-T Recommendation Y.1540. IP Packet Transfer and Availability Performance Parameters*, jul. 2016.
- [59] *ITU-T Recommendation Y.1541. Network performance objectives for IP-based services*, dic. 2011.
- [60] G. Baltoglou, E. Karapistoli, y P. Chatzimisios, “IPTV QoS and QoE measurements in wired and wireless networks”, en *2012 IEEE Global Communications Conference (GLOBECOM)*, Anaheim, CA, dic. 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*, 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)* [en línea], Disponible en: <http://www.vcodex.com/h264avc-context-adaptive-binary-arithmetic-coding-cabac/>
- [65] *H.264/AVC Context Adaptive Variable Length Coding* [en línea], disponible en: <http://www.vcodex.com/h264avc-context-adaptive-variable-length-coding/>
- [66] *Introducción a la compresión: Huffman y Entropía* [en línea], disponible en: <http://www.grc.upv.es/docencia/tdm/practicass/P1.pdf>
- [67] A. Mittal, M. A. Saad, y A. C. Bovik, “A completely blind video integrity Oracle”, *IEEE Transactions on Image Processing*, vol. 25, no. 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 Transactions on Image Processing*, vol. 19, no. 6, pp. 1427-1441, jun. 2010.
- [69] S. Chikkerur, V. Sundaram, M. Reisslein, y L. J. Karam, “Objective video quality assessment methods: A classification, review, and performance comparison”, *IEEE Transactions on Broadcasting*, vol. 57, no. 2, pp. 165-182, jun. 2011.
- [70] S. Wolf, y M. Pinson, “Video quality measurement techniques”. National Telecommunications and Information Administration (NTIA). Washington, DC, 2002.
- [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”, *Second Workshop on Digital Media and its Application in Museum Heritages (DMAMH 2007)*, Chongqing, China, dic. 2007, pp. 365-369.
- [73] L. Cruvinel y T. Vazao, “A simple metric for predicting video quality of experience”, *10° Conf. Sobre Redes Comput. CRC 2010*, Universidade do Minho, Braga, Portugal., pp. 77-82.
- [74] H. J. Kim y S. G. Choi, “QoE assessment model for multimedia streaming services using QoS parameters”, *Multimedia Tools and Applications*, vol. 72, no. 3, pp. 2163-2175, oct. 2014.

Referencias

- [75] R. Lübke, P. Büschel, D. Schuster, y A. Schill, “Measuring accuracy and performance of network emulators”, in *2014 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, Odessa, Ukraine, may. 2014, pp. 63-65.
- [76] I. Sobel y G. Feldman, “A 3x3 isotropic gradient operator for image processing, Stanford Artificial Intelligence Project (SAIL)”. Stanford, CA., 1968.
- [77] W. Song y D. W. Tjondronegoro, “Acceptability-Based QoE Models for Mobile Video”, *IEEE Transactions on Multimedia*, vol. 16, no. 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 Journal of Selected Topics in Signal Processing*, vol. 6, no. 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”, *Multimedia Tools and Applications*, vol. 75, no. 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”, *Multimedia Tools and Applications*, vol. 75, no. 23, pp. 16461-16485, dic. 2016.
- [81] C. Wang, X. Jiang, y Y. Wang, “Video Quality Assessment Models for IPTV Services”, *International Journal of Digital Content Technology and its Applications (JDCTA)*, vol. 7, no. 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)*, London, UK, jun. 2015 pp. 1723-1728.
- [83] D. Montgomery, *Diseño y análisis de experimentos*, Ciudad de México: Limusa, 2013.
- [84] *MSU Video Quality Measurement Tool*, dic. 2016, [en línea], disponible en: http://compression.ru/video/quality_measure/video_measurement_tool.html
- [85] H. Gutierrez y R. De la Vara, *Análisis y diseño de experimentos*, Madrid, España: McGraw-Hill Interamericana, 2012.

- [86] D. A. Lind, W. G. Marchal, y S. A. Wathen, *Estadística aplicada a los negocios y a la economía*, Ciudad de México: McGrawHill., 2012.
- [87] *Minitab* [en línea], Disponible en: <http://www.minitab.com/es-mx/>
- [88] A. N. Moldovan, I. Ghergulescu, y C. H. Muntean, “VQAMap: A novel mechanism for mapping objective video quality metrics to subjective MOS Scale”, *IEEE Transactions on Broadcasting*, vol. 62, no. 3, pp. 610-627, sep. 2016.
- [89] J. Gareth, D. Witten, T. Hastie, y R. Tibshirani, *An introduction to statistical learning*, New York, NY: Springer.
- [90] U. Engelke, M. Kusumab, H-J. Zepernicka, y M. Calderac “Reduced-reference metric design for objective perceptual quality assessment in wireless imaging”, *Signal Processing: Image Communication*, vol. 24, no. 7, pp. 525-547, ago. 2009.
- [91] Video Quality Experts Group, “Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment, phase II”. VQEG, Boulder, CO, 2003.
- [92] *What’s a Good Value for R-Squared?* [en línea], disponible en: <https://people.duke.edu/~rnau/rsquared.htm>
- [93] «VQEG Objective Video Quality Model Test Plan.» may-1998.
- [94] P. Corriveau, A. Webster, A. Rohaly y J. Libert, “Video Quality Experts Group, the quest for valid objective methods”, *Human Vision and Electronic Imaging, Proceedings of SPIE*, vol. 3959, pp. 129-139, 2000.
- [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”, in *2016 XI International Symposium on Telecommunications (BIHTEL)*, Sarajevo, Bosnia-Herzegovina, oct. 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.

Referencias

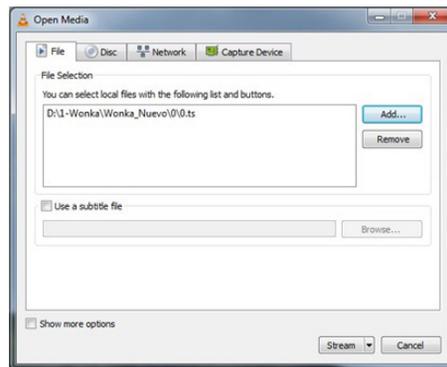
- [98] Y. A. Syahbana, Herman, A. A. Rahman, y K. A. Bakar, “Aligned-PSNR (APSNR) for Objective Video Quality Measurement (VQM) in video stream over wireless and mobile network”, en *2011 World Congress on Information and Communication Technologies*, Mumbai, India, dic. 2011, pp. 330-335.
- [99] T. Zinner, 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)*, Trondheim, Norway, jun. 2010, pp. 29-34.
- [100] Z. Wang y A. C. Bovik, “A universal image quality index”, *IEEE Signal Processing Letters*, vol. 9, no. 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, no. 11, pp. 1044-1055, jul. 2004.
- [103] *x-rite pantone* [en línea], disponible en: <http://www.xrite.com/categories/calibration-profiling/colormunki-display>

Anexo

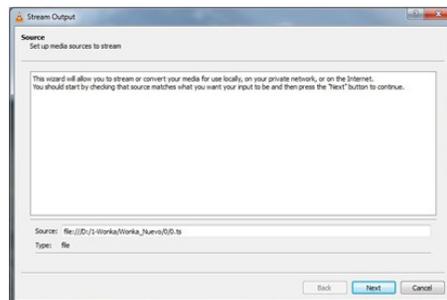
Configuración de VLC y NetEM

Configuración de VLC

1. En la barra superior del menú seleccione *Media>Streamy*. En la pantalla que aparece, seleccione el video a enviar presionando el botón *Add*.

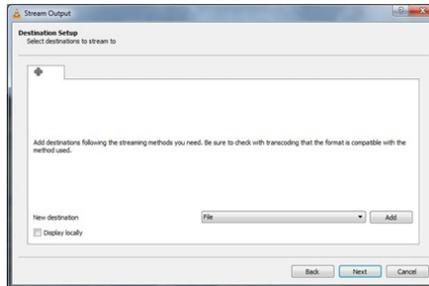


2. Presione *Stream*, aparecerá una nueva pantalla. Presione *Stream*, aparecerá una nueva pantalla.

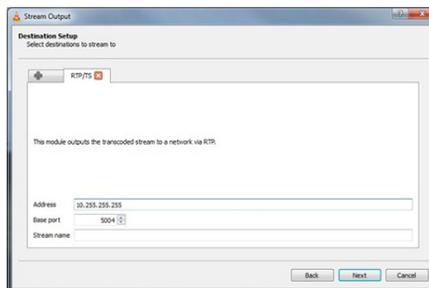


Anexo: configuración de VLC y NetEM

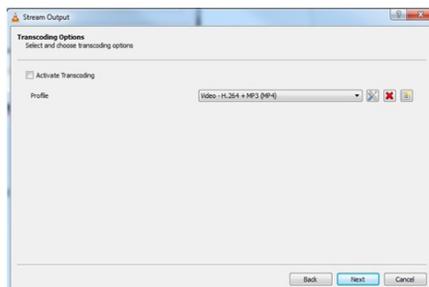
3. Presione *Next*, en la nueva pantalla presione *File*, elija la opción *RTP/MPEG Transport Stream* y asegúrese de que la opción *Display locally* esté chequeada.



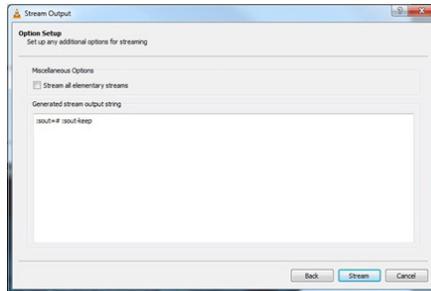
4. Presione *Add*, al hacerlo una nueva pantalla le permitirá configurar el puerto por donde se enviará el video (*Base port*) y la dirección de broadcast o multicast para su envío (*Address*).



5. Una vez configurados, presione *Next*, en la pantalla que aparece elija *Video - H.264 + MP3 (MP4)* de la opción *Profile* y asegúrese de tener deshabilitada la opción *Activate Transcoding*. Eso reducirá el consumo de recursos de procesamiento en el servidor, pues la transcoding, 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, no se requiere en este caso, pues solo se va a analizar el stream transmitido..



6. Presione Next, aparecerá una nueva pantalla, en ella presione Stream para iniciar la transmisión.



7. En el cliente VLC, en una ventana de comandos, digite: `vlc -vvv 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 ser recibido.
8. Luego del envío y la grabación del archivo en el cliente, se procede a la edición de los archivos con el fin de recortarlos a diez segundos y proceder con la medición de VQM con la herramienta MSU Video Quality Measurement Tool.

Configuración de NetEM

Con el fin de optimizar el uso de esta herramienta, se generaron tres *scripts* al hacer las pruebas, los cuales se encuentran en el CD adjunto, en la carpeta NetEm.

Los archivos que se deben instalar en la raíz de la sesión de administrador son: *EmuRed*, *ConfigPuente.shy* y *DetenerPuente.sh*. *Emured* es el archivo principal que se encarga de ejecutar *ConfigPuente.shy* y *DetenerPuente.sh*, que son los archivos en cargados de configurar el servidor en modo puente y de detener el modo puente del servidor al finalizar las pruebas se encargan. Adicionalmente *EmuRed* se encarga de recibir del usuario los valores de los parámetros de QoS que se van a configurar.

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% se muestra en la imagen de la página siguiente.

Anexo: configuración de VLC y NetEM

```
jcc@jcc-Vostro-230: ~/Netem
jcc@jcc-Vostro-230:~/Netem$ ls
ConfigPuenete.sh  DetenerPuenete.sh~  NetemManualScript
ConfigPuenete.sh~  EmuRed              RegistroTiempo.ods
DetenerPuenete.sh  NetemAutoScript    ScriptPruebaRed.ods
jcc@jcc-Vostro-230:~/Netem$ ./EmuRed
SCRIPT PARA EMULACIÓN DE RED
Selecclone 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 ns. 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 Cualquter 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'.
```


Este libro se terminó de imprimir y encuadernar en diciembre de 2018 en los talleres de Carvajal Soluciones de Comunicación, en la ciudad de Bogotá D.C. En su preparación, realizada desde la Editorial Universidad Icesi, se emplearon los tipos Gill Sans MT de 8, 10, 14 y 26 puntos, Baskerville MT Std de 9, 10, 12 y 30 puntos, Times New Roman de 10 puntos y Cambria Math de 12 puntos. La edición, que consta de 50 ejemplares, estuvo al cuidado de Claros Editores S.A.S.

Monitorear continuamente la calidad de la experiencia (QoE) en la transmisión de video en un servicio de televisión por IP (IPTV) es una actividad fundamental para garantizar la satisfacción del usuario. Si bien la comunidad de investigadores ya ha propuesto diferentes modelos que cumplen adecuadamente con este cometido, este libro propone una aproximación diferente, cuyo resultado es un modelo con ventajas prácticas: se puede realizar en línea sin necesidad de decodificar el video, no recurre a pruebas subjetivas y requiere una menor complejidad computacional que otros modelos basados en métodos objetivos.



ISBN: 978-958-8936-56-7

