

QoE-CAC: métrica de QoE em Sistemas de Comunicação Imersiva Consciente de Atenção e Comunicação

Rosana de Oliveira Santos, Boanerges Cipriano G. Junior, Carlos Eduardo da S. Santos, Ciro J. A. Macedo e Antonio Carlos de Oliveira-Jr.

Resumo— Sistemas de Comunicação Imersiva (SCI) abrangem serviços promissores, incluindo aplicações como jogos virtuais online, reuniões virtuais e outras. A viabilidade dessas aplicações depende, fundamentalmente, da Qualidade de Experiência (QoE), que exige condições rigorosas de recursos de Redes de Computadores, como latência ultrabaixa e altas taxas de transmissão de dados, para proporcionar uma experiência verdadeiramente imersiva. Este artigo propõe uma nova métrica de QoE, denominada QoE-CAC, que considera fatores de comunicação e de atenção do usuário à imagem, integrados em uma única métrica, contribuindo para guiar futuras estratégias de alocação de recursos de redes em SCI, com visão centrada no usuário.

Palavras-Chave— Sistemas de Comunicação Imersiva, QoE, recursos, *downlink*, *uplink*, latência.

Abstract— Immersive Communication Systems (ICS) encompass promising services, including applications such as online virtual games, virtual meetings, and others. The viability of these applications fundamentally depends on the Quality of Experience (QoE), which requires stringent Computer Network resource conditions, such as ultra-low latency and high data transmission rates, to provide a truly immersive experience. This paper proposes a new QoE metric, called QoE-CAC, which considers communication and user attention to the image factors, integrated into a single metric, contributing to guide future network resource allocation strategies in ICS, with a user-centric view.

Keywords— Immersive Communication Systems, QoE, resources, *downlink*, *uplink*, latency.

I. INTRODUÇÃO

A comunicação imersiva integra amplas interações entre usuários e aplicativos em ambientes altamente conectados, culminando no Metaverso. Este novo cenário incorpora tecnologias como Realidade Virtual (RV), Realidade Aumentada (RA) e Realidade Mista (RX), que fazem uso dos recursos de redes para processamento de dados. A integração dessas tecnologias com a infraestrutura de comunicação Pós-Quinta Geração de Redes Móveis Celulares (5G) permitirá avanços significativos

em diversos setores da sociedade. No campo da medicina, estas tecnologias podem possibilitar a realização de simulações precisas para treinamento cirúrgico, bem como diagnósticos e tratamentos mais eficazes por meio de visualizações tridimensionais detalhadas [1]. No setor de entretenimento, a combinação de RV, RA e RX com alta velocidade e baixa latência proporcionará experiências imersivas e interativas, elevando o nível de realismo em jogos, eventos esportivos e produções cinematográficas [2]. Na Indústria 4.0, essas tecnologias poderão facilitar a implementação de sistemas de manutenção preditiva, treinamento remoto e otimização de processos produtivos, resultando em maior eficiência e redução de custos [3].

No entanto, ainda existem desafios significativos relacionados à infraestrutura de comunicação necessária para viabilizar a comunicação imersiva. Sistemas imersivos dependem de tecnologias de rede avançadas para assegurar comunicações de alta qualidade em tempo real, utilizando vídeos em formatos como *point cloud* ou *light field* para criar cenários tridimensionais (3D) ultrarrealistas, semelhantes a hologramas [4].

O elevado custo de armazenamento e transmissão da mídia volumétrica representa uma limitação para o processo de comunicação, demandando grande largura de banda e latência extremamente baixa, o que torna as redes móveis atuais inadequadas para atender a esses requisitos. Assim sendo, a implementação de sistemas imersivos exige uma infraestrutura robusta capaz de compartilhar, em tempo real, imagens 3D e outros tipos de mídia, como vídeos e áudios, entre múltiplos usuários. [5], [6]

Para garantir uma experiência imersiva satisfatória, é essencial que esses sistemas capturem, renderizem e transmitam dados de forma eficiente nos vários pontos da comunicação. As exigências por baixa latência, alta largura de banda e estabilidade no *downlink*, aliadas à necessidade de manter a alta qualidade das imagens transmitidas, representam desafios críticos a serem superados. Outro fator que limita o desempenho dos Sistemas de Comunicação Imersiva (SCI) é a renderização em tempo real, que intensifica ainda mais as restrições de latência. Nesse contexto, a investigação de estratégias para alocação eficiente de recursos de infraestrutura de comunicação, além de técnicas de otimização de *streaming*, torna-se um tema de grande relevância.

Para enfrentar esses desafios, estudos recentes têm explorado novos modelos de avaliação de QoE que consideram tanto o consumo de recursos de rede quanto os fatores cognitivos

Rosana de Oliveira Santos, Instituto de Informática, Universidade Federal de Goiás, Goiânia-GO, e-mail: rosanasantos@discente.ufg.br; Boanerges Cipriano G. Junior, Instituto de Informática e Instituto de Matemática e Estatística, Universidade Federal de Goiás, Goiânia-GO, e-mail: boanergescg@discente.ufg.br; Carlos Eduardo da S. Santos, Instituto Federal de Educação, Ciência e Tecnologia do Tocantins, Campus Palmas, Palmas-TO, e-mail: carlosedu@iftto.edu.br; Ciro J. A. Macedo, Instituto Federal de Goiás, Campus Cidade de Goiás, Cidade de Goiás-GO, e-mail: ciro.macedo@ifg.edu.br; Antonio Carlos de Oliveira-Jr, Instituto de Informática, Universidade Federal de Goiás, Goiânia-GO e Fraunhofer Portugal AICOS, Porto, Portugal, e-mail: antoniojr@ufg.br.

e perceptuais que influenciam a atenção do usuário durante a interação com o conteúdo 3D [7]. As métricas de QoE são empregadas não apenas para avaliar o desempenho de aplicativos de comunicação imersiva, mas também para identificar oportunidades de aprimoramento. Seu valor vai além de medir a satisfação do usuário, servindo como um direcionador para os Provedores de Infraestrutura (InP) e Provedores de Serviços Virtuais (VSP). Com o objetivo final de melhorar a QoE, esses provedores podem planejar e implementar estratégias otimizadas para alcançar esse propósito.

Este trabalho propõe uma métrica de QoE, analisando a influência de indicadores-chave para SCIs relacionados à qualidade da imagem, como capacidade de renderização e fatores associados ao canal de comunicação, incluindo *downlink*, *uplink Bit Error Probability* (BEP) e latência, conforme ilustrado na Figura 1. A avaliação de QoE em SCI ainda é um tema pouco explorado, exigindo mais pesquisas e experimentações para impulsionar avanços na área. Especificamente, considerando os sistemas de comunicação imersiva e os recursos de redes de computadores utilizados para transmissão de vídeos imersivos, pretendemos responder às duas perguntas a seguir:

- É possível modelar uma métrica objetiva para QoE considerando a influência dos fatores de comunicação de rede e fatores subjetivos de atenção do usuário à imagem?
- Qual o impacto da QoE percebida pelo usuário nas aplicações imersivas e como ele pode orientar a alocação de recursos de rede?

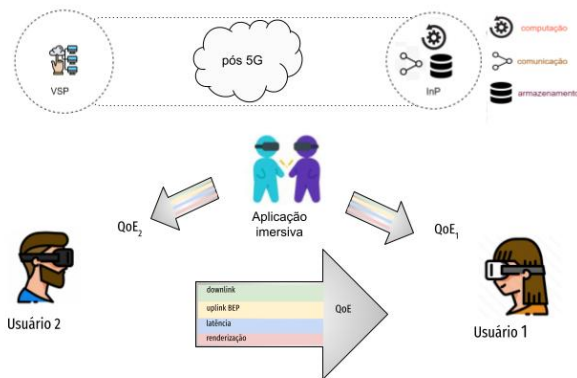


Fig. 1. A QoE dos usuários em aplicações de comunicação imersiva e seus fatores de influência

As contribuições desse trabalho de pesquisa são:

- definição de uma nova métrica objetiva de QoE para sistema de mídia volumétrica na perspectiva do usuário e com base nos recursos computacionais e de comunicação utilizados, além da qualidade do vídeo;
- realização de experimentos com a coleta de tráfego para validar a métrica em um ambiente realista, analisando o impacto da variação dos parâmetros de rede ao longo do tempo.

Para além desta Introdução, a Seção II apresenta os trabalhos relacionados. Na Seção III, a métrica de QoE Consciente de Atenção e Comunicação (QoE-CAC) é apresentada em detalhes, com destaque para os principais *Key Performance*

Indicators (KPI) de rede: *downlink*, *uplink* BEP e prejuízo da latência. A Seção IV discute os experimentos e resultados, enquanto a Seção V apresenta as considerações finais e os próximos passos do trabalho.

II. TRABALHOS RELACIONADOS

Diversos trabalhos discutem os problemas envolvidos em cenários de comunicação imersiva, especialmente no contexto de aplicações que combinam RA e RV em ambientes altamente interativos. Os sistemas de mídia volumétrica oferecem desafios quanto à transmissão de dados, tendo em vista que consomem altas taxas de transmissão, além de requerer ultra-baixa latência e sincronização rigorosa dos fluxos concorrentes. Para enfrentar esses desafios e dar suporte à próxima geração de sistemas de comunicação imersivos, a literatura tem trabalhado com abordagens e estratégias divididas em três camadas, conforme [4]: a camada de rede (com propostas de novas arquiteturas de *network slicing* e redes virtualizadas distribuídas), a camada de transporte e aplicação (com técnicas para compressão dos dados, bufferização e caching) e a camada de usuário final, onde são associadas técnicas para predição de movimento do usuário, segmentação espacial 3D e otimização na avaliação de QoS e QoE.

Os trabalhos de [8], [9] e [10] focam na compressão de dados como estratégia para reduzir a elevada taxa de transmissão da mídia imersiva. No entanto, esse método pode impactar negativamente a latência, uma vez que os processos de compressão e descompressão exigem um tempo considerável de processamento, o que compromete aplicações que dependem de transmissão em tempo real.

Um caminho a ser trilhado em busca de respostas para os desafios da comunicação imersiva diz respeito às pesquisas de como melhorar a QoE do usuário, de modo a proporcionar/garantir as características inerentes ao Metaverso. A QoE orienta o fluxo de dados a ser transmitido de maneira a garantir a qualidade suficiente para que o usuário se mantenha conectado e satisfeito durante a execução do serviço. O tema de avaliação de QoE em SCI ainda é pouco explorado e muito ainda precisa ser pesquisado e experimentado a fim de contribuir para o avanço na área. Vale destacar as recomendações ITU-T G.1035, G.1036 e P.1320 [11], [12], [13] que discutem os fatores de influência para serviços imersivos, bem como o *white paper* QUALINET [14] que oferece diretrizes e metodologias em busca dos estudos de métricas objetivas e subjetivas de QoE em sistemas imersivos, estendendo a noção de Qualidade de Serviço (QoS) centrada em rede e sistemas multimídia, adicionando o conceito de QoE, centrada no usuário.

A comunidade QUALINET [14] categoriza os fatores que influenciam a QoE imersiva em três grupos distintos: Fatores de Influência do Sistema (do inglês *System Influence Factors* (SIF)), Fatores de Influência do Contexto (do inglês *Context Influence Factors* (CIF)) e Fatores de Influência Humana (do inglês *Human Influence Factors* (HIF)). Esses fatores devem ser particularmente pesquisados conforme situações típicas de cada serviço imersivo e se referem às características próprias de um usuário, sistema, serviço, aplicativo ou contexto cujo

estado ou configuração real pode ter influência na experiência de mídia imersiva do usuário [14].

Grande parte dos trabalhos realiza avaliações e testes subjetivos através de questionários *Mean Opinion Score* (MOS) [15], com pesquisa de satisfação de usuários em aplicações de RV, RA e RX. Por exemplo, [16] realiza um estudo comparativo entre avaliações subjetivas (MOS) e objetivas, utilizando as métricas objetivas como o PSNR (*Peak Signal-to-Noise Ratio*) e SSIM (*Structured Similarity Index*). Há também aqueles que utilizam *Machine Learning* para estabelecer predição de valores de QoE a partir de questionários MOS [17]. Esses testes subjetivos geralmente seguem metodologias padrão, conforme as recomendações da ITU [11], [12], [13] e representam uma importante ferramenta para demonstrar a visão do usuário. Todavia, seu uso de forma isolada possui desvantagens, podendo introduzir vies. Portanto, testes subjetivos são mais eficazes quando complementados com métricas objetivas e análise técnica detalhada [17] e [15].

Poucos trabalhos estipulam métricas objetivas de QoE associadas à aplicação imersiva e que avaliam de forma ampla a visão do usuário, ampliando a noção de QoS centrada em rede. Em [18], os autores propõem um modelo generalizado de QoE para aplicações do tipo RA/RV com base no prejuízo de alguns KPI, incluindo taxa de congelamento, resolução de conteúdo e latência total. Porém, a modelagem não considera aspectos de contexto de influência humana em sua fórmula e não detalha a análise de experimentos práticos associados a esta modelagem. O mesmo autor contribui com um excelente trabalho em [19], o qual apesar de não fornecer explicitamente métricas de QoE, estabelece uma taxonomia clara e didática para os SCI.

Em [20] os autores propõem uma abordagem inovadora para a alocação de recursos com base na atenção do usuário, visando otimizar a QoE em serviços metaverso de comunicação ultra-confiável de baixa latência - (do inglês *Ultra-Reliable Low Latency Communications* (URLLC)). O trabalho introduz uma métrica denominada *Meta-Immersion* (MI), que combina KPIs objetivos, como taxa de transmissão e probabilidade de erro, com valores subjetivos de atenção dos usuários aos objetos virtuais. Os autores utilizam a teoria dos contratos para estabelecer um equilíbrio entre o InP e o VSP do metaverso, maximizando a utilidade de ambos os lados. Embora o esquema de alocação adaptativa de recursos demonstre ganhos significativos na QoE, em comparação com abordagens tradicionais de URLLC, os autores ignoram fatores adversos do ambiente, como congestionamento de rede e variabilidade de largura de banda, comuns em aplicações reais. Em eventos culturais, por exemplo, com um grande número de espectadores simultaneamente utilizando dispositivos de realidade imersiva, a latência e a largura de banda limitada poderiam degradar significativamente a QoE.

Assim, [20] avalia somente o impacto da renderização (sem avaliar os fatores de rede) frente ao cálculo de QoE; já que considera os parâmetros de *downlink* e *uplink* BEP como sendo ótimos; o trabalho de [20] é importante e merece destaque, todavia, carece de uma análise realista, envolvendo os KPIs de rede. Sua análise de performance e resultados apresentados basearam-se em comparação com dados analíticos e simulação

Monte Carlo. Além disso, [20] não considerou o prejuízo da latência na métrica de QoE. A análise da latência é crucial, pois ela oferece um impacto direto na QoE, principalmente quando avaliada a sensação de imersão e a náusea digital (*cybersickness*), quando envolve o uso de vestíveis, como os displays montados na cabeça (do inglês *Head-Mounted Display* (HMD)). Além disso, os aplicativos que envolvem comunicações em tempo real impõem requisitos de latência rigorosos no fluxo transmitido. Logo, incluir o KPI da latência na métrica de QoE é fundamental.

Neste contexto, este artigo propõe uma métrica de QoE, denominada QoE Consciente de Atenção e Comunicação (QoE-CAC), para sistemas imersivos com transmissão de vídeo 3D, baseada na correlação entre variáveis de comunicação e fatores atencionais, como a percepção de movimento e a nitidez da imagem. A métrica proposta abrange os principais KPIs de rede e o fator de renderização para avaliar a qualidade da imagem, correspondendo ao fator subjetivo de atenção do usuário em relação ao serviço utilizado.

É importante salientar que a modelagem da métrica QoE-CAC incorpora, em um único número real, fatores objetivos e subjetivos da experiência, evidenciando de forma mais ampla os aspectos de avaliação do usuário relacionados ao serviço (experiência subjetiva), bem como os aspectos relacionados ao ambiente (infraestrutura) necessários para manter a aplicação/serviço específico. Esta abordagem é extremamente importante, pois a métrica QoE-CAC poderá direcionar a qualidade do fluxo de dados a ser fornecido pela infraestrutura de rede, permitindo capturar uma visão mais completa da experiência do usuário, associada à visão da infraestrutura do sistema, desta forma, auxiliando a guiar uma possível orquestração de recursos, ao mesmo tempo que monitora a qualidade de experiência vivenciada pelo usuário.

III. MODELAGEM DA MÉTRICA DE QOE PROPOSTA

A modelagem da métrica QoE-CAC foi realizada considerando um cenário de aplicação imersiva que reúne dois ou mais usuários $\mathcal{U} = \{1, 2, \dots, u, \dots, U\}$, consumindo vídeos sob a cobertura da estação 5G. Aqui estamos particularmente interessados em aplicações que envolvam fluxo de comunicação em tempo real, por exemplo, um show musical imersivo ao vivo. Durante a transmissão, a QoE é calculada com base nos valores coletados de *downlink*, *uplink* BEP e latência. A renderização ocorre no HMD e a capacidade de renderização é definida proporcionalmente à atenção do usuário ao objeto do vídeo. Definimos a métrica QoE-CAC conforme a Equação 1 a seguir.

$$QoE-CAC(u) = (1 - \tau(I_u^{Lat})) \times \tau(R_u^D) \times (1 - \tau(E_u)) \times \sum_{n=1}^{N_{ou}} K_{n,u} \ln \left(\frac{P_{n,u}^{(R)}}{P^{(R)}} \right) \quad (1)$$

onde, durante cada transmissão de vídeo para um usuário u , I_u^{Lat} é o valor associado ao prejuízo da latência durante a transmissão, R_u^D é a taxa de *downlink*, E_u é a taxa de *uplink* BEP no intervalo conforme Tabela I, $K_{n,u}$ é o valor da atenção do usuário u ao objeto n , $P^{(R)}$ é a capacidade

de renderização mínima disponível, $P_{n,u}^{(R)}$ é a capacidade de renderização R para o usuário u visualizar o objeto n contido na mídia imersiva (imagem); $N_{O,u}$ denota o número de objetos virtuais que o usuário u visualiza em um serviço. A lei de

TABELA I
FATORES DE INFLUÊNCIA

parâmetro	descrição	fator	mínimo	máximo
Lat_u	Latência	SIF	10 ms	20 ms
R_u^D	Downlink	SIF	10 Mbits/s	42 Mbits/s
E_u	Uplink BEP	SIF	10^{-8}	10^{-2}
$P_{n,k}^R$	Rendering	HIF	15K	$ obj * 20$

Weber–Fechner [21] é aqui utilizada, para refletir a intensidade do estímulo no coeficiente de renderização [20]. Para ajustar as grandezas, os valores foram normalizados para eliminar o efeito de magnitude, conforme a Equação 2, em que t_{min} e t_{max} são os limites mínimo e máximo, respectivamente, fornecidos para cada uma das métricas da Equação 1.

$$\tau(t) = \frac{t - t_{min}}{t_{max} - t_{min}} \quad (2)$$

Para refletir o nível de degradação na QoE de sistemas imersivos, aproveitamos os métodos de imparidade ITU-T onde, um KPI_{I_x} é prejudicado por $I_x \in [0, 1]$ e $I_x = 0$ significa maior qualidade e $I_x = 1$ significa pior qualidade, isto é, prejuízo total [18]. O prejuízo da latência I_u^{Lat} é calculado conforme a Equação 3, em que o prejuízo $I_u^{Lat} \in [0, 1]$, com $I_u^{Lat} = 0$ significando maior qualidade (zero prejuízo) e $I_u^{Lat} = 1$ significando pior qualidade, isto é, prejuízo total [18]. Ainda conforme [18], as constantes a e b na Equação 3 são utilizadas para diferenciar dependências específicas entre aplicações, por exemplo, para serviço do tipo Jogos Virtuais, a latência pode ser mais rigorosa do que para uma Viagem Virtual. As constantes a e b na Equação 3 utilizadas foram $a = 0.5$ e $b = 5$.

$$I_u^{Lat} = \frac{1 + e^{-b}}{1 + e^{b \frac{t-a}{a}}}, \text{ onde } t \text{ é a latência total} \quad (3)$$

IV. EXPERIMENTOS E RESULTADOS

As etapas do experimento realizado são detalhadas passo-a-passo no Algoritmo 1. Utilizamos dois PCs com sistema operacional Ubuntu 22.04 LTS, com 8GB de memória RAM: um denominado cliente e outro servidor, conectados via *access point*. Para codificação do cálculo da QoE, a linguagem de programação MatLab R2022b foi utilizada. Inicialmente, para simular a transmissão com uso de óculos de RA, efetuamos o pré-processamento, gerando os frames de vídeos a partir das sequências de imagens do *dataset* UOAL¹. No lado do servidor, os vídeos são gerados e transmitidos para o cliente; neste momento a coleta de tráfego é feita, derivando os indicadores de serviço que irão compor a métrica QoE-CAC proposta. Para a coleta de tráfego, seguimos a mesma metodologia² dos autores de [22], com o uso do FFmpeg para geração dos vídeos e Gstreamer³ para transmissão dos

Algorithm 1 Etapas do experimento

- 1: **Entrada:** *Dataset* UOAL, valores K de predição do usuário para os objetos das imagens, capacidade de renderização $P_{n,u}^{(R)}$ para objetos n do usuário u , capacidade mínima de renderização $P^{(R)}$
- 2: Pré-processamento: **Geração dos vídeos**
- 3: **for** $u = 1$ to 50 **do**
- 4: **Transmissão do vídeo para usuário** u
- 5: **Coletar KPIs:** *downlink*, *uplink* BEP, latência
- 6: **return** $QoE-CAC(u)$
- 7: **end for**

mesmos. O Tshark foi utilizado para coleta de tráfego de rede e o mesmo foi salvo no formato PCAP. Para avaliação dos resultados, comparamos o valor e o comportamento da QoE-CAC com a QoE Optimal (H.Du) [20] que foca em avaliar o impacto da taxa de renderização dos objetos de atenção frente à QoE do usuário, considerando os parâmetros de comunicação como sendo ótimos, resultando conforme exige a Equação 4.

$$QoE\ Optimal(u) = \sum_{n=1}^{N_{O,u}} K_{n,u} \ln \frac{P_{n,u}^{(R)}}{P^{(R)}} \quad (4)$$

Na Figura 2 temos as duas estratégias ilustradas. Observe que a QoE-CAC possui valores mais baixos do que a QoE H.Du, pois a mesma identifica os prejuízos da latência quando está fora dos padrões esperados, além dos fatores de influência de *downlink* e *uplink* BEP. A QoE Optimal H.Du destaca a influência do custo de renderização associado às imagens visualizadas por cada usuário. Para o usuário 1, por exemplo, a QoE Optimal tem valor de 11.95. Esse valor corresponde à taxa de renderização da sequência de imagens visualizadas pelo usuário 1, ignorando os valores associados aos recursos de rede (*downlink* e *uplink* BEP). Precisamos, portanto, associar a esta experiência, as influências dos parâmetros de rede, na transmissão do vídeo. A variação dos valores dos parâmetros de rede ocorre dentro da janela de tempo. Em um cenário realista, os valores entregues nem sempre são constantes e próximos do prometido, conforme percebemos ao associar a Figura 2 e a Figura 3. Juntas, elas evidenciam que com a variação nos parâmetros de comunicação, a QoE pode responder positiva ou negativamente, conforme atinge os níveis esperados (ou não) para as taxas de comunicação. Perceba que os três primeiros usuários possuem um valor de QoE-Optimal similar, o que não ocorre com os valores da QoE-CAC. Isto pode ser justificado ao observarmos que o usuário 1 possui um menor valor de *downlink*, maior taxa de *uplink* BEP e maior latência. Já os usuários 2 e 3 compensam com taxas de *downlink* mais altas e menor latência, o que faz melhorar o valor da QoE-CAC.

V. CONCLUSÃO

As aplicações imersivas enfrentam o desafio contínuo de aprimorar a QoE para proporcionar um nível de imersão que atenda às expectativas do usuário. A definição de uma métrica de QoE que integre fatores de comunicação de rede, além dos

¹<https://github.com/fjiazhen-code/User-Object-Attention-Level>

²<https://github.com/dcomp-leris/VR-AR-CG-network-telemetry>

³<https://gstreamer.freedesktop.org/>

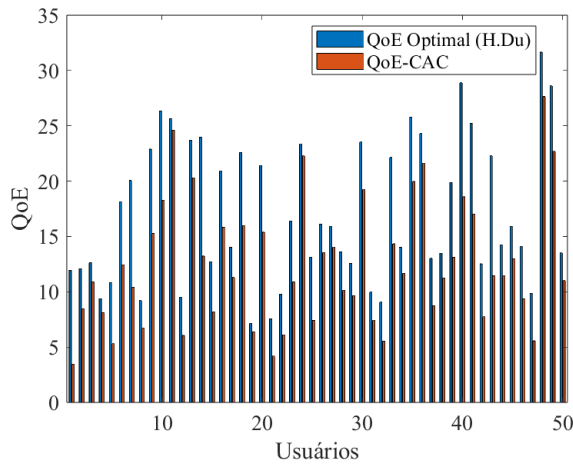


Fig. 2. Impacto dos fatores de atenção e comunicação sobre a QoE

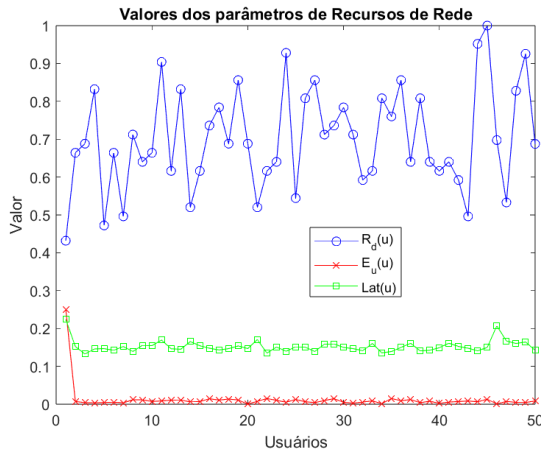


Fig. 3. Valores associados aos fatores de influência

aspectos perceptuais e atencionais, é fundamental para capturar as influências do contexto da aplicação imersiva.

Neste trabalho, propusemos um novo modelo de avaliação da QoE, incorporando parâmetros de comunicação – como latência, *downlink* e *uplink* BEP – em conjunto com o fator perceptivo da qualidade de renderização influenciada pela atenção do usuário. A métrica desenvolvida amplia a abordagem tradicional ao fornecer uma avaliação mais abrangente, permitindo não apenas medir a satisfação do usuário, mas também auxiliar os INPs e VSPs na implementação de estratégias otimizadas para maximizar a imersão e a qualidade percebida.

Os experimentos realizados demonstraram que a variação dos fatores de comunicação impacta significativamente a QoE, mesmo quando a renderização da imagem atinge um alto nível de qualidade. Como trabalho futuro, pretendemos desenvolver estratégias de otimização de latência e taxa de *downlink*, visando aprimorar a alocação de recursos e aumentar os ganhos relacionados à QoE do usuário em aplicações imersivas.

REFERÊNCIAS

- [1] J. H. Shuhaiber, “Augmented reality in surgery,” *Archives of surgery*, vol. 139, no. 2, pp. 170–174, 2004.
- [2] S. Z. A. Ansari, V. K. Shukla, K. Saxena, and B. Filomeno, “Implementing virtual reality in entertainment industry,” in *Cyber Intelligence and Information Retrieval: Proceedings of CIIR 2021*, pp. 561–570, Springer, 2022.
- [3] M. Eswaran and M. R. Bahubalendruni, “Challenges and opportunities on AR/VR technologies for manufacturing systems in the context of industry 4.0: A state of the art review,” *Journal of Manufacturing Systems*, vol. 65, pp. 260–278, 2022.
- [4] A. Clemm, M. T. Vega, H. K. Ravuri, T. Wauters, and F. De Turck, “Toward truly immersive holographic-type communication: Challenges and solutions,” *IEEE Communications Magazine*, vol. 58, no. 1, pp. 93–99, 2020.
- [5] N. Petkov, N. Christoff, A. Manolova, K. Tonchev, and V. Poulkov, “Comparative Study of Latent-Sensitive Processing of Heterogeneous Data in an Experimental Platform for 3D Video Holographic Communication,” in *2022 Global Conference on Wireless and Optical Technologies (GCWOT)*, pp. 1–6, IEEE, 2022.
- [6] J. van der Hooft, H. Amirpour, M. T. Vega, Y. Sanchez, R. Schatz, T. Schierl, and C. Timmerer, “A tutorial on immersive video delivery: From omnidirectional video to holography,” *IEEE Communications Surveys & Tutorials*, vol. 25, no. 2, pp. 1336–1375, 2023.
- [7] H. Du, J. Wang, D. Niyato, J. Kang, Z. Xiong, X. S. Shen, and D. I. Kim, “Exploring attention-aware network resource allocation for customized Metaverse services,” *IEEE Network*, 2022.
- [8] J. Li, C. Zhang, Z. Liu, W. Sun, and Q. Li, “Joint communication and computational resource allocation for QoE-driven point cloud video streaming,” in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2020.
- [9] J. Li, X. Wang, Z. Liu, and Q. Li, “A QoE Model in Point Cloud Video Streaming,” *CoRR*, vol. abs/2111.02985, 2021.
- [10] L. Wang, C. Li, W. Dai, J. Zou, and H. Xiong, “QoE-driven and tile-based adaptive streaming for point clouds,” in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1930–1934, IEEE, 2021.
- [11] ITU-T, “Quality of Experience (QoE) Influencing Factors for Augmented Reality (AR) Services,” Recommendation G.1036, 2022.
- [12] ITU-T, “Influencing factors on quality of experience for virtual reality services,” Recommendation G.1035, 2021.
- [13] ITU-T, “QoE assessment of extended reality (XR) meetings,” Recommendation P.1320, 2022.
- [14] A. Perkiš, C. Timmerer, S. Baraković, J. B. Husić, S. Bech, S. Bosse, J. Botev, K. Brunnström, L. Cruz, K. De Moor, et al., “QUALINET white paper on definitions of immersive media experience (IMEx),” *arXiv preprint arXiv:2007.07032*, 2020.
- [15] G. Zheng and L. Yuan, “A review of QoE research progress in meta-verse,” *Displays*, vol. 77, p. 102389, 2023.
- [16] J. van der Hooft, M. T. Vega, C. Timmerer, A. C. Begen, F. De Turck, and R. Schatz, “Objective and subjective QoE evaluation for adaptive point cloud streaming,” in *2020 twelfth international conference on quality of multimedia experience (QoMEX)*, pp. 1–6, IEEE, 2020.
- [17] M. S. Anwar, J. Wang, W. Khan, A. Ullah, S. Ahmad, and Z. Fei, “Subjective QoE of 360-Degree Virtual Reality Videos and Machine Learning Predictions,” *IEEE Access*, vol. 8, pp. 148084–148099, 2020.
- [18] B. Krogfoss, S. Duran, P. Pérez, and J. Bouwen, “Quantifying the Value of 5G and Edge Cloud on QoE for AR/VR,” *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*, pp. 1–4, 2020.
- [19] P. Perez, E. Gonzalez-Sosa, J. Gutiérrez, and N. García, “Emerging Immersive Communication Systems: Overview, Taxonomy, and Good Practices for QoE Assessment,” *Frontiers in Signal Processing*, vol. 2, p. 917684, 07 2022.
- [20] H. Du, J. Liu, D. Niyato, J. Kang, Z. Xiong, J. Zhang, and D. I. Kim, “Attention-Aware Resource Allocation and QoE Analysis for Metaverse xURLLC Services,” *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 7, pp. 2158–2175, 2023.
- [21] P. Reichl, S. Egger, R. Schatz, and A. D’Alconzo, “The Logarithmic Nature of QoE and the Role of the Weber-Fechner Law in QoE Assessment,” *2010 IEEE International Conference on Communications*, pp. 1–5, 2010.
- [22] A. Shirmarz, F. L. Verdi, S. K. Singh, and C. E. Rothenberg, “From Pixels to Packets: Traffic Classification of Augmented Reality and Cloud Gaming,” in *2024 IEEE 10th International Conference on Network Softwarization (NetSoft)*, pp. 195–203, 2024.