

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2020년 4월 2일 (02.04.2020)

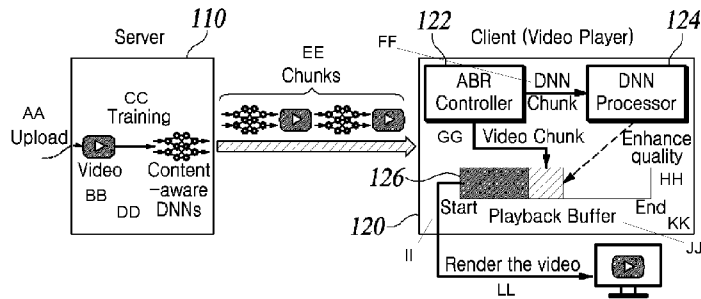


(10) 국제공개번호
WO 2020/067592 A1

- (51) 국제특허분류: H04N 21/466 (2011.01) H04N 21/2385 (2011.01)
H04N 21/2343 (2011.01) H04N 21/274 (2011.01)
- (21) 국제출원번호: PCT/KR2018/011602
- (22) 국제출원일: 2018년 9월 28일 (28.09.2018)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (71) 출원인: 한국과학기술원 (KOREA ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY) [KR/KR]; 34141 대전시 유성구 대학로 291, Daejeon (KR).
- (72) 발명자: 한동수 (HAN, Dongsu); 34141 대전시 유성구 대학로 291, Daejeon (KR). 여현호 (YEO, Hyunho); 34141 대전시 유성구 대학로 291, Daejeon (KR). 정영목 (JUNG, Youngmok); 34141 대전시 유성구 대학로 291, Daejeon (KR). 김재홍 (KIM, Jaehong); 34141 대전시 유성구 대학로 291, Daejeon (KR). 신진우 (SHIN, Jinwoo); 34141 대전시 유성구 대학로 291, Daejeon (KR).
- (74) 대리인: 이철희 (LEE, Chulhee); 06229 서울시 강남구 도곡로33길 26, 베리타스빌딩 2-4층, Seoul (KR).
- (81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(54) Title: METHOD AND APPARATUS FOR TRANSMITTING ADAPTIVE VIDEO IN REAL TIME BY USING CONTENT-AWARE NEURAL NETWORK

(54) 발명의 명칭: 콘텐츠 인지 신경망을 이용하여 실시간으로 적응형 비디오를 전송하는 방법 및 장치



- 110 ... Server
- 120 ... Client (Video Player)
- 122 ... ABR Controller
- 124 ... DNN Processor
- AA ... Upload
- BB ... Video
- CC ... Training
- DD ... Content-aware DNNs
- EE ... Chunks
- FF ... DNN Chunk
- GG ... Video Chunk
- HH ... Enhance quality
- II ... Start
- JJ ... Playback Buffer
- KK ... End
- LL ... Render the video

(57) Abstract: Provided is a method and apparatus for transmitting adaptive video in real time by using a content-aware neural network. According to an aspect of the present embodiment, a method for a server supporting a client transmitting an adaptive video in real time by using a content-aware neural network comprises the steps of: downloading a video; encoding the downloaded video according to at least one resolution; dividing the encoded video by a constant size; learning about the encoded video by using a content-aware deep neural network (DNN); generating a settings file including information regarding the learned content-aware DNN and information regarding the encoded video; and transmitting the settings file according to a request from the client.

WO 2020/067592 A1

(84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

공개:

— 국제조사보고서와 함께 (조약 제21조(3))

(57) 요약서: 콘텐츠 인지 신경망을 이용하여 실시간으로 적응형 비디오를 전송하는 방법 및 장치를 개시한다. 본 실시예의 일 측면에 의하면, 콘텐츠 인지 신경망을 이용하여 실시간으로 적응형 비디오를 전송하는 클라이언트를 지원하는 서버의 방법은, 비디오를 다운로드 받는 과정, 적어도 하나의 해상도에 따라 상기 다운로드 받은 비디오를 인코딩하는 과정, 상기 인코딩된 비디오를 일정 크기로 분할하는 과정, 상기 인코딩된 비디오를 콘텐츠 인식 DNN(Deep Neural Network)을 이용해 학습하는 과정, 상기 학습된 콘텐츠 인식 DNN에 관한 정보, 상기 인코딩된 비디오에 관한 정보를 포함하는 설정 파일을 생성하는 과정, 및 상기 클라이언트의 요청에 따라 상기 설정 파일을 전송하는 과정을 포함한다.

명세서

발명의 명칭: 콘텐츠 인지 신경망을 이용하여 실시간으로 적응형 비디오를 전송하는 방법 및 장치

기술분야

- [1] 본 발명은 콘텐츠 인지 신경망을 이용하여 실시간으로 적응형 비디오를 전송하는 방법 및 장치에 관한 것이다.

배경기술

- [2] 이 부분에 기술된 내용은 단순히 본 발명에 대한 배경 정보를 제공할 뿐 종래기술을 구성하는 것은 아니다.
- [3] 비디오 스트리밍 서비스는 지난 수십년간 빠르게 성장했다. 이러한 비디오 스트리밍 서비스의 품질은 전송 대역폭에 따라 결정되기 때문에 네트워크의 상태가 좋지 않으면 사용자의 QoE(Quality of Experience)도 저하된다. 이를 해결하기 위한 방법으로 서버 측에서는 분산 컴퓨팅 기술이 이용되고 있으며, 클라이언트(즉, 사용자) 측에서는 ABR(Adaptive Bit-Rate) 스트리밍이 전송 대역폭이 변경되거나 시간과 공간에서 변형되는 문제를 해결하는데 이용되고 있다. 그렇지만 이러한 기술들이 전송 대역폭으로부터 완전히 자유로울 수 있는 것은 아니다.
- [4] 그 외 비디오 스트리밍 서비스의 품질을 개선시키기 위해 더 좋은 코덱 선택, 적응형 비트 전송율 알고리즘의 최적화, 더 좋은 서버와 CDN(Content Distribution Networks) 선택, 중앙 제어 평면을 통한 클라이언트와 서버간의 코디네이션 등이 이용되고 있다.

발명의 상세한 설명

기술적 과제

- [5] 본 실시예는, 고품질의 비디오 스트리밍 서비스를 사용자에게 제공하는데 주된 목적이 있다. 본 실시예에 의하면 동일 품질의 비디오 스트리밍 서비스를 종래 기술보다 좁은 전송 대역폭으로 사용자에게 제공할 수 있다.

과제 해결 수단

- [6] 본 실시예의 일 측면에 의하면, 콘텐츠 인지 신경망을 이용하여 실시간으로 적응형 비디오를 전송하는 클라이언트를 지원하는 서버의 방법은, 비디오를 다운로드 받는 과정, 적어도 하나의 해상도에 따라 상기 다운로드 받은 비디오를 인코딩하는 과정, 상기 인코딩된 비디오를 일정 크기로 분할하는 과정, 상기 인코딩된 비디오를 콘텐츠 인식 DNN(Deep Neural Network)을 이용해 학습하는 과정, 상기 학습된 콘텐츠 인식 DNN에 관한 정보, 상기 인코딩된 비디오에 관한 정보를 포함하는 설정 파일을 생성하는 과정, 및 상기 클라이언트의 요청에 따라 상기 설정 파일을 전송하는 과정을 포함한다.
- [7] 상기 시스템의 실시예들은 콘텐츠 인지 신경망을 이용하여 실시간으로 적응형

비디오를 전송하는 클라이언트를 지원하는 서버 장치는, 비디오를 다운로드 받고, 상기 클라이언트의 요청에 따라 설정 파일을 전송하는 송수신부와, 적어도 하나의 해상도에 따라 상기 다운로드 받은 비디오를 인코딩하고, 상기 인코딩된 비디오를 일정 크기로 분할하고, 상기 인코딩된 비디오를 콘텐츠 인식 DNN(Deep Neural Network)을 이용해 학습하여, 상기 학습된 콘텐츠 인식 DNN에 관한 정보, 상기 인코딩된 비디오에 관한 정보를 포함하는 상기 설정 파일로 생성하는 제어부를 포함할 수 있다.

- [8] 본 실시예의 다른 측면에 의하면, 콘텐츠 인지 신경망을 이용하여 실시간으로 적응형 비디오를 전송하는 클라이언트의 방법은, 서버 장치로부터 다운로드 받을 비디오에 대한 설정 파일을 다운로드 받는 과정, 상기 설정 파일에 저장된 정보를 이용하여 콘텐츠 인식 DNN의 수행에 필요한 추론 시간을 측정하는 과정, 상기 클라이언트의 성능과 상기 측정된 추론 시간을 이용하여 다운로드 받을 대상을 결정하는 과정, 상기 결정된 대상에 대하여 상기 서버 장치로부터 다운로드 받는 과정, 상기 다운로드 받은 대상이 비디오이면 버퍼에 저장하고, 상기 다운로드 받은 대상이 콘텐츠 인식 DNN의 일부이면 상기 콘텐츠 인식 DNN에 추가하는 과정, 상기 콘텐츠 인식 DNN을 이용해 상기 버퍼에 저장된 비디오의 품질을 개선하는 과정, 및 상기 품질이 개선된 비디오를 실시간으로 재생하는 과정을 포함한다.

- [9] 본 실시예의 다른 측면에 의하면, 콘텐츠 인지 신경망을 이용하여 실시간으로 적응형 비디오를 전송하는 클라이언트 장치는, 서버 장치로부터 다운로드 받을 비디오에 대한 설정 파일을 다운로드 받고 결정된 대상에 대하여 상기 서버로부터 다운로드 받는 송수신부와 상기 설정 파일에 저장된 정보를 이용하여 콘텐츠 인식 DNN의 수행에 필요한 추론 시간을 측정하여, 상기 클라이언트의 성능과 상기 측정된 추론 시간을 이용하여 다운로드 받을 상기 대상을 결정하고, 상기 다운로드 받은 대상이 비디오이면 버퍼에 저장하고, 상기 다운로드 받은 대상이 콘텐츠 인식 DNN의 일부이면 상기 콘텐츠 인식 DNN에 추가하고, 상기 콘텐츠 인식 DNN을 이용해 상기 버퍼에 저장된 비디오의 품질을 개선하여, 상기 품질이 개선된 비디오를 실시간으로 재생하는 제어부를 포함한다.

발명의 효과

- [10] 이상에서 설명한 바와 같이 본 실시예에 의하면, 고품질의 비디오 스트리밍 서비스를 사용자에게 제공할 수 있다. 또한, 콘텐츠 인식 DNN(Deep Neural Network)을 이용함으로써, 고품질의 비디오 스트리밍 서비스를 실시간으로 사용자에게 제공할 수 있으며, 클라이언트는 자신의 하드웨어 자원(또는 성능)을 실시간으로 고려하여 비디오 스트리밍 서비스를 최적화할 수 있다. 본 실시예에 의하면, 콘텐츠 제공에 소요되는 전송 대역폭을 감소시켜 콘텐츠 제공자 혹은 CDNs 운영자는 같은 품질의 비디오를 더 적은 비용으로 전달할 수 있다.

도면의 간단한 설명

- [11] 도 1은 본 개시의 일 실시예에 따른 시스템의 개요를 나타낸 도면,
- [12] 도 2는 적응형 비트 전송율을 지원하는 DNN을 나타낸 도면,
- [13] 도 3은 본 개시의 일 실시예에 따라 적응형 비트 전송율을 지원하는 DNN을 나타낸 도면,
- [14] 도 4는 본 개시의 일 실시예에 따른 콘텐츠 인식 DNN을 간략히 나타낸 도면,
- [15] 도 5는 본 개시의 일 실시예에 따른 확장가능한 콘텐츠 인식 DNN을 나타낸 도면,
- [16] 도 6은 본 개시의 일 실시예에 따른 서버의 순서도를 나타낸 도면,
- [17] 도 7은 본 개시의 일 실시예에 따른 클라이언트의 순서도를 나타낸 도면,
- [18] 도 8은 9가지 비디오 에피소드에 대해 본 개시의 일 실시예와 종래 기술의 평균 QoE를 비교해 도시한 도면,
- [19] 도 9는 본 개시의 일 실시예와 종래 기술의 누적된 QoE를 비교해 도시한 도면이다.

발명의 실시를 위한 형태

- [20] 이하, 본 발명의 일부 실시예들을 예시적인 도면을 통해 상세하게 설명한다. 각 도면의 구성요소들에 참조부호를 부가함에 있어서, 동일한 구성요소들에 대해서는 비록 다른 도면상에 표시되더라도 가능한 한 동일한 부호를 가지도록 하고 있음에 유의해야 한다. 또한, 본 발명을 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략한다.
- [21] 또한, 본 발명의 구성 요소를 설명하는 데 있어서, 제 1, 제 2, A, B, (a), (b) 등의 용어를 사용할 수 있다. 이러한 용어는 그 구성 요소를 다른 구성 요소와 구별하기 위한 것일 뿐, 그 용어에 의해 해당 구성 요소의 본질이나 차례 또는 순서 등이 한정되지 않는다. 명세서 전체에서, 어떤 부분이 어떤 구성요소를 '포함' '구비'한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 '...부', '모듈' 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.
- [22] 도 1은 본 개시의 일 실시예에 따른 시스템의 개요를 나타낸 도면이다.
- [23] 도 1을 참조하면, 본 개시의 일 실시예에 따른 시스템인 NAS(Neural Adaptive Streaming, 이하 'NAS')는 적어도 하나의 비디오를 저장하고 있는 서버(110)와 서버(110)로부터 비디오를 다운로드하면서 실시간으로 재생할 수 있는 클라이언트(120)로 구성될 수 있다.
- [24] NAS는 DASH(Dynamic Adaptive Streaming over HTTP)에서 표준화된 HTTP 적응형 스트리밍 뿐만 아니라 다른 스트리밍에도 구현될 수 있다.

- [25] 서버(110)는 비디오가 전송되면 다양한 비트 전송율(bitrate)로 인코딩한 후 여러 개의 청크(chunk)로 나누어 저장한다. 또한, 서버(110)는 품질 향상을 위해 클라이언트(120)에서 이용할 콘텐츠 인식 DNN(Deep Neural Network, 이하 'DNN')을 학습시킨다. 이후 서버(110)는 학습된 콘텐츠 인식 DNN에 관한 정보, 비트 전송율, 해상도, 인코딩된 비디오를 다운로드 받을 수 있는 URL 등을 설정 파일(예컨대, manifest 파일 등)에 저장할 수 있다.
- [26] 서버(110)는 상기 비디오를 전송받고, 상기 설정 파일을 전송하는 송수신부(미도시)와 상기 다운로드 받은 비디오를 인코딩하여 일정 크기로 분할하고, 콘텐츠 인식 DNN을 이용해 학습하는 제어부(미도시)로 구성될 수 있다.
- [27] 클라이언트(120)는 비디오, 콘텐츠 인식 DNN, 설정 파일 등을 서버(110)로부터 다운로드 받을 수 있다. 클라이언트(120)는 먼저 상기 설정 파일을 다운로드 받아 클라이언트(120)에서 사용할 콘텐츠 인식 DNN과 비트 전송율을 결정하고, 결정된 콘텐츠 인식 DNN과 비트 전송율의 비디오를 다운로드 받을 수 있다. 클라이언트(120)는 ABR controller(122), DNN 프로세서(124)와 재생할 비디오를 저장할 버퍼(126)를 포함할 수 있다. ABR controller(122)는 현재 클라이언트(120)의 성능을 고려하여 재생할 비디오의 비트 전송율 및 다운로드할 대상으로 콘텐츠 인식 DNN 또는 비디오를 선택할 수 있다. DNN 프로세서(124)는 경량화된 메커니즘을 이용하여 클라이언트(120)의 리소스에 맞는 최적의 가용 메커니즘을 선택할 수 있다. 클라이언트(120)는 서버(110)로부터 콘텐츠 인식 DNN을 다운로드 받으면 DNN 프로세서(124)로 전달하고, 비디오를 다운로드 받으면 버퍼(126)로 전달한다. DNN 프로세서(124)가 콘텐츠 인식 DNN을 전달받으면 콘텐츠 인식 DNN을 초기화한다. 콘텐츠 인식 DNN은 비디오 화질(또는 품질)의 향상을 위해 프레임 단위로 수행된다. 이후, DNN 프로세서(124)는 다운로드되어 버퍼(126)에 저장된 비디오를 이용해 화질이 개선된 비디오를 생성해 다운로드된 비디오와 교체하여 버퍼(126)에 저장한다. 또는 DNN 프로세서(124)는 다운로드되어 버퍼(126)에 저장된 비디오를 이용해 화질이 개선된 비디오를 생성해 바로 재생할 수도 있다. 이로써 실질적으로 재생이 되는 비디오는 화질이 개선된 비디오가 될 수 있다. 디코딩, 콘텐츠 인식 DNN, 그리고 인코딩은 지연 시간을 최소화하기 위해 파이프라인과 병렬로 처리될 수 있다.
- [28] 클라이언트(120)는 서버(110)로부터 다운로드 받을 비디오에 대한 설정 파일을 다운로드 받고 비디오 및 콘텐츠 인식 DNN을 다운로드 받는 송수신부(미도시)와 상기 설정 파일에 저장된 정보를 이용하여 콘텐츠 인식 DNN의 수행에 필요한 추론 시간을 측정하여, 상기 클라이언트의 성능과 상기 측정된 추론 시간을 이용하여 다운로드 받을 상기 대상을 결정하고, 상기 다운로드 받은 대상이 비디오이면 버퍼에 저장하고, 상기 다운로드 받은 대상이 콘텐츠 인식 DNN의 일부이면 상기 콘텐츠 인식 DNN에 추가하고, 상기 콘텐츠

- 인식 DNN을 이용해 상기 버퍼에 저장된 비디오의 품질을 개선하여, 상기 품질이 개선된 비디오를 실시간으로 재생하는 제어부(미도시)로 구성될 수 있다.
- [29] 서버(110)와 클라이언트(120)의 구성을 제어부와 송수신부로 나누어 설명했으나, 하나의 구성으로 통합되어 구현될 수 있으며 또는 하나의 구성이 여러 개의 구성으로 나누어 구현될 수도 있다.
- [30] 또한, 도 1에서는 서버와 클라이언트로 나누어 설명하였으나, 서버와 클라이언트 사이에는 프록시(proxy) 서버가 더 포함될 수 있다. 이 경우 프록시 서버는 실시간 비디오 재생을 제외한 도 1에서 설명한 클라이언트의 기능을 수행하고, 클라이언트는 상기 프록시 서버로부터 실시간으로 재생할 수 있는 비디오를 수신하여 재생할 수 있다.
- [31] 도 2는 적응형 비트 전송율을 지원하는 DNN을 나타낸 도면이다.
- [32] 일반적인 DNN은 적응형 비트 전송율을 지원하기 어렵다. 적응형 비트 전송율을 지원하기 위해서는 다중 해상도를 입력으로 할 수 있어야 하며, DNN에 의한 추론이 실시간으로 수행되어야 한다. 즉, 시스템이 DNN을 이용해 비디오의 화질을 개선하고, 클라이언트가 적응형 비트 전송율(adaptive bitrate)을 지원한다면, 서버는 다양한 비트 전송율(또는 해상도)에 따라 DNN을 학습시켜야 한다. 여기서, 비트 전송율과 해상도는 특정한 관계에 있다. 예를 들면, 고해상도의 비디오를 실시간으로 재생하기 위해서는 비트 전송율은 높아야 하나, 저해상도의 비디오를 실시간으로 재생하는 경우에는 비트 전송율이 높을 필요는 없다. 그 외 비디오의 크기(가로, 세로)도 해상도와 비트 전송율에 영향을 미칠 수 있다.
- [33] 도 2는 예를 들어 해상도가 240p, 360p, 480p, 720p인 비디오를 입력해 1080p인 비디오를 출력하는 DNN을 나타낸 것이다. 이 때 DNN은 입력 해상도와 관계없이 적용될 블록(즉, 레이어)과 입력 해상도에 따라 다르게 적용될 블록으로 구성될 수 있다. 이와 같이 DNN을 이용해 서로 다른 저해상도의 비디오(또는 이미지)를 학습시켜 고해상도의 비디오를 출력시킬 수 있는 시스템을 MDSR(Multi-scale Deep Super Resolution) 시스템이라 한다.
- [34] 도 3은 본 개시의 일 실시예에 따라 적응형 비트 전송율을 지원하는 DNN을 나타낸 도면이다.
- [35] 도 2와 같은 DNN은 입력 비디오의 해상도와 관계없이 일부 블록을 공유하기 때문에 저장 공간을 줄일 수는 있으나, 입력 비디오의 해상도가 추론시간에 크게 영향을 미칠 수 있다. 예를 들어, 해상도가 720p인 비디오의 추론시간은 해상도가 240p인 비디오의 추론시간보다 약 4.3배 더 걸릴 수 있다. 따라서 해상도가 240p인 비디오는 실시간으로 재생이 가능하나, 해상도가 720p인 비디오는 실시간으로 재생하기 불가능할 수 있다. 만약 고해상도인 비디오를 실시간으로 재생하기 위해 DNN의 구성을 축소하게 되면 해상도가 낮은 비디오의 화질이 좋지 않을 수 있다.
- [36] 본 개시에서는 이를 해결하기 위해 입력 비디오의 해상도 별로 DNN을 따로

구성할 수 있다. 또한, 각 DNN은 클라이언트의 시간적 변화에 적응할 수 있도록 구성될 수 있다. 예를 들면, 해상도가 240p인 비디오의 DNN(510)은 14개의 레이어(layers)로 구성될 수 있으며, 클라이언트의 시간적인 성능 변화에 따라 8번째 레이어(512)를 실행 후 9번째 레이어(514)와 10번째 레이어(516)를 건너뛰고 11번째 레이어(518)를 실행할 수도 있다. 즉, DNN 추론시간이 비디오를 실시간으로 재생하기에 충분하다면 모든 레이어를 실행할 수 있으나, 그렇지 않다면 일부 레이어의 실행을 생략할 수 있다. 이와 관련하여서는 도 5에서 자세히 설명한다.

- [37] DNN에 관한 정보로는 입력 비디오의 해상도, 레이어의 수, 채널의 수, DNN의 용량, 출력 비디오의 품질 등이 될 수 있으며, 서버는 상기 DNN에 관한 정보를 설정 파일에 저장할 수 있다.
- [38] 도 4는 본 개시의 일 실시예에 따른 콘텐츠 인식 DNN을 간략히 나타낸 도면이다.
- [39] 비디오의 에피소드가 다양하기 때문에 모든 비디오에 잘 동작하는 범용 DNN 모델을 개발하는 것은 현실적으로 불가능하다. 이에 본 개시에서는 비디오의 에피소드(즉, 배경, 환경 등)마다 다른 DNN을 적용하는 콘텐츠 인식 DNN을 이용한다. 다만, 배경, 환경, 등장인물 등이 유사한 에피소드의 경우 동일한 콘텐츠 인식 DNN을 이용할 수도 있다. 입력 비디오의 해상도, 출력 비디오의 품질, 뿐만 아니라 비디오의 에피소드까지 고려한 콘텐츠 인식 DNN을 학습시키는 것은 시간이나 비용측면에서 비효율적일 수 있다. 본 개시에서는 이를 해결하기 위해 가장 많이 이용되는 에피소드를 일반 모델로 하여 학습한 후, 이를 기초로 다른 에피소드를 학습하여 학습 시간이나 비용 등을 줄일 수 있다. 가장 많이 이용되는 에피소드는 하나일 수도 있지만 여러 개일 수도 있다.
- [40] 도 5는 본 개시의 일 실시예에 따른 확장 가능한 콘텐츠 인식 DNN을 나타낸 도면이다.
- [41] 본 개시의 일 실시예에 따른 콘텐츠 인식 DNN은 확장이 가능하다. 상기 콘텐츠 인식 DNN은 필수 구성 요소(510)와 선택적 구성 요소(520)로 나뉠 수 있다. 필수 구성 요소(510)는 반드시 실행되어야 하나, 선택적 구성 요소(520)는 반드시 실행되어야 하는 것은 아니다. 필수 구성 요소(510)는 사전 처리 과정(512)과 사후 처리 과정(514)으로 구성될 수 있다. 선택적 구성 요소(520)는 다중 잔여 블록(522, 524, 526, 528)으로 구성될 수 있다. 또한, 다중 잔여 블록(522, 524, 526, 528)은 각각 2개의 컨볼루션 계층으로 구성될 수 있다. 선택적 구성 요소(520)를 실행하는 경우 재생되는 비디오의 품질은 더 좋아질 수 있다.
- [42] 서버가 상기 확장 가능한 콘텐츠 인식 DNN을 지원하는 경우, 상기 서버는 필수 구성 요소(510)와 선택적 구성 요소(520)의 포함 여부에 따른 모든 경로에 대해 비디오를 학습시켜야 한다. 따라서 상기 경로는 다양할 수 있다. 학습은 임의로 경로를 지정하여 이루어 질 수 있으며 출력되는 비디오와 원본 비디오 간의 오차가 적도록 학습된다. 학습이 완료되면, 상기 서버는 상기 확장 가능한

컨텐츠 인식 DNN을 청크 단위로 분할하여 저장하고, 상기 분할된 확장 가능한 컨텐츠 인식 DNN이 저장된 위치를 설정 파일에 URL로 저장할 수 있다.

- [43] 클라이언트가 상기 확장 가능한 컨텐츠 인식 DNN을 이용하는 경우, 먼저 필수 구성 요소(510)를 다운로드 받고, 상기 클라이언트의 실시간 성능(또는 자원)을 고려하여 상기 선택적 구성 요소(520)를 실행할지 여부 및 실행할 구성(522, 524, 526, 528) 등을 비디오 스트리밍 서비스 중에 결정할 수 있다. 상기 클라이언트는 처리중인 비디오 청크의 재생 시간까지 남은 시간을 계산해 이를 기초로 사용할 수 있는 상기 확장 가능한 컨텐츠 인식 DNN의 최대 레이어 수를 계산할 수 있다. 상기 클라이언트는 룩업(look-up) 테이블을 이용해 레이어 수와 추론 시간을 저장할 수 있다. 또한, 상기 클라이언트는 상기 선택적 구성 요소(520)를 서버로부터 다운받을지 여부에 대해서도 결정할 수 있다. 예를 들면, 클라이언트가 파일을 다운로드 받으면서 비디오 스트리밍 서비스를 이용한다면 상기 클라이언트는 상기 비디오 스트리밍 서비스에 이용할 자원이 충분치 않기 때문에 확장 가능한 컨텐츠 인식 DNN의 필수 구성 요소(510)만을 실행할 수 있다. 그러나, 클라이언트가 오로지 비디오 스트리밍 서비스만 이용한다면 필수 구성 요소(510)뿐만 아니라 선택적 구성 요소(520) 또한 실행할 수 있다. 클라이언트가 확장 가능한 컨텐츠 인식 DNN을 이용한다면 전송 시작시 필수 구성 요소(510)만을 수행함으로써 빠르게 비디오 스트리밍 서비스를 사용자에게 제공할 수 있다. 또한, 상기 클라이언트의 자원을 실시간으로 반영할 수 있기 때문에 사용자는 지연없이 비디오 스트리밍 서비스를 이용할 수 있다.
- [44] 도 6은 본 개시의 일 실시예에 따른 서버의 순서도를 나타낸 도면이다.
- [45] 상기 서버는 다른 장치로부터 비디오를 다운로드 받는다(610). 상기 비디오는 다양한 클라이언트들을 위해 제공되기 위한 것이다.
- [46] 상기 서버는 상기 다운로드 받은 비디오를 다양한 해상도 또는 비트 전송율에 따라 인코딩을 한다(620). 해상도와 비트 전송율은 특정한 관계에 있다. 예를 들면, 고해상도의 비디오를 실시간으로 재생하기 위해서는 비트 전송율은 높아야 하나, 저해상도의 비디오를 실시간으로 재생하는 경우에는 비트 전송율이 높을 필요는 없다. 또한 비디오의 크기(가로, 세로)도 해상도와 비트 전송율에 영향을 미칠 수 있다.
- [47] 상기 서버는 상기 인코딩된 비디오를 청크 단위로 분할한다(630).
- [48] 상기 서버는 상기 인코딩된 비디오를 이용해 컨텐츠 인식 DNN을 학습시킨다(640). 컨텐츠 인식 DNN은 해상도 별로 따로 학습될 수 있으며, 일반 모델을 이용해 학습될 수도 있다. 일반 모델이 없다면 초기화된 모델을 이용할 수도 있다.
- [49] 상기 서버는 상기 인코딩된 비디오의 저장 위치, 상기 컨텐츠 인식 DNN에 대한 정보 등을 설정 파일로 생성한다(650). 학습된 컨텐츠 인식 DNN에 관한 정보, 비트 전송율, 해상도, 인코딩된 비디오를 다운로드 받을 수 있는 URL 등이 설정 파일에 저장될 수 있다. 이때 학습된 컨텐츠 인식 DNN에 관한 정보는 인덱스

값일 수도 있다.

[50] 도 7은 본 개시의 일 실시예에 따른 클라이언트의 순서도를 나타낸 도면이다.

[51] 먼저, 상기 클라이언트는 서버로부터 다운로드 받은 비디오에 대한 설정 파일을 다운로드 받는다(710). 상기 설정 파일에는 상기 다운로드 받은 비디오에 관한 정보 외에도 상기 비디오의 품질을 향상시키는데 이용될 콘텐츠 인식 DNN에 대한 정보도 포함되어 있다. 상기 콘텐츠 인식 DNN에 대한 정보는 콘텐츠 인식 DNN의 인덱스 정보일 수 있다. 만약, 상기 클라이언트가 저장하고 있는 콘텐츠 인식 DNN이 있다면, 상기 서버와 이에 대한 정보를 공유할 수 있다. 이후 상기 클라이언트는 설정 파일에 저장하고 있는 콘텐츠 인식 DNN이 지시되면 상기 비디오만을 다운로드 받을 수 있다.

[52] 상기 클라이언트는 상기 설정 파일에 저장된 정보를 이용하여 콘텐츠 인식 DNN의 수행에 필요한 추론 시간을 측정한다(720). 상기 추론 시간을 측정하기 위해 모든 옵션에 대한 상기 설정 파일에 저장된 콘텐츠 인식 DNN을 다운로드할 수 있다. 그러나, 이 경우 불필요하게 많은 자원과 시간을 낭비할 수 있다. 따라서, 본 개시의 일 실시예에서 상기 클라이언트는 상기 콘텐츠 인식 DNN을 모두 다운로드 하지 않고 상기 설정 파일에 저장된 옵션에 대한 정보를 이용해 임의로 콘텐츠 인식 DNN을 구성할 수 있다. 즉, 실제 사용할 콘텐츠 인식 DNN을 다운로드 받지 않고, 상기 설정 파일에 저장된 입력 비디오의 해상도, 품질 레벨, 레이어의 수, 및 채널의 수 등을 이용해 임의로 콘텐츠 인식 DNN을 구성하여 콘텐츠 인식 DNN의 수행에 필요한 추론 시간을 측정할 수 있다. 이렇게 하는 경우, 상기 클라이언트가 4개의 DNN 옵션을 테스트하는데 소요되는 시간은 일반적인 비디오 청크의 재생 시간보다 적기 때문에 상기 클라이언트는 두 번째 비디오 청크를 재생하기 전에 어떤 옵션의 콘텐츠 인식 DNN을 사용할지 결정할 수 있다. 선택적으로, 상기 클라이언트는 DNN 옵션별로 수행하는데 필요한 추론 시간을 미리 저장하고 이를 이용해 사용할 콘텐츠 인식 DNN을 선택할 수도 있다.

[53] 상기 클라이언트는 상기 클라이언트의 성능과 상기 측정된 추론 시간을 이용하여 다운로드 받을 대상을 결정한다(730). 상기 클라이언트는 통합된 ABR 알고리즘(integrated adaptive bitrate algorithm)을 이용해 이후에 재생될 비디오를 다운로드 받거나 콘텐츠 인식 DNN을 다운로드 받을 수 있다. 상기 통합된 ABR 알고리즘은 직접 최적화하는 보강학습(Reinforcement Learning) 프레임 워크를 사용하고 A3C(asynchronous advantage actor-critic)을 딥러닝 알고리즘으로 선택한다. 구체적으로, 관측으로부터 전략(또는 정책)을 배우고 다운로드된 콘텐츠 인식 DNN의 비율, 콘텐츠 인식 DNN으로 인한 품질 향상, 네트워크 처리량 샘플 및 버퍼 점유율과 같은 원시 관측을 결정에 매핑할 수 있다.

[54] 보강학습에서 에이전트는 환경과 상호 작용을 한다. 반복되는 t 마다 에이전트는 환경에서 상태(s_t)를 관찰한 후 작업(a_t)을 수행한다. 이후, 환경은 보상(r_t)을 생성하고 상태를 s_{t+1} 로 갱신한다. 정책(π)은 주어진 상태(s_t)에서 행동(a

)을 취할 확률을 제공하는 함수로 다음과 같이 정의될 수 있다.

[55] $\pi(s_t, a_t): \rightarrow [0, 1]$

[56] 목표는 미래의 할인 보상의 합($\sum_{t=0}^{\infty} \gamma^t r_t$)을 최대화하는 정책(π)을

학습한다. 여기서, $\gamma \in (0, 1]$ 은 미래 보상에 대한 할인율이다.

[57] 또한, 작업의 셋($\{a_t\}$)는 콘텐츠 인식 DNN 청크를 다운로드할지 또는 특정 비트 전송율의 비디오 청크를 다운로드할지 여부로 지정할 수 있다. 보상(r_t)은 비트 전송율 유틸리티, 재버퍼링 시간 및 선택된 비트 전송율의 평활도의 함수로 QoE 측정항목이 될 수 있다. 상태(s_t)는 다운로드 받을 남은 콘텐츠 인식 DNN 청크의 개수, 처리량 측정치 및 클라이언트 측정치 (예를 들어, 버퍼 점유율, 과거의 비트 전송율)를 포함할 수 있다. 이후, 콘텐츠 인식 DNN 다운로드 및 품질 향상을 반영하여 환경 보상 및 상태를 업데이트한다. 콘텐츠 인식 DNN 다운로드는 '남은 콘텐츠 인식 DNN 청크의 개수'를 감소시켜 상태를 업데이트할 수 있다. 다운로드된 각 비디오 청크에 대한 보상은 콘텐츠 인식 DNN 기반 품질 향상에 반영하도록 업데이트될 수 있다. 상기 품질 향상은 다운로드된 콘텐츠 인식 DNN의 일부분에 대한 함수일 수 있다. 특히 콘텐츠 인식 DNN이 제공하는 평균 품질 향상에 따른 QoE의 비트율 유틸리티 구성 요소를 향상시킬 수 있다. 표 1은 본 개시에 따른 상태(S_t)를 나타낸다.

[58] [표1]

type	상태(S_t)
DNN status	남은 콘텐츠 인식 DNN 청크의 개수
Network status	과거 N개의 청크에 대한 처리량
	과거 N개의 청크를 다운로드 받는데 걸리는 시간
Client status	재생할 비디오를 저장할 버퍼 점유율
	다음 비디오 청크의 크기(sizes)
Video status	최근 비디오 청크의 비트 전송율
	남은 비디오 청크의 개수

[59] 보강학습에서는 정책을 나타내는 배우(actor)와 정책(policy)의 성능을 평가하는 데 사용되는 비평가라는 두 가지 신경 근사를 가지고 있다. 본 개시에서는 배우와 비평 네트워크를 교육시키기 위해 policy gradient method을 이용한다. 에이전트는 먼저 현재 정책($\pi_0(s_t, a_t)$)에 따라 궤적을 생성한다. 여기서, θ 는 배우의 신경망의 매개 변수(또는 가중치)를 나타낼 수 있다. 비평 네트워크는 이러한 궤적을 관찰하고 상태 s_t 에서 시작하여 정책 π_0 를 따르는 동작 z 를 취하는 것과 관련하여 예상되는 보상인 동작 값 함수 $Q^{\pi_0}(s_t, a_t)$ 를 추정하는 방법을

학습한다.

[60]

$$\theta \leftarrow \theta + \alpha \sum_t \nabla_{\theta} \log \pi_{\theta}(s_t, a_t) (Q^{\pi_{\theta}}(s_t, a_t) - V^{\pi_{\theta}}(s_t))$$

[61]

여기서 $V^{\pi_{\theta}}(s_t)$ 는 상태 s_t 에서 시작하는 π_{θ} 의 총 예상 보상을 나타내는 값 함수이고, α 는 학습 속도이다. 본 개시에 따른 보강학습에서 보상은 콘텐츠 인식 DNN이 제공하는 평균 QoE 향상을 반영하기 때문에 비평 네트워크는 업데이트된 총 보상을 예측하는 방법을 학습한다. 이를 통해 배우(actor)는 QoE를 최대화하기 위해 비디오와 DNN 다운로드의 균형을 유지하는 정책을 학습할 수 있다.

[62]

다시 도 7을 참조하여, 상기 클라이언트는 상기 다운로드 받은 대상이 비디오 청크이면 버퍼에 저장하고, 상기 다운로드 받은 대상이 콘텐츠 인식 DNN 청크이면 상기 콘텐츠 인식 DNN에 추가한다(740).

[63]

상기 클라이언트는 다운로드 받은 비디오 청크를 콘텐츠 인식 DNN을 이용해 화질을 개선한다(750). 상기 클라이언트는 DNN을 수행하기 위한 전용 DNN 프로세서를 이용할 수 있다.

[64]

상기 클라이언트는 화질이 개선된 비디오 청크를 실시간으로 재생한다(760).

[65]

도 6과 도 7에서는 각 과정을 순차적으로 실행하는 것으로 기재하고 있으나, 이는 본 발명의 일 실시예의 기술 사상을 예시적으로 설명한 것에 불과한 것이다. 다시 말해, 본 발명의 일 실시예가 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 일 실시예의 본질적인 특성에서 벗어나지 않는 범위에서도 도 6과 도 7에 기재된 순서를 변경하여 실행하거나 과정들 중 하나 이상의 과정을 병렬적으로 실행하는 것으로 다양하게 수정 및 변형하여 적용 가능할 것이므로, 도 6과 도 7은 시계열적인 순서로 한정되는 것은 아니다.

[66]

한편, 도 6과 도 7에 도시된 과정들은 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 즉, 컴퓨터가 읽을 수 있는 기록매체는 마그네틱 저장매체(예를 들면, 롬, 플로피 디스크, 하드디스크 등), 광학적 판독 매체(예를 들면, 시디롬, 디브이디 등) 및 캐리어 웨이브(예를 들면, 인터넷을 통한 전송)와 같은 저장매체를 포함한다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다.

[67]

도 8은 9가지 비디오 에피소드에 대해 본 개시의 일 실시예와 종래 기술의 평균 QoE를 비교해 도시한 도면이다.

[68]

구체적으로, 해상도 1080p이고 길이가 5분보다 긴 9가지의 비디오 에피소드를 이용해 학습을 위해 80%의 무작위 표본을 사용해 10시간 이상 학습시켰다.

Pensieve는 심층 보강학습을 이용하여 QoE를 최대화한 기술이며, MPC는 QoE를 최대화하는 비트 전송율을 선택하기 위해 다음 5개의 청크에 대한 버퍼 점유율과 처리량 예측을 이용한 기술이다. 또한, BOLA는 버퍼 점유율을 기반으로 Lyapunov 최적화한 기술이다.

- [69] QoE에 대한 측정항목으로는 QoE_{lin} , QoE_{log} , 및 QoE_{hd} 를 이용한다. QoE_{lin} 는 선형 비트 전송율 유틸리티, QoE_{log} 는 감소하는 한계 유틸리티를 나타내는 log 비트 전송율, 그리고 QoE_{hd} 는 고화질 비디오를 선호하는 정도를 나타낸다. 도 8에서 오차 막대는 평균으로부터의 표준 편차를 나타낸다. 본 개시의 일 실시예에 따르면 NAS는 세 가지 QoE 측정 항목 모두에서 모든 비디오 에피소드에 대해 가장 높은 QoE를 나타낸다. NAS는 지속적으로 Pensieve보다 QoE_{lin} 는 43.08%, QoE_{log} 는 36.26%, QoE_{hd} 는 42.57% 우수함을 보이고 있다. QoE_{lin} 을 사용하면 NAS가 Pensieve보다 평균 43.08% 성능이 우수하지만 Pensieve는 MPC보다 19.31% 향상된 성능을 보인다. BOLA와 비교하여도 NAS는 92.28% 향상됨을 알 수 있다. 장면 복잡성, 압축 아티팩트, 시간 중복성과 같은 많은 요소가 DNN 성능에 영향을 주기 때문에 QoE 개선은 Pensieve보다 비디오 에피소드에 따라 21.89%(뷰티)에서 76.04%(음악)까지 다양하였다.
- [70] 도 9는 본 개시의 일 실시예와 종래 기술의 누적된 QoE를 비교해 도시한 도면이다.
- [71] 도 9는 도 8의 9개의 비디오 에피소드 중 중간값을 나타낸 게임 에피소드를 이용하여 103개 이상의 네트워크 트래이스에 대한 QoE의 누적 분포를 도시한 것이다. NAS는 모든 네트워크 조건에서 이점을 제공한다. 일례로, NAS는 Pensieve보다 QoE_{lin} 중간값이 58.55% 향상되었다. 참고로, Pensieve는 주로 비트 전송율 유틸리티를 사용하여 재버퍼링을 줄임으로써 MPC에 비해 QoE 이득을 제공한다. 반대로 NAS는 클라이언트 측의 계산을 사용하기 때문에 이러한 절충을 나타내지 않는다. 다른 비디오 에피소드를 도시하지는 않았지만 비슷한 경향을 나타낸다.
- [72] 이상의 설명은 본 실시예의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 실시예가 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다. 따라서, 본 실시예들은 본 실시예의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예에 의하여 본 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 실시예의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.
- [73]

청구범위

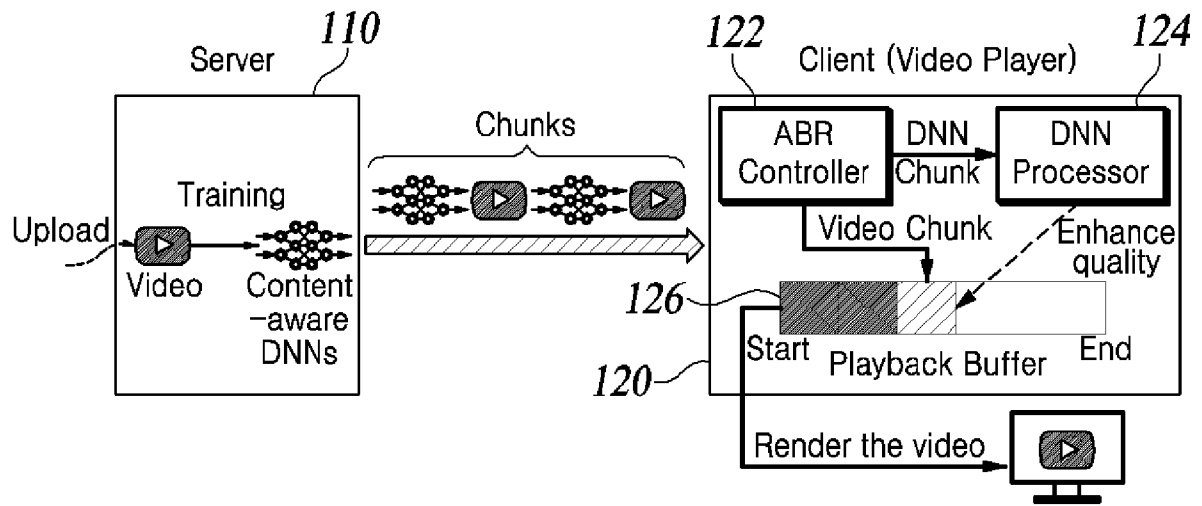
- [청구항 1] 콘텐츠 인지 신경망을 이용하여 실시간으로 적응형 비디오를 전송하는 서버의 방법에 있어서,
비디오를 다운로드 받는 과정,
적어도 하나의 해상도에 따라 상기 다운로드 받은 비디오를 인코딩하는 과정,
상기 인코딩된 비디오를 일정 크기로 분할하는 과정,
상기 인코딩된 비디오를 콘텐츠 인식 DNN(Deep Neural Network)을 이용해 학습하는 과정,
상기 학습된 콘텐츠 인식 DNN에 관한 정보, 상기 인코딩된 비디오에 관한 정보를 포함하는 설정 파일을 생성하는 과정, 및
클라이언트의 요청에 따라 상기 설정 파일을 전송하는 과정을 포함하는 방법.
- [청구항 2] 제1항에 있어서,
상기 인코딩된 비디오에 관한 정보는,
상기 인코딩된 비디오의 저장 위치, 해상도, 및 비트 전송율 중 적어도 하나를 포함하는 것을 특징으로 하는 방법.
- [청구항 3] 제1항에 있어서,
상기 학습된 콘텐츠 인식 DNN에 관한 정보는,
상기 학습된 콘텐츠 인식 DNN의 레이어 수, 채널 수, 저장 위치, 크기, 및 향상된 품질의 정도 중 적어도 하나를 포함하는 것을 특징으로 하는 방법.
- [청구항 4] 제1항에 있어서,
상기 콘텐츠 인식 DNN은 반드시 수행되는 필수 구성 요소와 선택적으로 수행될 수 있는 선택적 구성 요소로 구성됨을 특징으로 하는 방법.
- [청구항 5] 콘텐츠 인지 신경망을 이용하여 실시간으로 적응형 비디오를 전송하는 서버 장치에 있어서,
비디오를 다운로드 받고, 클라이언트의 요청에 따라 설정 파일을 전송하는 송수신부와,
적어도 하나의 해상도에 따라 상기 다운로드 받은 비디오를 인코딩하고, 상기 인코딩된 비디오를 일정 크기로 분할하고, 상기 인코딩된 비디오를 콘텐츠 인식 DNN(Deep Neural Network)을 이용해 학습하여, 상기 학습된 콘텐츠 인식 DNN에 관한 정보, 상기 인코딩된 비디오에 관한 정보를 포함하는 상기 설정 파일로 생성하는 제어부를 포함하는 서버 장치.
- [청구항 6] 제5항에 있어서,
상기 인코딩된 비디오에 관한 정보는,
상기 인코딩된 비디오의 저장 위치, 해상도, 및 비트 전송율 중 적어도 하나를 포함하는 것을 특징으로 하는 서버 장치.

- [청구항 7] 제5항에 있어서,
 상기 학습된 콘텐츠 인식 DNN에 관한 정보는,
 상기 학습된 콘텐츠 인식 DNN의 레이어 수, 채널 수, 저장 위치, 크기, 및
 향상된 품질의 정도 중 적어도 하나를 포함하는 것을 특징으로 하는 서버
 장치.
- [청구항 8] 제1항에 있어서,
 상기 콘텐츠 인식 DNN은 반드시 수행되는 필수 구성 요소와 선택적으로
 수행될 수 있는 선택적 구성 요소로 구성됨을 특징으로 하는 서버 장치.
- [청구항 9] 콘텐츠 인지 신경망을 이용하여 실시간으로 적응형 비디오를 다운로드
 받는 장치의 방법에 있어서,
 서버 장치로부터 다운로드 받을 비디오에 대한 설정 파일을 다운로드
 받는 과정,
 상기 설정 파일에 저장된 정보를 이용하여 콘텐츠 인식 DNN의 수행에
 필요한 추론 시간을 측정하는 과정,
 상기 장치의 성능과 상기 측정된 추론 시간을 이용하여 다운로드 받을
 대상을 결정하는 과정,
 상기 결정된 대상에 대하여 상기 서버 장치로부터 다운로드 받는 과정,
 상기 다운로드 받은 대상이 비디오이면 버퍼에 저장하고, 상기 다운로드
 받은 대상이 콘텐츠 인식 DNN의 일부이면 상기 콘텐츠 인식 DNN에
 추가하는 과정, 및
 상기 콘텐츠 인식 DNN을 이용해 상기 버퍼에 저장된 비디오의 품질을
 개선하는 과정을 포함하는 방법.
- [청구항 10] 제9항에 있어서,
 상기 품질이 개선된 비디오를 실시간으로 재생하거나 또는 전송하는
 과정을 포함하는 방법.
- [청구항 11] 제9항에 있어서,
 상기 추론 시간을 측정하는 과정은,
 상기 설정 파일에 저장된 콘텐츠 인식 DNN에 관한 정보를 이용해 임의로
 콘텐츠 인식 DNN을 구성해 상기 콘텐츠 인식 DNN의 수행에 필요한 추론
 시간을 측정하는 과정임을 특징으로 하는 방법.
- [청구항 12] 제9항에 있어서,
 상기 설정 파일은
 상기 다운로드 받을 비디오의 저장 위치, 해상도, 및 비트 전송율 중
 적어도 하나를 포함하는 것을 특징으로 하는 방법.
- [청구항 13] 제9항에 있어서,
 상기 설정 파일은
 상기 콘텐츠 인식 DNN의 레이어 수, 채널 수, 저장 위치, 크기, 및 향상된
 품질의 정도 중 적어도 하나를 포함하는 것을 특징으로 하는 방법.

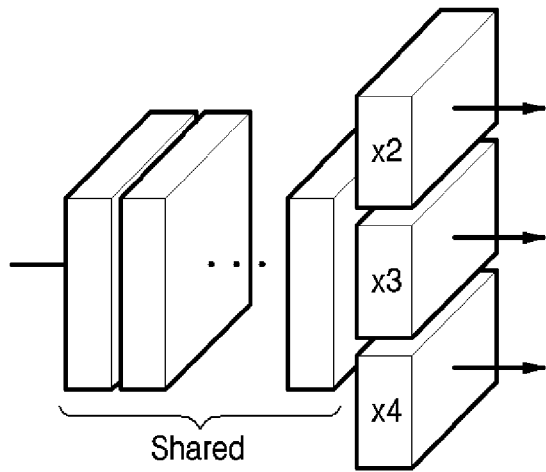
- [청구항 14] 제9항에 있어서,
상기 콘텐츠 인식 DNN은 반드시 수행되는 필수 구성 요소와 선택적으로 수행될 수 있는 선택적 구성 요소로 구성됨을 특징으로 하는 방법.
- [청구항 15] 제9항에 있어서,
상기 다운로드 받을 대상을 결정하는 과정은,
보강학습을 프레임 워크를 사용하고 A3C(asynchronous advantage actor-critic)을 딥러닝 알고리즘으로 선택하여 상기 다운로드 받을 대상을 결정하는 과정임을 특징으로 하는 방법.
- [청구항 16] 콘텐츠 인지 신경망을 이용하여 실시간으로 적응형 비디오를 다운로드 받는 장치에 있어서,
서버 장치로부터 다운로드 받을 비디오에 대한 설정 파일을 다운로드 받고 결정된 대상에 대하여 상기 서버로부터 다운로드 받는 송수신부와 상기 설정 파일에 저장된 정보를 이용하여 콘텐츠 인식 DNN의 수행에 필요한 추론 시간을 측정하여, 상기 클라이언트의 성능과 상기 측정된 추론 시간을 이용하여 다운로드 받을 상기 대상을 결정하고, 상기 다운로드 받은 대상이 비디오이면 버퍼에 저장하고, 상기 다운로드 받은 대상이 콘텐츠 인식 DNN의 일부이면 상기 콘텐츠 인식 DNN에 추가하고, 상기 콘텐츠 인식 DNN을 이용해 상기 버퍼에 저장된 비디오의 품질을 개선하는 제어부를 포함하는 장치.
- [청구항 17] 제16항에 있어서,
상기 제어부는,
상기 품질이 개선된 비디오를 실시간으로 재생함을 특징으로 하는 장치.
- [청구항 18] 제16항에 있어서,
상기 송수신부는,
상기 품질이 개선된 비디오를 실시간으로 전송함을 특징으로 하는 장치.
- [청구항 19] 제16항에 있어서,
상기 제어부는,
상기 설정 파일에 저장된 콘텐츠 인식 DNN에 관한 정보를 이용해 임의로 콘텐츠 인식 DNN을 구성해 상기 콘텐츠 인식 DNN의 수행에 필요한 추론 시간을 측정함을 특징으로 하는 장치.
- [청구항 20] 제16항에 있어서,
상기 설정 파일은
상기 다운로드 받을 비디오의 저장 위치, 해상도, 및 비트 전송율 중 적어도 하나를 포함하는 것을 특징으로 하는 장치.
- [청구항 21] 제16항에 있어서,
상기 설정 파일은
상기 콘텐츠 인식 DNN의 레이어 수, 채널 수, 저장 위치, 크기, 및 향상된 품질의 정도 중 적어도 하나를 포함하는 것을 특징으로 하는 장치.

- [청구항 22] 제16항에 있어서,
상기 콘텐츠 인식 DNN은 반드시 수행되는 필수 구성 요소와 선택적으로
수행될 수 있는 선택적 구성 요소로 구성됨을 특징으로 하는 장치.
- [청구항 23] 제16항에 있어서,
상기 제어부는,
보강학습을 프레임 워크를 사용하고 A3C(asynchronous advantage
actor-critic)을 딥러닝 알고리즘으로 선택하여 상기 다운로드 받을 대상을
결정함을 특징으로 하는 장치.

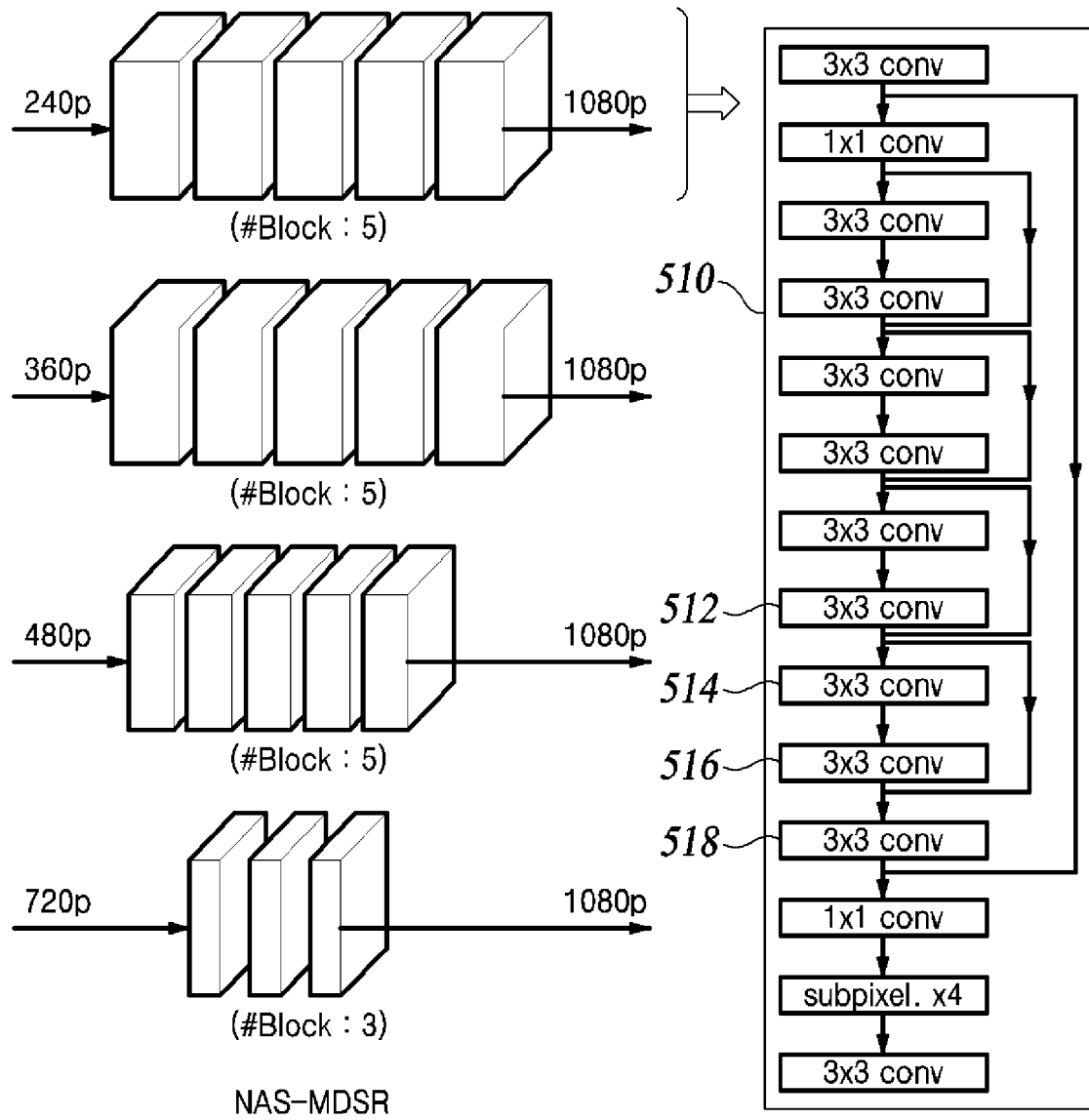
[도1]



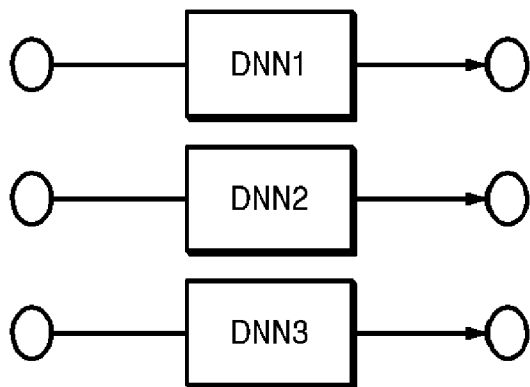
[도2]



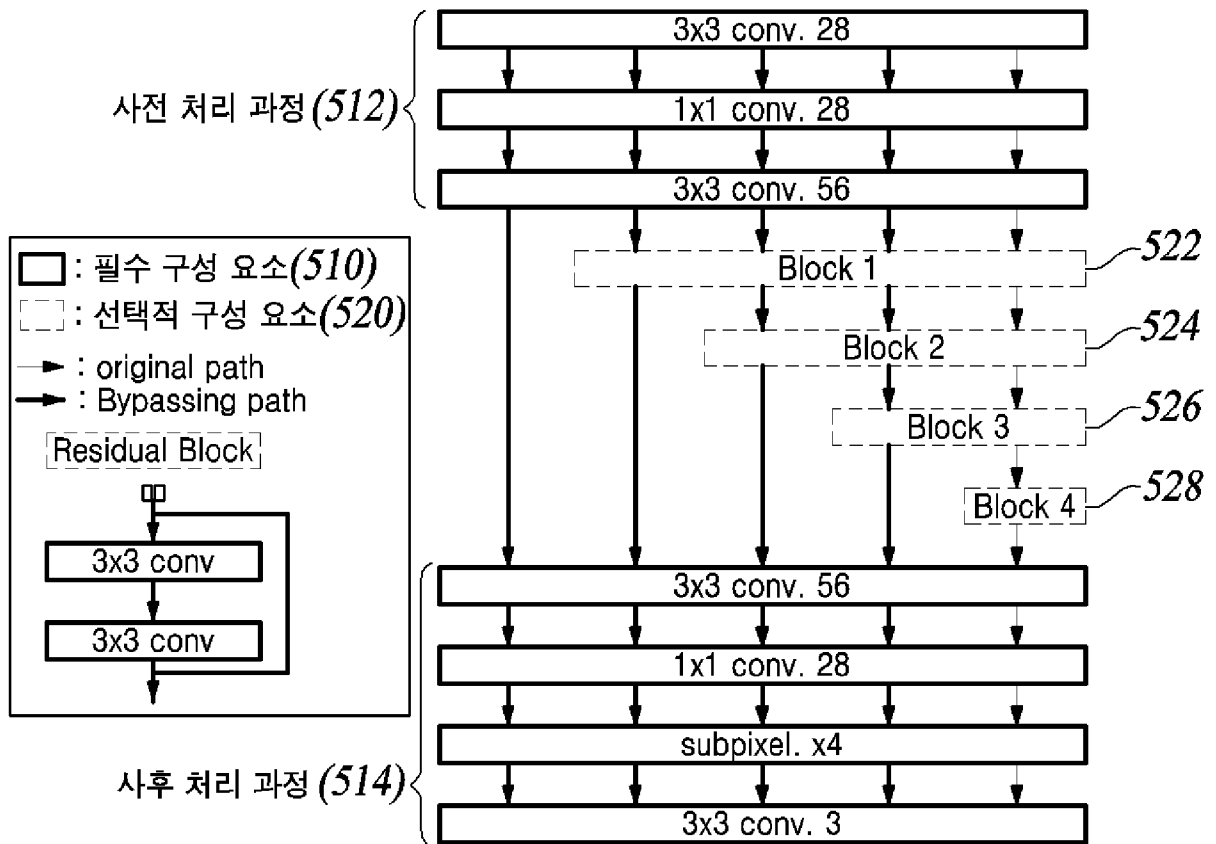
[도3]



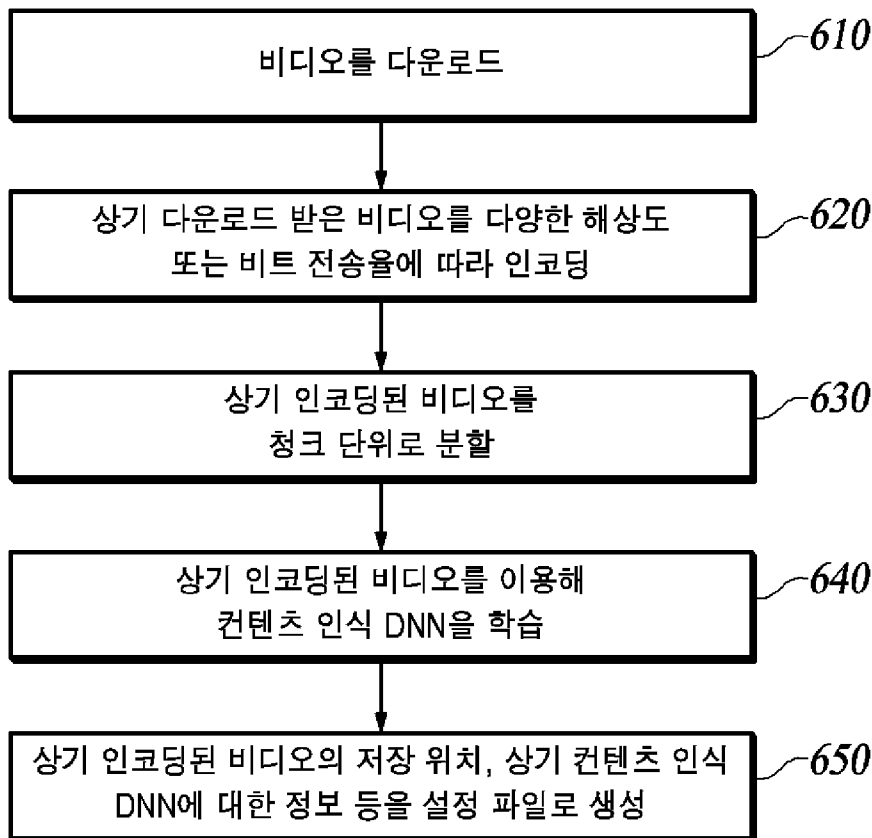
[도4]



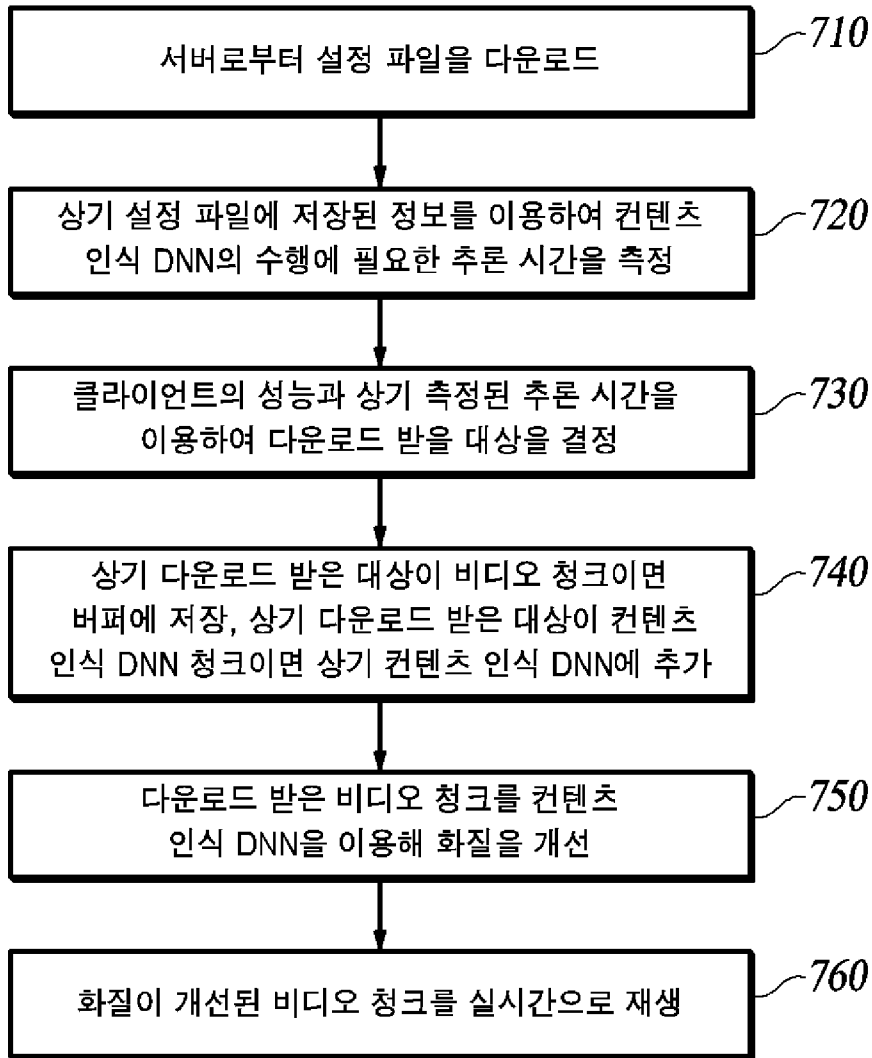
[도5]



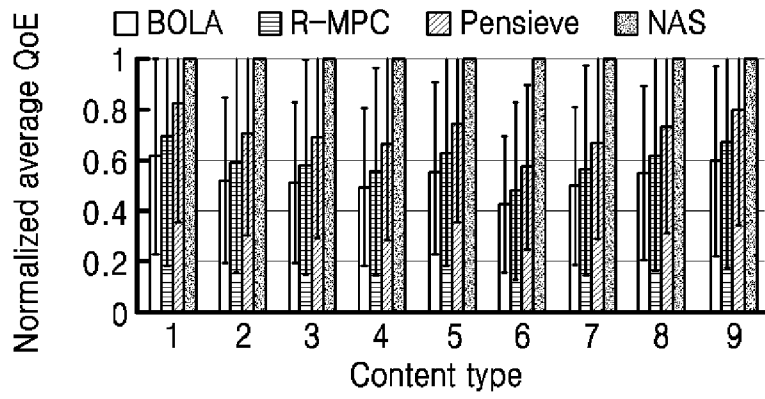
[도6]



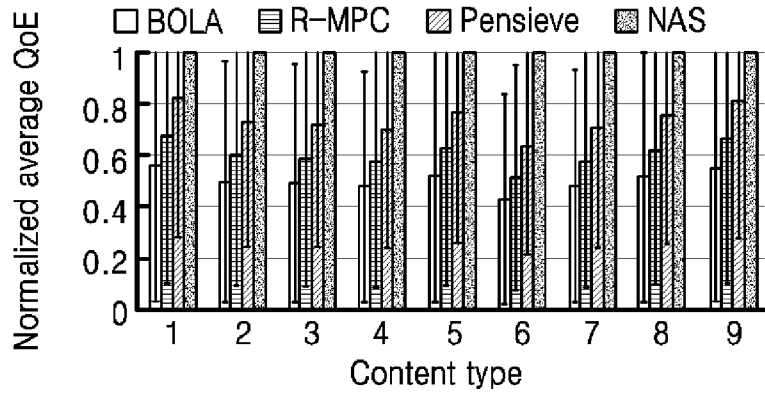
[도7]



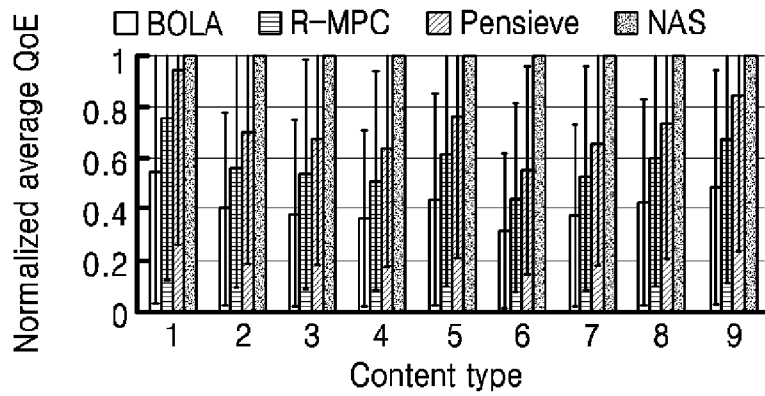
[도 8]



(a) QoE_{lin}

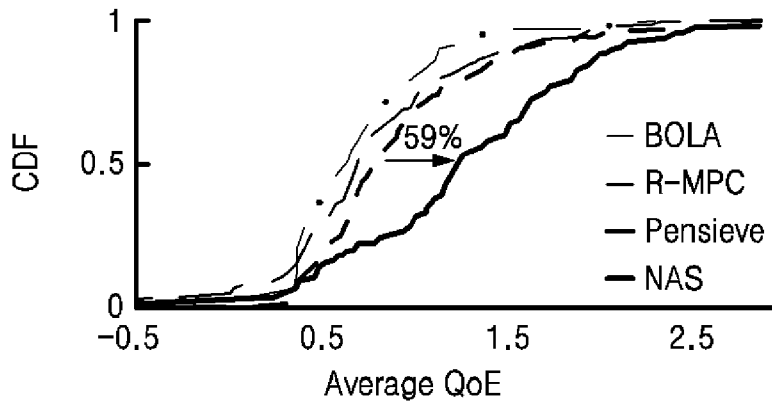


(b) QoE_{log}

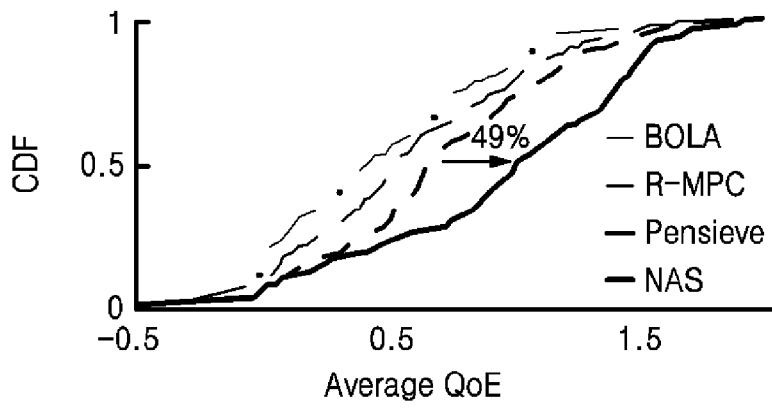


(c) QoE_{hd}

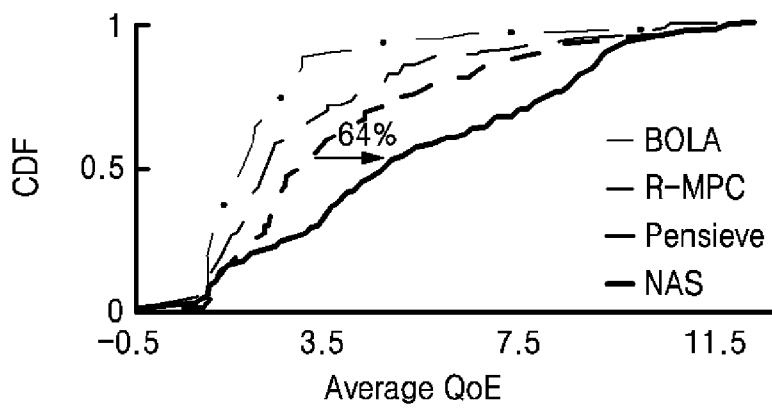
[도9]



(a) QoE_{lin}



(b) QoE_{log}



(c) QoE_{hd}

INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2018/011602

A. CLASSIFICATION OF SUBJECT MATTER

H04N 21/466(2011.01)i, H04N 21/2343(2011.01)i, H04N 21/2385(2011.01)i, H04N 21/274(2011.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04N 21/466; G06F 9/50; G06K 9/62; G06N 3/08; G06T 7/00; G06T 9/00; H04N 21/2343; H04N 21/2385; H04N 21/274

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean utility models and applications for utility models: IPC as above

Japanese utility models and applications for utility models: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS (KIPO internal) & Keywords: content, recognize, neural network, real time, transmit

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	YEO, Hyunho et al. How will Deep Learning Change Internet Video Delivery?. Hot Nets-XVI Proceedings of the 16th ACM Workshop on Hot Topics in Networks. Palo Alto, CA, USA. 30 November-01 December 2017, pages 57-64 See pages 57-62; and figure 1.	1-23
A	KR 10-2018-0094024 A (GOOGLE LLC.) 22 August 2018 See claims 1-8; and figure 4.	1-23
A	KR 10-2018-0004898 A (WAFFLE APPS CO., LTD.) 15 January 2018 See claims 1-9; and figures 1, 3.	1-23
A	WO 2017-055609 A1 (PIKSEL, INC.) 06 April 2017 See claims 1-11; and figures 1-3.	1-23
A	KR 10-2016-0034814 A (SAMSUNG ELECTRONICS CO., LTD.) 30 March 2016 See claims 1-11; and figure 1.	1-23

 Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

20 JUNE 2019 (20.06.2019)

Date of mailing of the international search report

21 JUNE 2019 (21.06.2019)

Name and mailing address of the ISA/KR



Korean Intellectual Property Office
Government Complex Daejeon Building 4, 189, Cheongsa-ro, Seo-gu,
Daejeon, 35208, Republic of Korea

Facsimile No. +82-42-481-8578

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2018/011602

Patent document cited in search report	Publication date	Patent family member	Publication date
KR 10-2018-0094024 A	22/08/2018	CN 108475214 A EP 3380939 A1 JP 2019-510287 A US 2017-0277994 A1 WO 2017-171949 A1	31/08/2018 03/10/2018 11/04/2019 28/09/2017 05/10/2017
KR 10-2018-0004898 A	15/01/2018	None	
WO 2017-055609 A1	06/04/2017	EP 3353745 A1 US 2018-0288440 A1	01/08/2018 04/10/2018
KR 10-2016-0034814 A	30/03/2016	CN 105447498 A US 2016-0086078 A1	30/03/2016 24/03/2016

A. 발명이 속하는 기술분류(국제특허분류(IPC)) H04N 21/466(2011.01)i, H04N 21/2343(2011.01)i, H04N 21/2385(2011.01)i, H04N 21/274(2011.01)i		
B. 조사된 분야 조사된 최소문헌(국제특허분류를 기재) H04N 21/466; G06F 9/50; G06K 9/62; G06N 3/08; G06T 7/00; G06T 9/00; H04N 21/2343; H04N 21/2385; H04N 21/274 조사된 기술분야에 속하는 최소문헌 이외의 문헌 한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우)) eKOMPASS(특허청 내부 검색시스템) & 키워드: 콘텐츠(content), 인지(recognize), 신경망(neural network), 실시간(real-time), 전송(transmit)		
C. 관련 문헌		
카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
X	HYUNHO YEO 등, `How will Deep Learning Change Internet Video Delivery?`, Hot Nets-XVI Proceedings of the 16th ACM Workshop on Hot Topics in Networks, Palo Alto, CA, USA, 2017.11.30-12.01, 페이지 57-64 페이지 57-62; 및 도면 1 참조.	1-23
A	KR 10-2018-0094024 A (구글 엘엘씨) 2018.08.22 청구항 1-8; 및 도면 4 참조.	1-23
A	KR 10-2018-0004898 A (주식회사 와플애플) 2018.01.15 청구항 1-9; 및 도면 1, 3 참조.	1-23
A	WO 2017-055609 A1 (PIKSEL, INC.) 2017.04.06 청구항 1-11; 및 도면 1-3 참조.	1-23
A	KR 10-2016-0034814 A (삼성전자주식회사) 2016.03.30 청구항 1-11; 및 도면 1 참조.	1-23
<input type="checkbox"/> 추가 문헌이 C(계속)에 기재되어 있습니다. <input checked="" type="checkbox"/> 대응특허에 관한 별지를 참조하십시오.		
* 인용된 문헌의 특별 카테고리: “A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌 “E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌 “L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌 “O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌 “P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌 “T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌 “X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다. “Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다. “&” 동일한 대응특허문헌에 속하는 문헌		
국제조사의 실제 완료일 2019년 06월 20일 (20.06.2019)	국제조사보고서 발송일 2019년 06월 21일 (21.06.2019)	
ISA/KR의 명칭 및 우편주소 대한민국 특허청 (35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사) 팩스 번호 +82-42-481-8578	심사관 권성호 전화번호 +82-42-481-3547	

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2018-0094024 A	2018/08/22	CN 108475214 A EP 3380939 A1 JP 2019-510287 A US 2017-0277994 A1 WO 2017-171949 A1	2018/08/31 2018/10/03 2019/04/11 2017/09/28 2017/10/05
KR 10-2018-0004898 A	2018/01/15	없음	
WO 2017-055609 A1	2017/04/06	EP 3353745 A1 US 2018-0288440 A1	2018/08/01 2018/10/04
KR 10-2016-0034814 A	2016/03/30	CN 105447498 A US 2016-0086078 A1	2016/03/30 2016/03/24