

High Quality Video Streaming System in Ultra-Low Latency over 5G-MEC

Jeongseok Kim[†] · Jaeho Lee^{††}

ABSTRACT

The Internet including mobile networks is developing to overcoming the limitation of physical distance and providing or acquiring information from remote locations. However, the systems that use video as primary information require higher bandwidth for recognizing the situation in remote places more accurately through high-quality video as well as lower latency for faster interaction between devices and users. The emergence of the 5th generation mobile network provides features such as high bandwidth and precise location recognition that were not experienced in previous-generation technologies. In addition, the Mobile Edge Computing that minimizes network latency in the mobile network requires a change in the traditional system architecture that was composed of the existing smart device and high availability server system. However, even with 5G and MEC, since there is a limit to overcome the mobile network state fluctuations only by enhancing the network infrastructure, this study proposes a high-definition video streaming system in ultra-low latency based on the SRT protocol that provides Forward Error Correction and Fast Retransmission. The proposed system shows how to deploy software components that are developed in consideration of the nature of 5G and MEC to achieve sub-1 second latency for 4K real-time video streaming. In the last of this paper, we analyze the most significant factor in the entire video transmission process to achieve the lowest possible latency.

Keywords : Ultra-Low Latency, Video Streaming, 5G, MEC, SRT

5G-MEC 기반 초저지연 고화질 영상 전송 시스템

김 정 석[†] · 이 재 호^{††}

요 약

모바일 네트워크 및 인터넷의 발전은 물리적인 거리의 한계를 극복하고 원격지의 정보를 제공하거나 획득하는데 기여하고 있다. 그러나 영상 전송을 주요 정보 제공 수단으로 사용하는 시스템은 여전히 고대역폭과 저지연 전송을 요구하고 있으며, 전송된 영상을 기반으로 상황을 판단하고 실시간 피드백을 제공하기 위해서는 전송된 영상의 품질뿐만 아니라 데이터 신뢰성과 전송 지연시간 문제는 극복해야 할 중요한 부분이다. 5세대 모바일 네트워크의 출현은 이전 세대의 기술에서 경험할 수 없었던 고대역폭과 정밀한 위치 인식 등의 특성을 제공하여, 원격 진료 및 수술, 사회 안전망을 위한 무선 원격 비디오 감시 시스템, 차량의 자율 주행 뿐만 아니라 UAV/UGV의 비가시권 제어를 실현할 수 있는 기반이 되고 있다. 또한 모바일 네트워크의 특성을 고려하여 네트워크 지연 시간을 최소화하는 Mobile Edge Computing 기술은 기존의 스마트 단말과 고가용성 서버 시스템으로 구성되던 시스템 아키텍처에 대한 변화를 요구하고 있다. 그러나 여전히 무선 구간에서 발생하는 네트워크 불확실성은 고해상도 영상을 전송할 때 영상 품질의 문제로 이어지며, 캐시를 활용한 전통적인 해결 방법은 지연 시간의 증가로 이어지게 되어 5G-MEC로 극복한 문제에 대한 근본적인 해결책이 되지 못한다. 본 연구에서는 Forward Error Correction과 Fast Retransmission을 이용하는 SRT 프로토콜을 기반으로 초저지연 고화질 영상 전송 시스템을 제안하고 각 시스템 컴포넌트를 5G-MEC의 특성을 고려하여 배치하여 4K 영상 전송시에도 종단간 지연시간을 1초 이하로 제한할 수 있음을 실험 결과로 제시하고 있다. 또한 실시간 고화질 영상 전송시 고려해야 하는 요소로, 영상의 품질과 카메라-사용자 간의 최종 지연 시간 및 지연시간에 영향을 미치는 구간을 분석하고 추가적으로 개선할 수 있는 부분을 찾아 제시하도록 한다.

키워드 : 초저지연, 영상전송, 5G, MEC, SRT

1. 서 론

영상 전송은 CDN(Content Delivery Network)처럼 다수

의 최종 사용자에게 동일한 영상을 전달하는 분야와 화상 회의와 같이 지정된 일부 사용자들간에 영상을 교환하기 위한 분야로 구분할 수 있다. 전자의 경우에는 전송 지연 시간보다는 고화질의 영상을 전송하는 것이 중요시되나, 후자의 경우는 참여자간의 원활한 대화를 위해서 영상 품질보다는 짧은 지연시간이 상대적으로 강조되는 특성을 가지고 있다. 영상 전송 지연 시간이 중요한 또 다른 영역은 드론의 비가시권 제어(BVLOS: Beyond Visual Line Of Sight)를 예로 들수 있는데, 드론에서 획득한 영상을 기반으로 기체를 제어하려고 할때에는 영상

※ 이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No. 2020-0-00096, 클라우드에 연결된 개별 로봇 및 로봇그룹의 작업 계획 기술 개발)

† 준 회원 : 서울시립대학교 전자전기컴퓨터공학부 박사과정

†† 종신회원 : 서울시립대학교 전자전기컴퓨터공학부 교수

Manuscript Received : November 5, 2020

First Revision : December 28, 2020

Accepted : January 6, 2021

* Corresponding Author : Jaeho Lee(jaeho@uos.ac.kr)

의 안정적인 전송 품질 확보와 명령과 제어(Command and Control)를 위한 저지연 전송 모두 해결해야 하는 문제로 남게 된다. 고화질 영상 전송에는 더 많은 대역폭을 요구하기 때문에 지연 시간은 상대적으로 더 늘어날 수 밖에 없다.

5G 모바일 네트워크의 출현은 기존 세대보다 높은 가용 대역폭을 제공하여 Full HD 해상도를 넘어 UHD(4K) 영상을 전송할 수 있는 토대뿐만 아니라 고대역의 무선 주파수를 활용하여 무선 구간에서 빠른 응답 속도를 기대할 수 있게 되었다. 단순히 직렬화된 짧은 데이터를 다량으로 처리하는데 있어서는 5G 기반 통신으로도 충분할지 모르나, 실시간 고화질 영상 전송의 경우에는 무선 구간에서 데이터 전송 신뢰성 및 사용자 단말까지 연결된 인터넷 구간을 고려한 시스템의 설계가 필요하게 된다. 이에 MEC(Mobile Edge Computing)의 등장은 무선 구간에 네트워크 거리가 좀 더 가까운 곳에서 데이터 처리를 수행하여 수 밀리초의 응답 속도를 구현할 수 있는 환경을 제공하고 있다.

본 연구에서는 원격지에서 전송하는 고 해상도 영상을 기반으로 상황을 판단하고 직접적인 명령을 수행할 수 있는 수준을 목표로 하고 있으며, 구체적으로는 5G 모바일 네트워크와 MEC를 활용하여 무선 구간에서 4K 해상도의 영상을 전송하면서도 카메라-사용자 중단간의 초저지연 전송을 달성하기 위한 소프트웨어 모듈과 시스템 아키텍처를 제안하고자 한다. 제 2장 관련 연구에서는 각 응용 도메인 별 요구하는 지연시간을 먼저 확인해보고, 5G와 MEC가 목표로하는 지연시간과 네트워크 품질에 대해서 살펴보고자 한다. 또한 지연시간과 관련하여 기존에 사용하고 있는 표준 기술들의 상태를 점검하고, 제안 시스템에서 도입하고 있는 새로운 프로토콜의 강점을 기술 하였다. 제 3장 제안 아키텍처에서는 5G-MEC를 고려하여 개발한 각 소프트웨어 컴포넌트와 이 컴포넌트를 이용한 서비스를 제공하기 위한 논리적 배치를 논의 하였으며, 제 4장에서는 본 제안 시스템이 달성한 결과와 그에 대한 분석을 제공하고 있다. 마지막으로 제 5장에서는 제안한 시스템의 현재 수준과 남은 과제 및 발전 방향에 대하여 기술하고 있다.

2. 관련 연구

원격 제어가 가능한 수준의 실시간 영상 전송 시스템 구현하기 위해서는 미디어의 처리, 데이터 프로토콜, 모바일 네트워크의 인프라스트럭처의 최적의 조합을 제시할 필요가 있다. 또한 막연하게 기대하는 정성적인 목표가 아닌, 기존 연구에서 확인된 사항과 실험을 통해 달성할 수 있는 이론적인 최대치를 추정하고, 이를 토대로 제안하는 시스템이 달성한 수준에 대해서 확인하고자 한다.

2.1 응용 분야별 지연시간 요구사항

전송하는 데이터의 양보다는 응답 속도의 관점에서 각 응

Table 1. Latency Requirements of Applications

Domain	Latency Requirement	Application
Healthcare	< 10 ms RTT	Haptic
Automotive	< 10 ms One-Way	Cooperative Driving
Entertainment	< 7 ms	AR/VR
Manufacturing	< 1 ms One-Way	Control and Command

용 영역을 보았을때, M. A. Lema et al.[1]은 각 영역에서 다음과 같은 지연시간 혹은 응답 속도를 만족해야한다고 정리하였다.

그러나 Table 1에서 정리한 지연 시간은 장치-사립간의 상호작용을 의미하는 것이 아니라, 장치-장치간 통신시에 확보해야하는 패킷 단위의 지연 시간을 의미한다.

영상 기반의 비가시권 제어에 대한 구체적인 언급은 없으나, 손바닥으로 물체를 만졌을 때 감각을 느끼는 속도를 근거로 하여 로봇틱스와 텔레프레즌스를 실현할 때 목표하는 지연시간을 1 ms로 제시하고 있다[2]. 그러나 이 연구에서 언급한 텔레프레즌스를 위한 응답 시간은 감각을 인지하는 처리를 포함한 것으로 앞서 제시한 패킷 단위의 지연시간과 비교한다면 매우 달성하기 어려운 목표치임을 알수 있다.

2.2 모바일 네트워크와 MEC

지연 시간 관점에서 네트워크의 성능은 5G 모바일 네트워크의 무선 특성 및 네트워크 토폴로지에 따라 달라지는데 I. Parvez at el.[3]의 연구에 따르면, 5G RAN(Radio Access Network)에서 무선 특성, 코어 네트워크의 구성과 콘텐츠 캐시의 사용 유무에 따라 실현 가능한 지연 시간이 ≤ 1 ms 부터 ≤ 17 ms의 범위임을 보이고 있다.

낮은 지연시간과 높은 신뢰성, 그리고 실시간에 가까운 핸드오버 기능은 4세대 기술인 LTE(Long Term Evolution)에서도 강조되었던 항목이다. 5G 네트워크에서 또한 유사한 항목을 강조하고 있으나, M. Lauridsen at el.[4]은 구체적인 목표 값으로 RTT(Round Trip Time)의 경우 1 ms, 연결 지연 시간(Idle to Active Time)은 10 ms, 핸드오버 시간은 0 ms를 제시하고 있다.

짧은 핸드오버 시간은 5G RAN 간의 이동시에도 데이터 전송에 영향을 미치지 않기 때문에, 네트워크 가상화(Network Virtualization)[5]를 기반으로 RAN을 에지 클라우드로 코어 네트워크를 코어 클라우드로 구성할 수 있는 기반이 된다. 또한 이렇게 구성된 클라우드는 각 단말 혹은 서비스의 성격에 따라 하나 혹은 다수의 물리적인 네트워크를 논리적인 여

Table 2. Comparison of Target Values of LTE and 5G

	LTE Target	5G Target
RTT	20 ms	1 ms
Idle to Active Time	100 ms	10 ms
Handover	49.5 ms	0 ms

러개의 독립적인 네트워크로 구분하여 서로 간의 간섭이 없도록 하는 네트워크 슬라이스가 가능하도록 한다. 이렇게 구성된 에지 클라우드와 코어 클라우드를 5G 네트워크를 제공하기 위한 용도를 벗어나서, 각 클라우드의 일부 리소스를 사용자에게 제공하고자 하는 전략으로 MEC(Mobile Edge Computing 혹은 Multi-access Edge Computing)[6]이 출현하게 되었다. MEC는 클라우드 기반으로 구성되어 필요한 리소스를 유연하게 할당 받아 사용할 수 있을 뿐 아니라, *무선 단말-5G 네트워크-인터넷* 구간이 아닌, *무선 단말-5G RAN* 혹은 *무선 단말-5G RAN-코어 네트워크*로 상대적인 전송 구간이 짧아 낮은 지연시간을 요구하는 응용 서비스에 적합한 인프라스트럭처이다.

2.3 영상 전송 기술

영상 전송은 목적에 따라 다른 기술을 사용하여 개발되고 있다. 특히 지연시간에 따른 영상 전송 프로토콜과 기술의 분포[7]는 Wowza Media Systems에 의하여 정리된 바있다. 그러나 각 기술 항목을 연관된 지연 시간만을 이용하여 1차원으로 분류하는 것은 해당 기술의 특성과 해결하고자 하는 과제를 설명하기에 부족하기 때문에 Fig. 1에서 지연 시간에 따른 분포에 영상 전송 품질을 확보해야하는 경우를 고려하여 일반적으로 적용되는 분야와 프로토콜을 정리하였다.

방송 영역(Broadcast)에서는 고화질 영상 우선 전송 및 범용성 확보를 통한 다양한 서비스 제공을 목적으로 하기 때문에 상대적으로 높은 지연 시간 분포를 가진다.

이와 반대로 화상 회의(Video Conference Call)의 경우는 반대로 양방향 대화에 어려움이 없어야하기 때문에 낮은 지연시간을 우선으로 하되 간헐적으로 발생하는 영상 손실에 대해서는 다른 영역에 비하여 관대한 편이다.

영상 기반 감시 시스템(Video Surveillance)의 경우에는 실시간 모니터링과 사건 발생에 대한 증거 확보를 위한 녹화를 동시에 수행하기 때문에 영상 품질 확보와 낮은 지연시간 달성의 두가지 항목을 어느 정도 절충하는 성격을 보인다. 마지막으로 본 연구에서 목적으로 하는 영역인 비가시권 제어(BVLOS)를 위한 영상 전송의 경우에는 상충되는 요구사항인 안정적인 고화질 영상 확보와 제어가 가능한 수준의 지연 시

간을 동시에 달성할 것을 요구하는 분야이다.

고화질 영상을 전송하기 위하여 높은 대역폭을 사용하면서도 모바일 네트워크를 통한 저지연 전송을 위해서는 영상을 고속으로 압축하고 전송할 수 있어야 하나, 모바일 네트워크 인프라스트럭처의 도움 없이 단일 기술의 발전으로는 달성하기 어려운 과제로 본 제안 시스템에서 달성하고자하는 최종 목표가 되는 영역이다.

2.4 WebRTC

데이터가 아닌 영상을 주요 정보를 획득하고 판단하는 수단으로 사용하는 실시간 영상 전달에서 지연시간은 매우 중요한 요소가 된다. 실시간 영상 전송에 주로 사용되는 기술로는 RTMP와 WebRTC가 있으나, RTMP의 경우 Adobe Flash 플랫폼 기반의 기술로 대부분의 웹 브라우저가 지원을 중단[8]하고 한시적으로 사용이 가능한 상태이므로 현재로서는 WebRTC만이 웹을 이용하여 실시간 영상 전송을 수행할 수 있는 유일한 기술이라고 할 수 있다.

WebRTC는 원격지에서 다자간의 실시간 대화를 목표로 하는 기술이기 때문에 낮은 지연 시간의 영상 전송에 적합한 기술로 평가되고 있다. WebRTC는 웹 브라우저, 모바일 응용 프로그램 간의 실시간 영상, 음성 및 데이터를 전송하기 위한 오픈 소스 프로젝트로 W3C에서는 WebRTC 1.0을 통하여 웹 기반의 P2P 연결을 가능하게 하는 HTML5 태그를 표준화하기 위하여 노력[9]하고 있으며, IETF는 실질적인 커뮤니케이션 프로토콜에 대한 표준화를 위한 활동[10]을 하고 있다.

WebRTC를 통한 영상 전송 모델은 접속 제공 방법에 따라 Mesh, Multipoint Control Unit(MCU), Selective Forwarding Unit(SFU)와 같은 세가지[11]로 나뉘어진다. Mesh 연결은 대화에 참여하는 모든 노드가 각자의 연결 상태를 유지 하기 때문에 사용자의 리소스를 많이 사용하는 구조로 알려져있다. 그러나 구현이 쉬운 편이기 때문에 상대적으로 적은 수의 사용자를 위한 서비스에 적합한 구조이다.

MCU의 경우는, 다수의 사용자가 중앙의 한 포인트로 연결을 하고, 사용자는 단일 연결을 통해 대화에 참여한 모든 사용자의 스트림을 받을 수 있는 구조이다. Mesh와 달리 사용자는 하나의 연결을 관리하기에 많은 자원을 요구하지 않으나, 서비스 제공자는 모든 트래픽을 관리해야하기 때문에 서비스 유지 관리의 복잡도가 상승하게 된다. 마지막으로 SFU의 경우, MCU와 유사하게 사용자의 모든 트래픽이 중앙의 관리 시스템으로 연결되지만, 요구하는 스트림에 대해서 선별적으로 전송하는 방식을 택하여 Mesh와 MCU 구조의 장점을 합한 구조이다. 어떠한 방식의 구성을 선택하더라도, WebRTC는 P2P기반의 프로토콜이기 때문에, Fig. 2에 묘사한 것처럼, Signal Server를 경유하여 상대방과 연결할 수 있는 경로를 확인한 후, 네트워크 상황에 따라 STUN(Session Traversal Utilities for NAT) 혹은 TURN(Traversal Using Relays around NAT) 방식으로 접속하게 된다.

오픈 소스 프로젝트로 구현된 다양한 WebRTC 플랫폼에

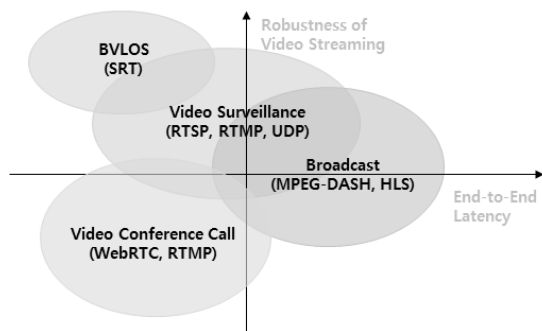


Fig. 1. Video Streaming Latency and Relative Protocol Continuum

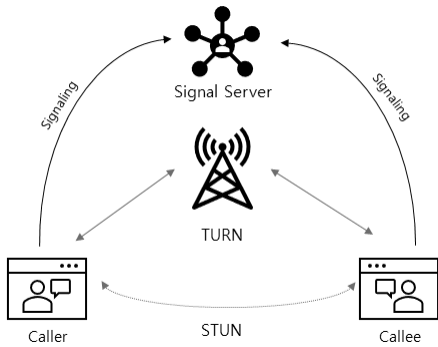


Fig. 2. WebRTC Full Monty: Singaling, TURN and STUN

서 SFU로 구성하였을 때, 540x360, 30fps의 영상을 전송하여 7명의 사용자로 구성된 화상 회의 공간을 최대 40개까지 생성하여 50ms에서 1초 이상의 지연 시간이 발생하는 것을 확인한바 있다[12]. 또한 D. Vucic at el.[13]이 WebRTC를 이용하여 수행한 실험에서도 전송 해상도가 증가함에 따라 영상 품질(Quality of Experience)은 좋지 않은 결과를 보여주고 있다. 위 실험환경이 모바일 네트워크가 아닌 WIFI를 활용한 네트워크로 한정하여 4G LTE 대비 충분한 가용 대역폭을 확보했음에도 이러한 결과를 보인 이유는 특정 조건에서 수신 버퍼량을 초과하는 데이터가 수신되어 초과분은 유실(Overrun)이 되고, 유실된 데이터 조각은 다시 영상 재생시에 버퍼가 부족한 상태(Underrun)를 초래하여 정상적인 영상 재생이 불가능한 상태가 반복되었기 때문이다.

2.5 SRT(Secure Reliable Transport)

인터넷을 통한 데이터 전송에서 대용량 데이터를 전송하는 경우 높은 BDP (Bandwidth-Delay Product)는 가용 대역폭을 충분히 활용하지 못하는 문제를 발생시킨다[14]. 예를 들어 1 Gbits/s의 대역폭을 제공하는 네트워크라도, 1ms의 RTT를 갖는 경우 BDP는 125kB¹⁾의 값을 가지나, RTT가 10ms로 증가하는 경우 BDP 값은 1250kB²⁾로 송신자가 수신자에게 보낸 데이터가 도착하기 전에 네트워크 상에 머무르는 데이터의 양이 10배가 차이나는 것을 알 수 있다. 이는 데이터 전달의 불확실성을 야기하며 전체적인 데이터 전송 처리량(Network throughput)의 감소로 나타나게 된다.

대용량 데이터 전송시에 발생하는 병목 현상을 극복하기 위해, Y. Gu와 R. L. Grossman은 UDP 기반으로 최적화된 혼잡 제어(Congestion Control) 방법을 사용한 UDT(UDP-based Data Transfer protocol)를 제안[15]하여, TCP를 이용한 데이터 전송보다 CPU 사용량은 다소 높더라도 대용량 데이터 전송에서 가용 대역폭을 효율적으로 사용할 수 있음을 증명한 바 있다. 그러나 실시간 영상 전송의 관점에서 본다면, 제한 시간 내에 도달하지 않은 영상 프레임은 재생할 필요가

1) 1Gbits/s × 1 · 10⁻³s = 1Mbits = 125kB
 2) 1Gbits/s × 10 · 10⁻³s = 10Mbits = 1250kB

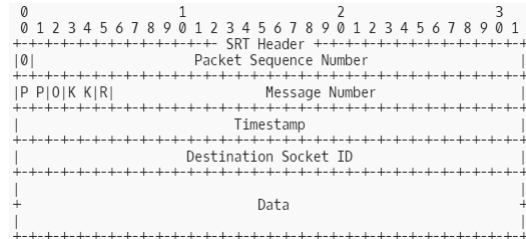


Fig. 3. Data Packet Structure of SRT Protocol

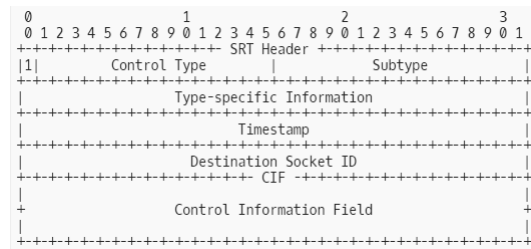


Fig. 4. Control Packet Structure of SRT Protocol

없으므로, UDT 프로토콜에 전송 시간과 도달 제한 시간 정보를 삽입하여 라이브 스트림을 지원할 수 있는 기능을 추가한 SRT(Secure Reliable Transport) 오픈 소스 프로젝트 [16]가 출현하게 되었다.

SRT 오픈 소스 프로젝트는 Haivision사에서 최초 주도한 프로젝트이나 소스코드는 Mozilla Public License로 완전히 대중에게 개방되었으며 Haivision과 SK Telecom은 공동으로 표준 스펙 제정[17]을 위해 노력하고 있다. 오픈 소스 프로젝트로 구현체 및 관련된 내용을 인용하는데 자유롭기 때문에 본 연구에서 제안하는 시스템의 기반 프로토콜로 사용하고 있다.

SRT는 UDP 기반의 데이터 전송을 위한 프로토콜로 Control과 Data 타입의 두가지 패킷 구조를 갖고 있으며, 두 타입 모두 Timestamp를 항상 명시하도록 정의되어있다. 특히 Control Packet은 Handshake, Keep-Alive, ACK, NACK, SHUTDOWN, ACKACK 등의 연결을 시도하거나 연결 상태를 유지하기 위한 정보를 교환하도록 정의하여 전송 신뢰성(Reliability)을 확보하고, UDP의 연결절차 없이 빠른 전송이 가능한 장점과 TCP처럼 연결 지향적인 특성을 모두 구현하고 있다.

SRT는 전송하는 특성에 따라 Message, Buffer, Live 세가지 모드를 지원한다. Message와 Buffer 모드는 Non-Realtime 방식으로 데이터 전송에 사용할 수 있으며, Buffer 모드의 경우, 대용량의 파일을 전송하는데 적합하다. Live 모드의 경우에는 앞서 기술한 데이터 패킷에 삽입된 시간 정보를 이용하는 것으로, 도달 유효 시간내에 도착하지 못하는 데이터 패킷을 버리거나 사용할지 여부를 결정할 수 있도록 하여 라이브 스트림에 사용할 수 있도록 고안되었다. 이는 실시간 영상 전송시에 멀티미디어 데이터를 분석하지 않고 데이터의 시간에

대한 유효성을 미리 알수 있어 불필요한 연산을 억제하는데 도움을 준다. 또한 FEC (Forward Error Correction)과 ARQ (Automatic Repeat reQuest) 방법을 이용하여 높은 BDP에서도 데이터 전송 신뢰성을 확보하여 UHD 혹은 그 이상의 고해상도 실시간 영상을 전달할 수 있는 기술로 발전하고 있다.

3. 제안 아키텍처

본 연구에서 높은 네트워크 대역폭을 요구하는 고화질 영상을 실시간으로 전송하기 위하여 SRT를 기반 프로토콜로 시스템을 구현하고, 5G 모바일 네트워크와 MEC를 활용하여 사용자 체감도(Quality of Experience)를 높일 수 있는 아키텍처를 제안한다.

3.1 시스템 컴포넌트 설계

제안 아키텍처를 구현하기 위하여 SRT 프로토콜을 기반으로 하는 실시간 영상 전송 시스템을 영상 송출(Source), SRT 중계(Relay), 영상 변환(Transcoder) 세 부분으로 정의한다.

소스 에이전트의 경우, 카메라 제어 및 영상 인코딩을 담당하고, 릴레이 에이전트는 SRT 프로토콜의 관점에서 소스 에이

전트의 데이터를 요청하는 다른 사용자 혹은 서비스에 전달하는 역할을 수행한다. 트랜스코더 에이전트는 사용자 관점에서 편리하게 확인할 수 있도록 소스 에이전트에서 송출한 영상 스트림을 웹 브라우저에서 간단히 확인할 수 있는 MPEG 스트림으로 변환하는 역할을 수행한다. 다만 이 경우에는 범용적인 영상 확인을 위한 수단으로 사용된 것이기 때문에 지연 시간 혹은 영상 품질에 대한 보장을 하지 못하게 된다.

3.2 시스템 컴포넌트의 배치

앞서 제시한 각 에이전트는 실제 환경에서 다음과 같은 응용할 수 있다.

- 1) 소스 에이전트: 드론/로봇 혹은 웨어러블 기기의 영상 송출
- 2) 릴레이 에이전트: 영상 스트림의 재분배
- 3) 트랜스코더 에이전트: 영상의 변환 혹은 영상 정보 프로세싱

제안 시스템의 각 에이전트는 Fig. 6에 도식화한 것 처럼 MEC를 기준으로 배치되어있으며, 사용자 또한 5G 네트워크를 사용하는 경우와 일반 인터넷을 이용하여 접근하는 경우로 나뉘어지게 된다.

MEC는 5G 무선 모바일 단말의 데이터를 가장 먼저 취득하게 되는 구간으로 네트워크 홉이 가장 적기 때문에 가장 낮은 지연 시간을 기대할 수 있으나, 서비스를 제공할 수 있는 대상이 (A)와 (B)와 같이 5G 무선 네트워크로 접속하는 경우로 제한되는 한계가 있다. 그렇기 때문에 모바일 네트워크에서 수신한 실시간 영상 스트림을 분배(SRT Protocol Relay Pool)하는 (C)를 MEC 내에 배치하고, 이를 인터넷 구간으로 송출하거나 혹은 빠른 응답 속도를 제공해야하는 특수 목적의 서비스(D)로 전달하는 구성을 필요로 한다. MEC 내에서 데이터를 처리할 수 있는 경우로 한정한다면, (A)에서 송출한 데이터를 (B)가 인지해야하는 경우 인터넷 구간을 거칠 필요가 없기 때문에 상대적으로 빠른 응답 시간을 확보하는데 유리하게 된다.

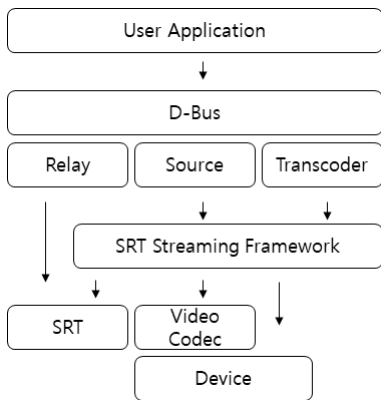


Fig. 5. Common Software Stack of Proposed System

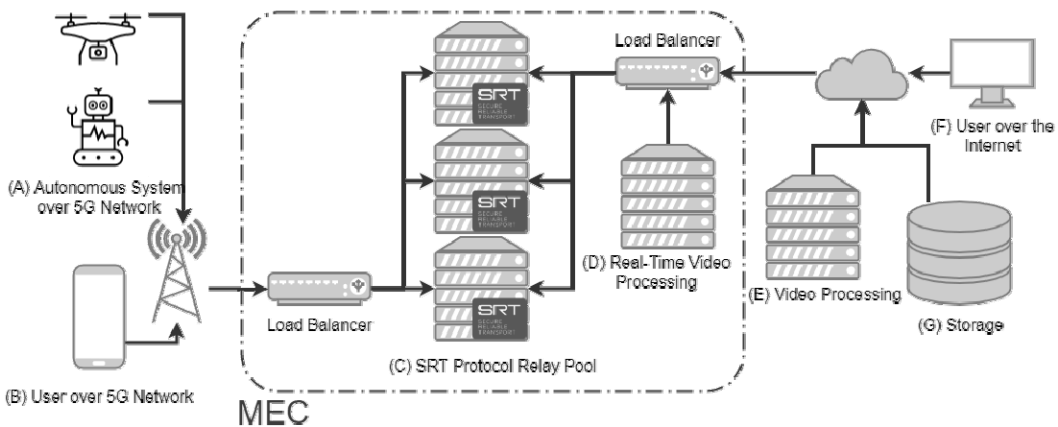


Fig. 6. Overview of Ultra-low Latency Video Streaming Agent Deployment in MEC-enabled 5G Mobile Network

반면에 일반적으로 상대적으로 느린 응답 시간을 허용하는 서비스의 경우에는 MEC와 직접 연결이 가능한 네트워크에 배치하여 자원의 효율적인 사용을 가능하도록 하며, 저장 장치(G) 혹은 비디오 인덱싱 혹은 썸네일 생성기등 과 같은 지연 시간 제약에 대해 비교적 자유로운 서비스(E)의 경우 인터넷 구간을 경유하도록 하여 유연한 서비스 구성을 가능하도록 한다.

각 서비스 컴포넌트들이 MEC, MEC와 연결 가능한 네트워크, 인터넷 공간에 분산 배치되기 때문에, 사용자의 경우에는 유즈 케이스에 따라 연결하는 네트워크를 분리하여야한다. 예를 들어 5G 모바일 네트워크를 이용한 접속인 경우 MEC에서 제공하는 서비스를 직접 접근할 수 있으나, 인터넷 접속을 하는 경우에는 MEC에서 제공하는 서비스에는 다른 서비스 컴포넌트를 경유하여 접근하기 때문에 상대적으로 높은 지연 시간을 경험하게 된다.

5G-MEC를 이용한 비가시권 제어와 같이 미션 크리티컬한 시스템을 구성하는 경우, 5G 모바일 네트워크를 사용하는 단말 혹은 MEC에 직접 접근할 수 있는 사용자에게만 장치에 대한 제어권을 주고, 인터넷 구간을 통해 접속하는 사용자에게는 시청 권한만을 부여하여 지연 시간차이로 인한 오동작을 방지하는 등의 노력이 필요한 것을 의미한다.

3.3 네트워크 상태에 따른 버퍼 설정

SRT 프로토콜의 장점중에 하나는 일정한 지연 시간을 보장한다는 것이다. 그러나 무선 구간의 불확실성은 패킷 유실에 대해서 FEC로 복원이 어려운 경우 ARQ로 재전송 요구가 가능할지라도 설정한 지연 시간 이내에 수행이 불가능하여 최종적으로는 해당 데이터를 유실할 수 있음을 의미한다.

$$FCW = \frac{B_{link} \cdot RTT}{P} \tag{1}$$

$$BS = (L_{user} + \frac{RTT}{2}) \cdot B_{estimated} \tag{2}$$

이는 5G 네트워크 구간에서는 동일 시간내에 이전 세대의 네트워크에 비하여 현저히 많은 데이터를 유실할 수 있다는 것을 의미하며, 고해상도 영상 전송에 적용하는 경우 수신한 영상 해석에 문제가 발생하게 된다는 것을 의미한다.

이를 극복하기 위하여 본 시스템에서는 앞서 살펴본 BDP와 유사한 개념으로 Equation (1)에서는 Flow Control Window을 5G 네트워크가 제공가능한 대역폭(B_{link}), 측정가능한 RTT 값, 단위 패킷의 크기(P)를 이용하여 확인하도록 하였다.

또한 Equation (2)에서는 사용자가 설정 가능한 지연 시간(L_{user})과 데이터를 전송한 후 계산 가능한 대역폭($B_{estimated}$)을 이용하여 SRT를 이용한 전송에 사용할 버퍼 크기(BS)를 결정하도록 하였다. 시스템에서 BS 값을 적용하기 위해서는 NACK 혹은 ACK를 수신하지 않고 전송되고 있는 상태의 코

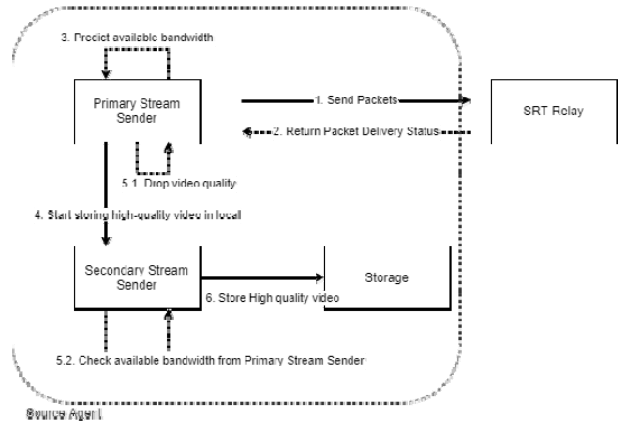


Fig. 7. An Example of High-quality Video Preserving Sequence by Communicating between Source Agent and SRT Relay Agent

기 단위를 의미하는 FCW 보다 작아서는 안된다.

Fig. 7은 이렇게 계산된 값을 이용하여, 대역폭이 순간적으로 낮아지는 경우, 영상 압축 설정값을 조절하여 실시간 영상의 송출 데이터 양을 제한하고, 고해상도 영상은 장치내에 보존하는 방식을 취하도록 하는 사용 예를 보여주고 있다.

4. 성능 검증

이번 섹션에서는 제안된 시스템을 설계 및 구현[18]하여 고해상도 영상을 전송하였을때 카메라-사용자 간의 영상 지연시간과 전송된 영상의 품질을 기존 연구와 비교 분석한 결과를 기술한다. 다만 모바일 네트워크의 특성상, 네트워크 상태 조건이 공간 및 사용자의 사용 빈도에 따라 변화가 심하기 때문에 RTT 변화 및 패킷 유실율을 시뮬레이션한 환경에서 영상 지연 시간을 확인하고, 실제 환경에서 실측한 지연 시간과 시뮬레이션 한 결과가 유사한지 검증하였다.

4.1 테스트 환경

영상 송출 장치는 4K 해상도의 영상을 전송할 수 있도록 다음과 같은 환경을 구성하였다.

- CPU: Intel 8th Gen i5-8365UE @1.6GHz(8core)
- Memory: 16GB
- Camera: Logitech 4K brio webcam

SRT 릴레이는 AWS(Amazon Web Service)에서 MEC 영역으로 제공하는 Wavelength zone[19]에 배포하고, 영상 송출 여부를 확인하기 위하여 Wavelength zone에 연결된 별도의 서브넷을 구축하여 트랜스코더 에이전트를 동작하도록 구성하였다. 다만, 카메라-사용자 지연시간을 측정하는 용도로는 MPEG-TS over SRT를 바로 재생할 수 있는 소프트웨어를 SRT 릴레이에 직접 연결하여 측정하도록 하였다. AWS내에 구성한 노드는 CPU 사용율과 가용한 네트워크 대

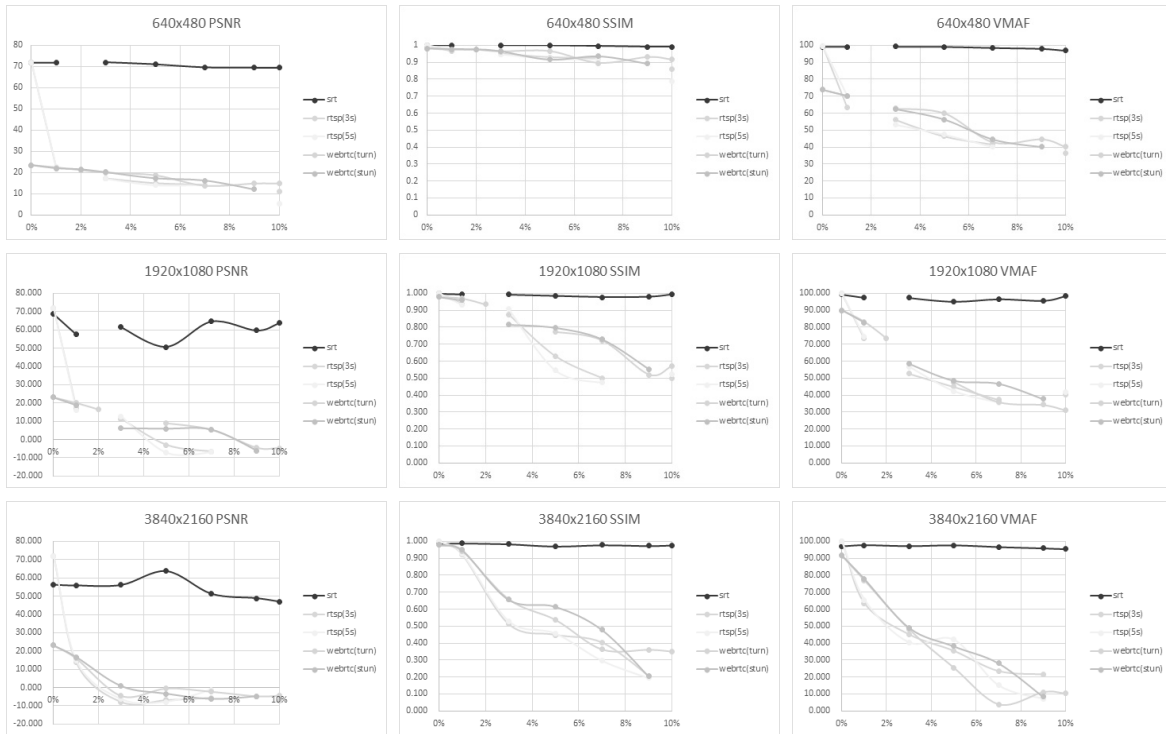


Fig. 8. Video Quality Index(PSNR, SSIM, VMAF) Comparison by Resolution According to Packet Loss Rate

역폭을 고려하여 모두 r5d.2xlarge(8 vCPU, 64GB Memory, 10 Gbps Network bandwidth)를 사용하도록 하였다.

높은 BDP 조건에서 제한된 시스템의 성능을 확인하고 기존 연구와 비교하기 위해서는 데이터의 양 혹은 RTT 값이 증가함에 따른 양상을 비교 분석할 필요가 있다. 영상 전송 시스템에서는 해상도의 증가가 전송 데이터의 증가를 의미하므로 해상도 및 RTT 값의 변화에 따른 성능을 테스트하였다.

4.2 패킷 유실율에 따른 영상 품질 변화

제한된 시스템의 데이터 전송 신뢰성(Reliability)을 확인하기 위해 패킷 유실율을 0부터 10%까지 증가하며 원본 영상과 전송된 영상의 PSNR(Peak Signal-to-Noise Ratio), SSIM(Structural SIMilarity), VMAF(Video Multimethod Assessment Fusion) 값을 비교 분석하였다.

PSNR의 경우 원본과 대조 영상의 픽셀 값의 차이에 대한 연산으로 실제 사람이 느끼는 영상의 품질과는 다소 거리감이 있고, 무손실 영상에 대해서는 PSNR 값이 무한대로 수렴하기도하나, 전송된 영상이 원본과 다른 비율을 표시하는 참고 지표로 활용할 수 있다.[20]

제한된 시스템이 전송한 영상 품질의 대조군으로는 RTSP의 jitter buffer를 3초와 5초로 설정한 전송, WebRTC의 TURN과 STUN 연결을 설정하였으며 비교 결과는 Fig. 7의 그래프와 같이 나타났다.

낮은 BDP 값을 갖는 경우인 640x480 해상도의 영상 전송에서는 대체적으로 패킷 유실율이 증가하더라도 SSIM 값

이 0.8 이상으로 나타나, 모든 시스템에서 영상 식별이 가능한 수준의 지표를 보여주고 있다. SSIM[21]은 원본과 대조 영상의 밝기(luminance), 색상 대비(contrast), 그리고 구조적 유사성을 표현하는 지표로 -1 부터 1의 값을 가지며 1에 가까울 수록 원본과 유사함을 표시하는 지표이다.

VMAF는 비교적 최근에 Netflix에서 오픈 소스[22]로 제안된 머신러닝 기반의 영상 품질 지표 계산 방식으로 내부적으로 VIF(Video Information Fidelity)[23] 혹은 DLM(Detail Loss Metric)[24]를 이용하여 영상의 각 프레임의 품질 지표를 계산한다. VMAF 점수는 0에서 100의 값을 가지며 원본과 유사할수록 100에 근접한 수치를 나타낸다.

전송 영상의 해상도가 1920x1080인 경우 패킷 유실율이 2%이상으로 증가하게 되면, 제한된 시스템(SRT)를 제외한 나머지 방법으로는 PSNR의 경우 음수의 값을 기록하고, SSIM은 0.8이하, VMAF 지표도 80 이하로 영상 식별의 어려움이 있음을 알 수 있다.

4K 해상도인 3840x2160을 전송하는 경우, 패킷 유실율에 따른 영상 식별 지표는 급격하게 감소하는 것으로 보이나 여전히 제한된 시스템의 영상 품질 지표는 식별 가능한 수준으로 나타나고 있다.

실제로 Fig. 8은 5% 패킷 유실율 조건하에서 전송한 영상 일부를 발췌한 것으로 제한된 시스템에서 전송한 영상은 원본과 거의 일치하는 품질로 전송했으나, 다른 전송 방식을 사용하는 경우에는 영상 식별이 불가능하였다.

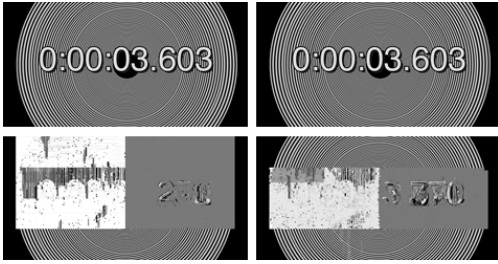


Fig. 9. 4K Video Streaming Quality Comparison in 5% Packet Loss Condition; a) Original (Upper Left), b) Proposed System (Upper Right), c) RTSP - 3s Jitter Buffer (Lower Left), d) WebRTC - STUN (Lower Right)

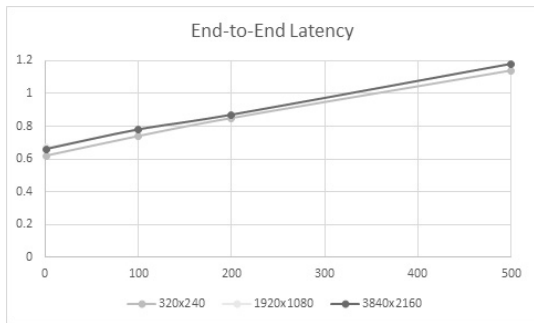


Fig. 10. Estimated End-to-End Latency by RTT(ms)

4.3 해상도 및 RTT 변화에 따른 종단간 영상 지연 시간 변화

앞의 실험을 통하여 패킷 유실율이 발생하더라도 제안된 시스템을 이용한 영상 전송은 QoE를 보장함을 알 수 있다. Fig. 9는 BDP 증가에 따른 카메라-사용자간에 실측된 영상 지연 시간의 변화를 나타낸다. 네트워크의 전송 지연 시간을 0부터 500 ms까지 증가하면서 카메라로부터 송출되는 영상과 재생되는 영상의 시간차이를 초 단위로 표기한 그래프이다.

영상 지연시간을 측정하기 위한 방법으로는 O. Boyaci, et al.[25]이 사용한 방법을 따라 영상 송출 장치와 수신 장치의 시간을 동기화한 후 영상 송출 장치의 시간 정보를 카메라를 통해 라이브 스트림으로 송출 후, 재생하는 장치의 시간과 차이를 분석하도록 하였다.

결과적으로 Fig. 10의 그래프는 RTT가 0에서 500ms로 점진적으로 증가할 때 종단간 영상 전송 지연 시간도 500ms에서 1초 정도까지 증가하는 것을 확인할 수 있으며, 이는 제안한 시스템이 의도한 대로 해상도가 변화하더라도 큰 차이가 발생하지 않았다. 이는 본 제안 시스템에서 종단간 영상 전송 지연 시간을 네트워크 상태에 따라 예측이 가능할 뿐만 아니라 더 나아가서는 복잡한 네트워크 구성을 확인하지 않고도 단순히 ICMP를 이용한 RTT 값을 구하는 것으로 영상 전송 지연 시간을 추정할 수 있다는 것을 의미한다. 또한 Table 3은 여러가지 조건이 제약되어있는 시뮬레이션 환경이 아닌, 실제 5G 및 MEC 네트워크에서 다른 시간에 측정된 지연 시간으로 위 실험으로 얻은 수치와 매우 유사한 결과를 보여주고 있다.

Table 3. End-to-End Latency Measurement Data in ms by Video Resolution in 5G Network using MEC

Latency	640×480	1920×1080	3840×2160
Min.	0.61	0.63	0.66
Max.	0.8	0.88	0.763

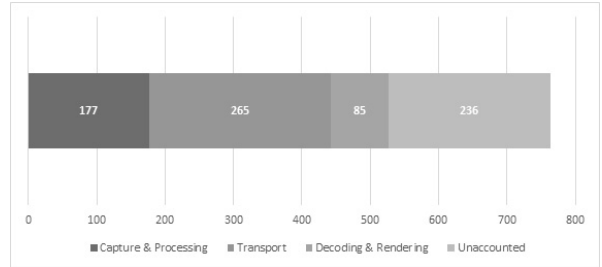


Fig. 11. End-to-End Latency Analysis

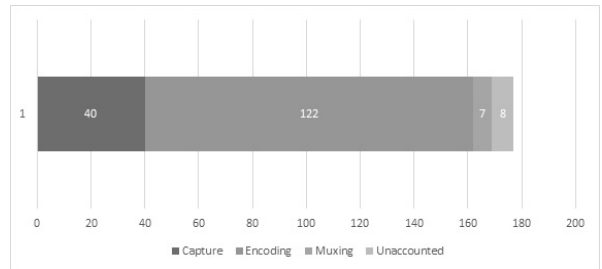


Fig. 12. Capture & Processing Time Analysis

4.4 종단간 영상 전송 지연 시간 분석

종단간 영상 전송시에는 카메라부터 영상 확인을 위한 모니터까지 본 제안 시스템에서 제어하지 못하는 많은 요소를 포함하고 있다. 그러나 테스트에서 사용한 장치들은 객관적인 데이터를 확보하기 위하여 일반적으로 사용할 수 있는 범용 장비로 수행을 하였기 본 제안 시스템에서 구현한 소프트웨어 컴포넌트외의 요소에서 발생하는 지연 시간을 분석하여 추후 최적화 할 항목들을 살펴볼 필요가 있다.

Fig. 11은 종단간 영상 전송 지연시간이 763ms로 측정된 샘플 데이터의 각 구간별 처리 시간을 분석한 결과로, 카메라의 영상을 송출하기 위해 처리하는 구간(Capture & Processing)은 177ms, SRT 전송(Transport)은 265ms, 영상을 수신한 후 출력하기 직전(Decoding & Rendering)까지는 85ms로 기록됨을 확인할 수 있었다. 그러나 나머지 236ms에 대해서는 처리 시간을 분석할 수 없었는데, 이는 USB 카메라가 처리하는 시간과 영상 디코딩 후 모니터에 출력하는데 필요한 처리 시간으로 추정된다.

보다 상세한 분석을 위하여, USB 카메라를 장착한 영상 송출 장치에서 Capture & Processing 구간을 분석한 결과는 Fig. 12에 표시한 것처럼, 영상 데이터의 획득(Capture)은 40ms, 영상 압축(Encoding)은 122ms, 미디어 멀티플렉싱(Muxing)의 경우는 7ms의 소요 시간을 기록하고 있다. 또한

설명되지 않는 8ms의 처리 시간이 기록되고 있는데 이는 앞서 분석한것처럼 USB 카메라 내부의 처리시간으로 추정된다.

실험에서 측정된 SRT 프로토콜을 통한 전송 지연시간은 전체 종단간 지연시간의 35%에 불과하기 때문에 영상 처리에 대한 최적화가 진행된다면 종단간 지연 시간은 개선될 여지가 충분히 있는 것으로 판단된다.

5. 결 론

특정한 공간의 정보를 취득하기 위해서 다양한 센서와 전송 및 처리 방법이 발전하고 있지만, 영상 전송은 다른 어떠한 수단보다도 풍부한 데이터를 제공하는 방법 중에 하나이다. 실시간 영상 전송을 통한 상황 인지를 위해서는 앞서 기술한 바와 같이, 짧은 지연 시간 내에 해당 영상을 전송하는 것 뿐만 아니라, 영상의 품질을 높이는 것도 매우 중요한 부분이다. 5G 모바일 네트워크의 출현은 이전세대와 비교하여 현저하게 짧은 지연 시간과 높은 가용 대역폭을 제공하여 고화질 영상의 실시간 전송을 가능하도록 하였으나, 높은 BDP는 5G 모바일 네트워크에서도 병목 현상을 일으켜 가용 대역폭 사용 효율을 떨어뜨리게 된다.

본 연구에서는 가용 대역폭을 효율적으로 사용하기 위하여 새롭게 고안된 SRT 프로토콜을 기반으로 시스템을 설계하여 대용량 데이터의 고속 전송에 적합함을 알수 있었다. 또한 고화질 영상의 실시간 전송은 단일 패킷에 대한 처리가 아닌 대용량 데이터의 실시간 처리에 가깝기 때문에, 모바일 네트워크에서 서비스를 요구하는 단말 혹은 사용자까지 지연시간 없이 전송할 수 있는 방법으로 소프트웨어적인 구현뿐만 아니라 5G 모바일 네트워크의 인프라스트럭처를 이용한 MEC를 적극 활용하여 RTT 값에 따른 종단간 지연 간에 대해서 실험으로 4K 고해상도 실시간 영상 전송의 지연시간을 1초 미만이 가능함을 증명하였다. 다만 실험에 사용한 환경이 영상 처리에 최적화되지 않은 상태에서 5G-MEC 실험을 진행하였기에 측정된 종단간 지연시간은 최적의 값이 아니며 추후 영상 처리를 최적화하여 단축할 수 있는 여지가 남아있음을 알 수 있다. 또한 실사용 환경에서 측정한 5G 네트워크의 RTT는 회선 사용량에 따라 변동이 있었기 때문에 향후 5G 네트워크의 목표 RTT 값인 1ms 에 근접한다면 본 연구에서 제안한 시스템으로 달성할 수 있는 종단간 지연시간은 더 낮아질 수 있을 것으로 기대한다.

References

- [1] M. A. Lema et al., "Business Case and Technology Analysis for 5G Low Latency Applications," in *IEEE Access*, Vol.5, pp.5917-5935, 2017, doi: 10.1109/ACCESS.2017.2685687.
- [2] M. A. Lema et al., "5G case study of Internet of Skills: Slicing the human senses," *2017 European Conference on Networks and Communications (EuCNC)*, Oulu, pp.1-6, 2017, doi: 10.1109/EuCNC.2017.7980762.
- [3] I. Parvez, A. Rahmati, I. Guvenc, A. I. Sarwat, and H. Dai, "A Survey on Low Latency Towards 5G: RAN, Core Network and Caching Solutions," in *IEEE Communications Surveys & Tutorials*, Vol.20, No.4, pp.3098-3130, 2018, doi: 10.1109/COMST.2018.2841349.
- [4] M. Lauridsen, L. C. Gimenez, I. Rodriguez, T. B. Sorensen, and P. Mogensen, "From LTE to 5G for Connected Mobility," in *IEEE Communications Magazine*, Vol.55, No.3, pp.156-162, March 2017, doi: 10.1109/MCOM.2017.1600778CM.
- [5] N.M. Mosharaf Kabir Chowdhury and Raouf Boutaba, "A survey of network virtualization," *Computer Networks*, Vol.54, Iss.5, pp.862-876, 2010, <https://doi.org/10.1016/j.comnet.2009.10.017>.
- [6] S. Kekki, et al., "MEC in 5G networks," *ETSI White Paper*, No.28, pp.1-28, Jun. 2018.
- [7] Low Latency Streaming [Internet], <https://www.wowza.com/low-latency>
- [8] Adobe Flash Player EOL General Information Page [Internet], <https://www.adobe.com/products/flashplayer/end-of-life.html>
- [9] C. Jennings, H. Bostrom, and J. Bruaroey, "WebRTC 1.0: Real-time communication between browsers," 2020. W3C Working Draft. <https://www.w3.org/TR/webrtc/>
- [10] Rtcweb Status Pages [Internet], <https://tools.ietf.org/wg/rtcweb/>
- [11] M. Waters, "GStreamer WebRTC," in *GStreamer Annual Conference*, Oct. 2017.
- [12] E. Andre, N. Le Breton, A. Lemesle, L. Roux, and A. Gouaillard, "Comparative Study of WebRTC Open Source SFUs for Video Conferencing," *2018 Principles, Systems and Applications of IP Telecommunications (IPTComm)*, Chicago, IL, pp.1-8, 2018, doi: 10.1109/IPTCOMM.2018.8567642.
- [13] D. Vucic and L. Skorin-Kapov, "The impact of mobile device factors on QoE for multi-party video conferencing via WebRTC," *2015 13th International Conference on Telecommunications (ConTEL)*, Graz, pp.1-8, 2015, doi: 10.1109/ConTEL.2015.7231206.
- [14] V. Jacobson and R. Braden, "TCP Extensions for Long-Delay Paths", RFC 1072, DOI 10.17487/RFC1072, Oct. 1988. <http://www.rfc-editor.org/info/rfc1072>
- [15] Y. Gu and R. L. Grossman, "UDT: UDP-based data transfer for high-speed wide area networks," *Computer Networks*, Vol.51, No.7, pp.1777-1799, 2007.
- [16] SRT(Secure Reliable Transport) [Internet], <https://github.com/Haivision/srt>

[17] M. Sharabayko, M. Sharabayko, J. Dube, J. Kim, and J. Kim, "The SRT Protocol," Work in Progress, Internet-Draft, draft-sharabayko-mops-srt-01, 9 Sept. 2020. <https://www.ietf.org/id/draft-sharabayko-mops-srt-01.txt>

[18] Hwangsaeul project [Internet], <https://github.com/hwangsaeul>

[19] AWS Wavelength [Internet], <https://aws.amazon.com/wavelength>

[20] Q. Huynh-Thu and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," *Electronics Letters*, Vol.44, Iss.13, pp.800-801, 2008, doi: 10.1049/el:20080522.

[21] Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," in *IEEE Transactions on Image Processing*, Vol.13, No.4, pp.600-612, Apr. 2004, doi: 10.1109/TIP.2003.819861.

[22] VMAF source code [Internet], <https://github.com/Netflix/vmaf>

[23] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," in *IEEE Transactions on Image Processing*, Vol.15, No.2, pp.430-444, Feb. 2006, doi: 10.1109/TIP.2005.859378.

[24] S. Li, F. Zhang, L. Ma, and K. N. Ngan, "Image Quality Assessment by Separately Evaluating Detail Losses and Additive Impairments," in *IEEE Transactions on Multimedia*, Vol.13, No.5, pp.935-949, Oct. 2011, doi: 10.1109/TMM.2011.2152382.

[25] O. Boyaci, A. Forte, S. A. Baset, and H. Schulzrinne, "vDelay: A Tool to Measure Capture-to-Display Latency and Frame Rate," *2009 11th IEEE International Symposium on Multimedia*, San Diego, CA, pp.194-200, 2009, doi: 10.1109/ISM.2009.46.



김 정 석

<https://orcid.org/0000-0001-6431-7910>

e-mail : justinkim@uos.ac.kr,

jeongseok.kim@sk.com

2001년 서울시립대학교 환경공학부(학사)

2003년 서울시립대학교 전자전기컴퓨터공학부
(석사)

2009년 ~ 현 재 서울시립대학교 전자전기컴퓨터공학부 박사과정

2018년 ~ 현 재 에스케이텔레콤 연구원

관심분야 : 인공지능, 암호학, 미디어



이 재 호

<https://orcid.org/0000-0002-3332-3207>

e-mail : jaeho@uos.ac.kr

1985년 서울대학교 계산통계학과(학사)

1987년 서울대학교 계산통계학과(석사)

1997년 University of Michigan(박사)

1998년 ~ 현 재 서울시립대학교

전자전기컴퓨터공학부 교수

관심분야 : 인공지능, 지능 로봇