

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2022년 1월 20일 (20.01.2022) WIPO | PCT

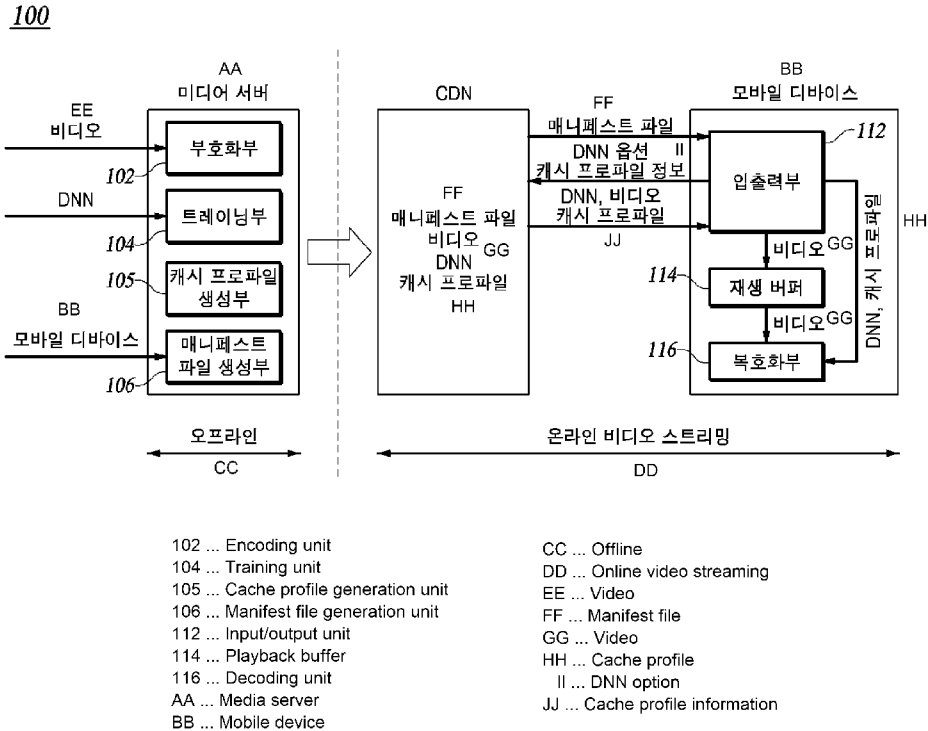


(10) 국제공개번호
WO 2022/014801 A1

- (51) 국제특허분류:
H04N 21/2343 (2011.01) G06N 3/08 (2006.01)
H04N 19/57 (2014.01) G06T 3/40 (2006.01)
- (21) 국제출원번호: PCT/KR2020/018519
- (22) 국제출원일: 2020년 12월 17일 (17.12.2020)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보:
10-2020-0088762 2020년 7월 17일 (17.07.2020) KR
- (71) 출원인: 한국과학기술원 (KOREA ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY)
[KR/KR]: 34141 대전시 유성구 대학로 291, Daejeon (KR).
- (72) 발명자: 한동수 (HAN, Dong Su); 34141 대전시 유성구 대학로 291, Daejeon (KR). 여현호 (YEO, Hyun Ho); 34141 대전시 유성구 대학로 291, Daejeon (KR). 정영목 (JUNG, Young Mok); 34141 대전시 유성구 대학로 291, Daejeon (KR). 정찬주 (CHONG, Chan Ju); 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, IT, JO, JP, KE, KG, KH, KN, KP, 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,

(54) Title: DEVICE AND METHOD FOR SUPER-RESOLUTION ACCELERATION IN REAL-TIME VIDEO STREAMING

(54) 발명의 명칭: 실시간 비디오 스트리밍에서의 초해상도 가속을 위한 장치 및 방법



(57) Abstract: An objective of the present embodiment is to enable reception of real-time video streaming on a mobile device with the computing capacity of the mobile device, while maintaining user quality of experience (QoE) even under bandwidth constraints. To this end, a deep neural network (DNN)-based SR is applied to a small number of pre-selected frames, and the resolution of the remaining frames is improved using the frames to which SR has been applied, wherein the frame to which SR is to be applied is selected within a preconfigured quality margin. The present invention provides a device and a method for SR acceleration, in which, for real-time video streaming according to the computing capacity of a mobile device and under video-specific constraints, a server transfers multiple options for a deep neural network and a cache profile including SR application information, and the mobile device

[다음 쪽 계속]

PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(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))

is capable of selecting an option appropriate for the computing capacity.

(57) 요약서: 본 실시예는, 모바일 디바이스 상에서의 실시간 비디오 스트리밍(real-time video streaming)에 있어서, 모바일 디바이스의 컴퓨팅 용량(computing capacity)으로 수용이 가능하면서도 대역폭 제약 하에서도 사용자 QoE(Quality of Experience)를 유지하는 것을 목적으로 한다. 이를 위하여, 심층신경망(Deep Neural Network: DNN) 기반의 SR을 소수의 기 선정된 프레임에 적용하고, SR이 적용된 프레임을 이용하여 나머지 프레임의 해상도를 향상시키되, 기 설정된 품질 마진 이내에서 SR을 적용하기 위한 프레임을 선별한다. 또한, 모바일 디바이스의 컴퓨팅 용량 및 비디오 특화적인(video-specific) 제약 하에서의 실시간 비디오 스트리밍을 위하여, 서버는 심층신경망에 대한 다중 옵션(multiple option), 및 SR 적용 정보를 포함하는 캐시 프로파일(cache profile)을 전달하고, 모바일 디바이스는 컴퓨팅 용량에 적합한 옵션을 선택하는 것이 가능한 SR 가속을 위한 장치 및 방법을 제공한다.

명세서

발명의 명칭: 실시간 비디오 스트리밍에서의 초해상화 가속을 위한 장치 및 방법

기술분야

- [1] 본 개시는 실시간 비디오 스트리밍에서의 초해상화 가속을 위한 장치 및 방법에 관한 것이다. 더욱 상세하게는, 모바일 디바이스 상에서의 실시간 비디오 스트리밍에 있어서, 모바일 디바이스의 컴퓨팅 용량으로도 수용이 가능하면서도 대역폭 제약 하에서도 사용자 QoE를 유지하는 것이 가능한 초해상화 가속을 위한 장치 및 방법에 대한 것이다.

배경기술

- [2] 이하에 기술되는 내용은 단순히 본 발명과 관련되는 배경 정보만을 제공할 뿐 종래기술을 구성하는 것이 아니다.
- [3] 스마트폰, 태블릿 등과 같은 모바일 디바이스의 보편화에 따라 모바일 비디오 스트리밍(mobile video streaming)에 대한 수요가 폭증하고 있다. 이러한 수요 팽창과 연계하여, 품질에 대한 기대치를 충족시키기 위해 높은 QoE(Quality of Experience)를 사용자에게 제공하는 것이 매우 중요해지고 있다. 증가하는 비디오 수요를 맞추기 위하여, 모바일 캐리어들(mobile carriers) 및 CDN(Contents Delivery Network)들은 대역폭(bandwidth) 확장을 위해 부단히 노력하고 있다. 또한 대역폭 제약(bandwidth constraint)의 상황에서 QoE를 최대화하기 위한 노력에 따라 적응적 스트리밍(adaptive streaming), 초해상화(Super-Resolution: SR) 등의 분야에서 의미있는 진보가 달성되고 있다.
- [4] 대역폭 제약에 대처하기 위한 기술의 하나로서 적응적 스트리밍이 있다(비특허문헌 1). 적응적 스트리밍에서, 서버는 비디오를 복수의 비트레이트(multiple bitrates)로 인코딩한 후, 적당한 길이(예컨대, 2-10 초)의 비디오 뭉치(video chunk)로 분할한다. 클라이언트는 ABR(Adaptive Bitrate) 알고리즘을 이용하여 대역폭 상황에 맞는 적절한 품질의 비디오 뭉치를 선택한다. 비트레이트 및 서버 선택을 최적화하기 위한 개선된 성과에도 불구하고, 적응적 스트리밍은 비디오 품질에 따른 사용자 QoE가 유용한 망 대역폭에 의존한다는 근본적인 문제가 있다.
- [5] 한편, 저해상도 비디오의 품질을 개선하기 위한 기술의 하나로서 SR이 있다. 최근 심층신경망(Deep Neural Network: DNN)을 기반으로 SR의 속도 및 성능 개선에 가시적인 진보가 성취되고 있다. 그러나, 이러한 진보에도 불구하고, SR은 매우 비싼 기술로서, 탑재되는 디바이스의 컴퓨팅 용량에 따라 비디오의 품질이 좌우된다는 문제가 있다.
- [6] ABR 알고리즘과 SR을 결합한 기술의 하나로서 콘텐츠인지 기반의 신경 적응 스트리밍(content-aware Neural Adaptive Streaming: NAS)이 있다(비특허문헌 2

참조). NAS에서, 서버는 비디오에 대하여 콘텐츠인지 기반의 복수의 SR DNN을 트레이닝시킨 후, 트레이닝된 복수의 SR DNN 및 상응하는 비디오 뭉치를 제공한다. 클라이언트는 집적(integrated) ABR 알고리즘을 이용하여 SR DNN 또는 비디오 뭉치를 전송받을지를 결정할 수 있다. 복수의 SR DNN을 획득한 후, 클라이언트는 전송받은 비디오에 SR을 적용하여 고해상도의 비디오를 생성할 수 있다. 대역폭 제약에 따라 낮은 비트레이트로 전송된 저해상도 비디오에 대하여, SR DNN을 이용하여 망 대역폭과는 독립적으로 비디오의 품질 개선이 가능함에도 불구하고, NAS는 클라이언트 측의 컴퓨팅 용량(computing capacity)에 전적으로 의존하므로, 모바일 디바이스에서 실시간 비디오 스트리밍이 허용되기 어렵다는 문제가 있다.

- [7] 기존 기술에서 사용된 데스크탑 수준(desktop-class)의 GPU(Graphics Processing Unit)와 비교하여, 본질적으로 모바일 디바이스는 강한 파워 제약(power constraint)과 연계되어 상대적으로 약한 연산 용량을 가질 수 있다. 예컨대, 영상 처리(image processing)를 위한 SOTA(State of the Art) 모바일 SR조차도(비특허문헌 3 참조), 실시간 비디오 처리에 있어서는 한계를 보인다.
- [8] 따라서, 실시간 비디오 스트리밍에 있어서, 모바일 디바이스의 컴퓨팅 용량으로 수용이 가능하면서도 대역폭 제약 하에서도 사용자 QoE를 유지하는 것이 가능한 SR 가속을 위한 방법이 요구된다.
- [9] (선행기술문헌)
- [10] (비특허문헌)
- [11] (비특허문헌 1) MPEG-DASH Specification. <https://dashif.org/>.
- [12] (비특허문헌 2) Hyunho Yeo, Youngmok Jung, Jaehong Kim, Jinwoo Shin, and Dongsu Han. 2018. Neural adaptive content-aware internet video delivery. In 13th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 18). 645-661.
- [13] (비특허문헌 3) Royson Lee, Stylianos I Venieris, Lukasz Dudziak, Sourav Bhattacharya, and Nicholas D Lane. 2019. MobiSR: Efficient On-Device Super-Resolution through Heterogeneous Mobile Processors. In The 25th Annual International Conference on Mobile Computing and Networking. 1-16.
- [14] (비특허문헌 4) Webm Official Website. <https://www.webmproject.org/>.
- [15] (비특허문헌 5) Google's Exoplayer Official Website. <https://developer.android.com/guide/topics/media/exoplayer>.
- [16] (비특허문헌 6) Xiaomi Mi9 Specifications. https://www.gsmarena.com/xiaomi_mi_9-9507.php.
- [17] (비특허문헌 7) <https://www.youtube.com/>
- [18] (비특허문헌 8) Wowza's DASH bitrate recommendation. <https://www.wowza.com/docs/how-to-encode-source-video-for-wowza-streaming-cloud>.

- [19] (비특허문헌 9) Hongzi Mao, Ravi Netravali, and Mohammad Alizadeh. 2017. Neural Adaptive Video Streaming with Pensieve. In Proceedings of the ACM Special Interest Group on Data Communication (SIGCOMM). 197-210.

발명의 상세한 설명

기술적 과제

- [20] 본 개시는, 본 개시는, 모바일 디바이스 상에서의 실시간 비디오 스트리밍(real-time video streaming)에 있어서, 모바일 디바이스의 컴퓨팅 용량(computing capacity)으로 수용이 가능하면서도 대역폭 제약 하에서도 사용자 QoE(Quality of Experience)를 유지하는 것을 목적으로 한다. 이를 위하여, 심층신경망(Deep Neural Network: DNN) 기반의 SR을 소수의 기 선정된 프레임에 적용하고, SR이 적용된 프레임을 이용하여 나머지 프레임의 해상도를 향상시키되, 기 설정된 품질 마진(quality margin) 이내에서 SR을 적용하기 위한 프레임을 선별한다. 또한, 모바일 디바이스의 컴퓨팅 용량 및 비디오 특화적인(video-specific) 제약 하에서의 실시간 비디오 스트리밍을 위하여, 서버는 심층신경망에 대한 다중 옵션(multiple option), 및 SR 적용 정보를 포함하는 캐시 프로파일(cache profile)을 전달하고, 모바일 디바이스는 컴퓨팅 용량에 적합한 옵션을 선택하는 것이 가능한 SR 가속을 위한 장치 및 방법을 제공하는 데 주된 목적이 있다.

기술적 해결방법

- [21] 본 개시의 실시예에 따르면, 모바일 디바이스가 실행하는 SR(Super-Resolution) 가속을 위한 방법에 있어서, 비디오에 대한 매니페스트 파일(manifest file)을 서버로부터 획득하는 과정; 상기 매니페스트 파일로부터 SR DNN(Deep Neural Network)에 대한 옵션(option)과 캐시 프로파일(cache profile)에 대한 정보를 선택하는 과정; 상기 옵션과 캐시 프로파일에 대한 정보를 상기 서버 측으로 전송하여 상기 옵션에 대응하는 SR DNN, 캐시 프로파일 및 비디오 뭉치(chunk)를 상기 서버로부터 획득하는 과정; 및 상기 SR DNN을 이용하여 SR 집적 복호화기(SR-integrated decoder)를 초기화한 후, 상기 SR 집적 복호화기를 이용하여 상기 캐시 프로파일을 기반으로 상기 비디오 뭉치를 형성하는 현재 프레임에 대한 해상도를 향상시키는 과정을 포함하는 것을 특징으로 하는 SR 가속을 위한 방법을 제공한다.
- [22] 본 개시의 다른 실시예에 따르면, 모바일 디바이스 상에 탑재되는 SR(Super-Resolution) 가속을 위한 장치에 있어서, 비디오에 대한 매니페스트 파일(manifest file)을 서버로부터 획득하고, 상기 매니페스트 파일로부터 SR DNN(Deep Neural Network)에 대한 옵션(option)과 캐시 프로파일(cache profile)에 대한 정보를 선택하며, 상기 옵션과 캐시프로파일에 대한 정보를 상기 서버 측으로 전송하여 상기 옵션에 대응하는 SR DNN, 캐시 프로파일 및 비디오 뭉치(chunk)를 상기 서버로부터 획득하는 입출력부; 및 상기 SR DNN을

이용하여 SR 집적 복호화기(SR-integrated decoder)를 초기화한 후, 상기 SR 집적 복호화기를 이용하여 상기 캐시 프로파일을 기반으로 상기 비디오 문치를 형성하는 현재 프레임에 대한 해상도를 향상시키는 복호화부를 포함하는 것을 특징으로 하는 SR 가속을 위한 장치를 제공한다.

- [23] 본 개시의 다른 실시예에 따르면, 서버가 실행하는 SR(Super-Resolution) 가속을 위한 방법에 있어서, 비디오를 획득하는 과정; 영상 부호화기(image encoder)를 이용하여 상기 비디오로부터 복수의 비트레이트(multiple bitrate)에 해당하는 비디오 버전(video version)을 생성하는 과정; 상기 비디오 및 상기 복수의 비디오 버전을 이용하여 상응하는 복수의 SR 심층신경망(Deep Neural Network: DNN)을 트레이닝시키는 과정; 기 설정된 품질 마진 이내에서 상기 비디오 버전으로부터 상기 각 SR DNN에 대한 앵커 프레임(anchor frame)을 선별하고, 상기 앵커 프레임에 대한 정보를 포함하는 캐시 프로파일(cache profile)을 생성하는 과정; 및 상기 캐시 프로파일에 대한 정보, 및 상기 복수의 SR DNN에 대한 정보를 포함하는 매니페스트 파일(manifest file)을 생성하는 과정을 포함하는 것을 특징으로 하는 SR 가속을 위한 방법을 제공한다.

- [24] 본 개시의 다른 실시예에 따르면, 모바일 디바이스가 실행하는 SR 가속을 위한 방법이 포함하는 각 단계를 실행시키기 위하여 컴퓨터로 읽을 수 있는 기록매체에 저장된 컴퓨터프로그램을 제공한다.

- [25] 본 개시의 다른 실시예에 따르면, 서버가 실행하는 SR 가속을 위한 방법이 포함하는 각 단계를 실행시키기 위하여 컴퓨터로 읽을 수 있는 기록매체에 저장된 컴퓨터프로그램을 제공한다.

발명의 효과

- [26] 이상에서 설명한 바와 같이 본 실시예에 따르면, 모바일 디바이스 상에서의 비디오 스트리밍에 있어서, 심층신경망 기반의 SR을 소수의 기 설정된 프레임에 적용하고, SR이 적용된 프레임을 이용하여 나머지 프레임의 해상도를 향상시키는 SR 가속을 위한 장치 및 방법을 제공함으로써, 모바일 디바이스의 비디오 처리량(processing throughput)을 증가시키고, 에너지 소모를 감소시켜 발열을 적정 수준으로 유지하며, 대역폭 제약 하에서도 사용자 QoE를 개선시키는 것이 가능해지는 효과가 있다.

- [27] 또한 본 실시예에 따르면, 모바일 디바이스 상에서의 비디오 스트리밍에 있어서, 서버는 심층신경망에 대한 다중 옵션, 및 SR 적용 정보를 포함하는 캐시 프로파일을 전달하고, 모바일 디바이스는 컴퓨팅 용량에 적합한 옵션을 선택하는 것이 가능한 SR 가속을 위한 장치 및 방법을 제공함으로써, 모바일 디바이스의 컴퓨팅 용량 및 비디오 특화적인(video-specific) 제약 하에서의 실시간 비디오 스트리밍이 가능해지는 효과가 있다.

도면의 간단한 설명

- [28] 도 1은 본 개시의 일 실시예에 따른 SR 가속을 위한 장치에 대한 개략적인

블록도이다.

- [29] 도 2는 본 개시의 일 실시예에 따른 SR 가속을 위한 방법의 순서도이다.
- [30] 도 3은 영상 복호화기 내 인터 예측을 개념적으로 설명한 예시도이다.
- [31] 도 4는 상용 영상 부호화기가 생성한 프레임 의존성을 나타내는 그래프의 예시도이다.
- [32] 도 5는 본 개시의 일 실시예에 따른 앵커 프레임을 선별하는 방법의 순서도이다.
- [33] 도 6은 본 개시의 일 실시예에 따른 품질 이득의 집합에 대한 예시도이다.
- [34] 도 7은 본 개시의 일 실시예에 따른 SR 집적 복호화기의 동작을 개념적으로 설명한 블록도이다.
- [35] 도 8은 본 개시의 일 실시예에 따른 비앵커 프레임에 대한 인터 예측의 개념적인 예시도이다.

발명의 실시를 위한 형태

- [36] 이하, 본 발명의 실시예들을 예시적인 도면을 참조하여 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 실시예들을 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 실시예들의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략한다.
- [37] 또한, 본 실시예들의 구성요소를 설명하는 데 있어서, 제 1, 제 2, A, B, (a), (b) 등의 용어를 사용할 수 있다. 이러한 용어는 그 구성요소를 다른 구성요소와 구별하기 위한 것일 뿐, 그 용어에 의해 해당 구성요소의 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.
- [38] 첨부된 도면과 함께 이하에 개시될 상세한 설명은 본 발명의 예시적인 실시형태를 설명하고자 하는 것이며, 본 발명이 실시될 수 있는 유일한 실시형태를 나타내고자 하는 것이 아니다.
- [39] 본 실시예는 실시간 비디오 스트리밍에서의 SR(Super-Resolution) 가속을 위한 장치 및 방법에 대한 내용을 개시한다. 보다 자세하게는, 모바일 디바이스 상에서의 실시간 비디오 스트리밍에 있어서, 모바일 디바이스의 컴퓨팅 용량(computing capacity)으로도 수용이 가능하면서도 대역폭 제약 하에서도 사용자 QoE(Quality of Experience)를 유지하는 것이 가능한 SR 가속을 위한 장치 및 방법에 대한 것이다.

- [40] 이하, 본 개시에 따른 SR 가속을 위한 장치 및 방법은 클라이언트로서 스마트폰, 태블릿 등과 같은 모바일 디바이스를 이용하고 있으나, 반드시 이에 한정하는 것은 아니며, 클라이언트의 대상이 데스크탑, 셋탑 박스 등과 같이 컴퓨팅 용량이 충분한 장치로도 확장될 수 있다.
- [41] 이하, 하나의 비디오는 복수의 비디오 뭉치(video chunk)를 포함하는 것으로 가정한다. 하나의 비디오 뭉치는 적어도 하나의 GOP(Group of Pictures)를 포함하고, GOP는 복수의 프레임을 포함하는 것으로 가정한다.
- [42] 도 1은 본 개시의 일 실시예에 따른 SR 가속을 위한 장치에 대한 개략적인 블록도이다.
- [43] 모바일 디바이스 상에서의 실시간 비디오 스트리밍에 있어서, 본 실시예에 따른 SR 가속을 위한 장치(100)는 심층신경망(Deep Neural Network: DNN) 기반의 SR을 소수의 기 선정된 프레임에 적용하고, SR이 적용된 프레임을 이용하여 나머지 프레임의 해상도를 향상시키되, 기 설정된 품질 마진 이내에서 SR을 적용하기 위한 프레임을 선별한다. SR 가속을 위한 장치(100, 이하 'SR 가속장치')는 미디어 서버(media server, 이하 '서버'), CDN(Contents Delivery Network) 및 모바일 디바이스에 분산되어 탑재됨으로써 실시간 비디오 스트리밍을 구현할 수 있다.
- [44] 서버에 탑재된 SR 가속장치(100)는 실시간 비디오 스트리밍을 위한 준비 과정을 실행하는데, 이러한 준비 과정은 오프라인(offline)으로 실행될 수 있다. 서버는 준비한 데이터를 CDN 측으로 전달한다. 한편, CDN과 모바일 디바이스에 탑재된 SR 가속장치(100) 간의 실시간 비디오 스트리밍 과정은 온라인(online)으로 실행될 수 있다.
- [45] 도 1의 도시는 본 실시예에 따른 예시적인 구성이며, 실시간 비디오 스트리밍을 위한 미디어 서버 구조와 동작, CDN의 구조와 동작, 및 모바일 디바이스의 구조와 동작에 따라 다른 구성요소 또는 구성요소 간의 다른 연결을 포함하는 다양한 구현이 가능하다.
- [46] 실시간 비디오 스트리밍을 준비하기 위하여 서버는 비디오로부터 상이한 비트레이트를 갖는 복수의 비디오 버전(video version)을 생성하고, 비디오 버전 각각에 상응하는 SR DNN을 트레이닝시키며, 기 설정된 품질 마진(quality margin) 이내에서 SR을 적용하기 위한 앵커 프레임(anchor frame)을 각 SR DNN마다 선별한다. 이를 위하여 서버는 부호화부(102), 트레이닝부(104), 캐시 프로파일(cache profile) 생성부(105) 및 매니페스트 파일(manifest file) 생성부(106)의 전부 또는 일부를 포함할 수 있다.
- [47] 부호화부(102)는, 서버에 고해상도 비디오가 업로드(upload)되면, 영상 부호화기(image encoder)를 이용하여 고해상도 비디오로부터 상이한 비트레이트를 갖는 복수의 비디오 버전을 생성한다. 예컨대, 고해상도 비디오가 1080p인 경우, 복수의 비디오 버전은 240p, 360p, 또는 480p 등과 같은 저해상도의 비디오를 포함할 수 있다.

- [48] 트레이닝부(104)는 고해상도 비디오 및 복수의 비디오 버전을 이용하여 상응하는 복수의 SR DNN을 트레이닝시킨다.
- [49] 본 실시예에 있어서 DNN으로는 영상 신호처리(image signal processing)에 적합한 것으로 알려진 CNN(convolutional Neural Network) 기반의 신경망이 이용되나(비특허문헌 2 참조), 반드시 이에 한정하는 것은 아니며, 영상 변환이 가능한 어느 형태의 신경망이든 이용될 수 있다. 이하, 특별한 언급이 없다면 DNN과 SR DNN은 동일한 의미이고, 상호 교차적으로 사용될 수 있다.
- [50] 캐시 프로파일 생성부(105) 품질 마진 이내에서 비디오 버전으로부터 SR을 적용하기 위한 앵커 프레임(anchor frame)을 각 SR DNN마다 선별하고, 앵커 프레임에 대한 정보를 포함하는 캐시 프로파일을 생성한다. 예를 들어, 앵커 프레임에 대한 정보로서, 앵커 프레임과 나머지 프레임을 구별하기 위한 1 비트의 데이터가 각 프레임 별로 할당될 수 있다.
- [51] 여기서, 품질 마진은 설정이 가능한 파라미터로서, 비디오 또는 비디오 버전의 매 프레임에 SR을 적용한 경우의 품질에 대한 마진이다. 서버는 품질 마진 및 각각의 비디오 버전에 따라 상이한 앵커 프레임을 선별하고, 상응하는 캐시 프로파일을 생성할 수 있다.
- [52] 매니페스트 파일 생성부(106)는 캐시 프로파일에 대한 정보, 및 복수의 SR DNN에 대한 정보를 포함하는 매니페스트 파일을 생성한다. 여기서, 캐시 프로파일에 대한 정보는 CDN 측에 저장되는 캐시 프로파일의 위치로서, 예를 들어, 캐시 프로파일을 전송받을 수 있는 URL(Uniform Resource Locator)에 대한 정보일 수 있다.
- [53] 각각의 모바일 디바이스(또는 모바일 디바이스가 포함하는 프로세서)는 상이한 컴퓨팅 용량을 갖는다. 따라서, 이러한 제약 조건들 하에서 실시간 처리를 지원하기 위하여, 서버는 각 비디오 버전 별로 상이한 성능을 갖는 복수의 옵션들(options)을 제공할 수 있다. 여기서, 복수의 옵션들은 복수의 DNN을 의미할 수 있다. 따라서 서버는 비디오 버전 및 성능 옵션에 따라 상이한 DNN을 제공할 수 있다. 트레이닝부(104)는 비디오 버전 및 성능 옵션에 따른 상이한 복수의 DNN을 트레이닝시킬 수 있다.
- [54] 매니페스트 파일은 각 비디오 버전과 컴퓨팅 용량에 따른 DNN의 옵션들, 및 품질 마진과 DNN 옵션들에 따른 캐시 프로파일에 대한 정보를 포함한다. 또한 매니페스트 파일은 각 DNN의 옵션들 각각이 지원할 수 있는 모바일 디바이스의 리스트를 포함할 수 있다.
- [55]
- [56] *서버는 비디오 버전, DNN, 캐시 프로파일 및 매니페스트 파일을 포함하는 데이터를 CDN에 제공하고, CDN은 제공받은 데이터를 이용하여 모바일 디바이스에 대한 실시간 비디오 스트리밍을 실행한다.
- [57] 모바일 디바이스가 비디오를 요청하는 경우, CDN은 해당하는 매니페스트 파일을 모바일 디바이스 측으로 전송한다. 또한 매니페스트 파일로부터 선택된

DNN에 대한 옵션, 및 캐시 프로파일에 대한 정보를 모바일 디바이스로부터 획득한 후, CDN은 옵션에 맞는 DNN, 캐시 프로파일 및 비디오 버전의 뭉치를 모바일 디바이스 측으로 전송한다.

- [58] 모바일 디바이스는 매니페스트 파일로부터 DNN에 대한 옵션과 캐시 프로파일에 대한 정보를 선택한다. 또한 모바일 디바이스는 선택된 옵션과 캐시 프로파일에 대한 정보를 CDN 측으로 전송하여 DNN, 캐시 프로파일 및 비디오를 형성하는 프레임들을 CDN으로부터 획득하고, DNN을 이용하여 SR 집적 복호화기(SR-integrated decoder)를 초기화한 후, SR 집적 복호화기를 이용하여 캐시 프로파일을 기반으로 각 프레임에 대한 해상도를 향상시킨다. 이를 위하여 모바일 디바이스는 입출력부(112), 재생 버퍼(playback buffer, 114) 및 복호화부(116)의 전부 또는 일부를 포함할 수 있다.
- [59] 입출력부(112)는 재생을 원하는 비디오에 대한 매니페스트 파일을 CDN으로부터 획득하여, 매니페스트 파일로부터 비디오 버전 별 DNN 각각에 대한 옵션과 캐시 프로파일에 대한 정보를 선택한다. DNN 각각에 대한 옵션을 선택하는 것은, 비디오 버전 별로 DNN을 각각 선택한다는 의미이다. 매니페스트 파일은 각 DNN의 옵션들이 지원하는 모바일 디바이스 리스트를 포함하고, 이 리스트를 이용하여 모바일 디바이스의 컴퓨팅 용량(computing capacity)에 적합한 옵션이 선택될 수 있다.
- [60] 본 실시예는 적응적 스트리밍이 적용되는 상황을 가정하므로, 가용한 대역폭에 의존하여 상이한 비디오 버전의 뭉치가 전달될 수 있다. 따라서, 모바일 디바이스는 각 비디오 버전 별 DNN을 필요로 하고, 또한, 각 DNN 별로 모바일 디바이스의 컴퓨팅 용량에 적합한 옵션을 선택할 수 있다.
- [61] 입출력부(112)는 선택한 옵션과 캐시 프로파일에 대한 정보를 CDN 측으로 전송하여 옵션에 맞는 DNN, 캐시 프로파일 및 비디오 버전의 뭉치(chunk)를 CDN으로부터 획득한다. 입출력부(112)는 각 비디오 버전 별 옵션에 적합한 DNN, 캐시 프로파일 및 비디오 버전의 뭉치(chunk)를 획득한다. 이때, 입출력부(112)는 집적 ABR 알고리즘을 이용할 수 있다(비특허문헌 2 참조). 따라서, 각 DNN 옵션에 대하여, DNN 전체가 획득된 후에 모바일 디바이스는 SR 가속을 실행하거나, 확장가능(scalable) DNN인 경우 전체 DNN의 일부가 수신되면 SR 가속을 실행할 수 있다.
- [62] 재생 버퍼(114)는 다운로드된 비디오 버전의 뭉치를 저장한다.
- [63] 복호화부(116)는 DNN을 이용하여 SR 집적 복호화기를 초기화한 후, SR 집적 복호화기를 이용하여 캐시 프로파일을 기반으로 각 프레임에 대한 해상도를 향상시킨다. 전송한 바와 같이 캐시 프로파일은 비디오 뭉치를 형성하는 프레임이 DNN을 적용하기 위한 앵커 프레임인지 여부에 대한 정보를 포함한다.
- [64] 복호화부(116)는 SR 집적 복호화기를 이용함에 있어서, 캐시 프로파일에 저장된 앵커 프레임 정보에 따라, 앵커 프레임인 경우 DNN을 적용하거나, 앵커 프레임이 아닌 경우 캐시된 고해상도의 이전 프레임을 이용하여, 현재

- 프레임(current frame)에 대한 고해상도 프레임을 생성한 후 추후 이용을 위하여 캐시할 수 있다.
- [65] 복호화부(116)는 비디오 버전 별 DNN을 이용하여 SR 집적 복호화기를 초기화할 수 있다. 따라서, 전송받은 비디오 버전에 따라 SR 집적 복호화기가 이용하는 DNN이 변경될 수 있다.
- [66] 도 2는 본 개시의 일 실시예에 따른 SR 가속을 위한 방법의 순서도이다.
- [67] 도 2a의 도시는 서버/CDN에서 진행하는 SR 가속을 위한 방법에 대한 순서도이다.
- [68] 서버는 영상 부호화기(image encoder)를 이용하여 고해상도 비디오로부터 상이한 비트레이트를 갖는 복수의 비디오 버전을 생성한다(S200).
- [69] 서버는 고해상도 비디오 및 복수의 비디오 버전을 이용하여 상응하는 복수의 SR DNN을 트레이닝시킨다(S202).
- [70] 서버는 앵커 프레임에 대한 정보를 포함하는 캐시 프로파일을 생성한다(S203). 캐시 프로파일을 생성하기 위하여, 서버는 기 설정된 품질 마진 이내에서 SR을 적용하기 위한 앵커 프레임을 각 SR DNN마다 선별한다. 품질 마진 및 각각의 비디오 버전에 따라 상이한 앵커 프레임이 선별되고, 상응하는 캐시 프로파일이 생성될 수 있다.
- [71] 서버는 캐시 프로파일에 대한 정보, 및 복수의 SR DNN에 대한 정보를 포함하는 매니페스트 파일을 생성한다(S204).
- [72] 서버는 각 DNN 별로 상이한 사양을 갖는 복수의 옵션들을 제공할 수 있다. 여기서, 복수의 옵션들은 복수의 DNN을 의미할 수 있다. 따라서, 매니페스트 파일은 각 비디오 버전과 컴퓨팅 용량에 따른 DNN에 대한 복수의 옵션들, 및 CDN 측에 저장되는 캐시 프로파일의 위치에 대한 정보를 포함할 수 있다.
- [73] 비디오 및 모바일 디바이스 특화적인(specific) 제약에 대하여, 서버가 앵커 프레임을 선별하는 방법, 및 복수의 성능 옵션을 준비하는 과정은 추후에 상세하게 설명하기로 한다.
- [74] 서버는 비디오 버전, DNN, 캐시 프로파일 및 매니페스트 파일을 포함하는 데이터를 CDN에 제공하고, CDN은 제공받은 데이터를 이용하여 모바일 디바이스에 대한 실시간 비디오 스트리밍을 실행한다.
- [75] 모바일 디바이스가 비디오를 요청하는 경우, CDN은 해당하는 매니페스트 파일을 모바일 디바이스 측으로 전송한다(S206).
- [76] 매니페스트 파일로부터 선택된 SR DNN에 대한 옵션과 캐시 프로파일에 대한 정보를 모바일 디바이스로부터 획득한 후, CDN은 옵션에 맞는 SR DNN, 캐시 프로파일 및 비디오 버전의 뭉치를 모바일 디바이스 측으로 전송한다(S208).
- [77] 도 2b의 도시는 모바일 디바이스에서 진행하는 SR 가속을 위한 방법에 대한 순서도이다.
- [78] 모바일 디바이스는 재생을 원하는 비디오에 대한 매니페스트 파일을 CDN으로부터 획득한다(S220).

- [79] 모바일 디바이스는 매니페스트 파일로부터 DNN에 대한 옵션과 캐시 프로파일에 대한 정보를 선택한다(S222). 매니페스트 파일은 각 DNN의 옵션들이 지원하는 모바일 디바이스 리스트를 포함하고, 이 리스트를 이용하여 모바일 디바이스의 컴퓨팅 용량(computing capacity)에 적합한 옵션이 선택될 수 있다.
- [80] 모바일 디바이스는 선택한 옵션과 캐시 프로파일에 대한 정보를 CDN 측으로 전송하여 옵션에 맞는 DNN, 캐시 프로파일 및 비디오의 뭉치를 CDN으로부터 획득한다(S224).
- [81] 모바일 디바이스는 DNN을 이용하여 SR 집적 복호화기를 초기화한 후, SR 집적 복호화기를 이용하여 캐시 프로파일을 기반으로 각 프레임에 대한 해상도를 향상시킨다(S226). 캐시 프로파일은 비디오 뭉치를 형성하는 프레임이 DNN을 적용하기 위한 앵커 프레임인지 여부에 대한 정보를 포함한다. 캐시 프로파일을 이용하여 SR 집적 복호화기가 실행하는 영상 복호화 과정에 대하여는 추후에 상세하게 설명하기로 한다.
- [82] 이하, 도 3 내지 도 6의 도시를 이용하여 캐시 프로파일에 포함되는 앵커 프레임을 선별하는 방법에 대하여 설명한다.
- [83] 객체 분류(object classification) 및 검출(detection)에 적용되는 DNN은 후단 레이어 측으로 갈수록 특성 맵(feature map)을 축소시킨다. 따라서 DNN 내 앞단의 컨볼루션 레이어(convolution layer) 정보를 캐시 후 재사용하는 방식이, 풍부한 정보를 재활용한다는 측면에서 분류 성능의 향상에 도움이 되는 것으로 알려져 있다.
- [84] 이에 반하여, 고해상도 이미지를 복원(reconstruction)하기 위해 SR DNN은 후단 레이어 측으로 갈수록 특성 맵을 확대시킨다. 또한, 컴퓨팅 특성(computing characteristics) 측면에서, SR DNN 내 마지막에 위치하는 컨볼루션 레이어가 연산 지연(latency)의 대부분을 차지한다. 따라서, 연산량 절약 및 정보량 측면에서 전단에 위치하는 레이어를 캐시하는 것은 효과적이지 못하므로, 본 실시예에서는 SR DNN의 출력이 캐시된 후 재사용된다.
- [85] 도 3에 도시된 바와 같이, GOP 중 첫 프레임을 제외하고(첫 프레임은 인트라 예측(intra prediction)을 이용), 일반적인 영상 복호화기(image decoder)는 이미 복원된 채로 캐시되어 있는 프레임에 의존하는 인터 예측을 이용하여 현재 프레임을 복원한다. 영상 복호화기는 참조 인덱스(reference index)가 지정하는 참조 프레임(reference frame)을 획득하고, 참조 프레임 및 움직임벡터(motion vector)를 이용하여 인터 예측(inter prediction) 블록을 생성한다. 영상 복호화기는 인터 예측 블록과 영상 부호화기로부터 전달된 잔차 신호(residue signal)를 가산함으로써 최종적인 복원 블록(reconstruction block)을 생성할 수 있다. 영상 복호화기는 복원 블록들을 결합하여 복원 프레임을 생성하고, 생성된 복원 프레임은 다시 캐시되어, 이후의 인터 예측에 참조 프레임으로 이용될 수 있다.
- [86] 따라서, 인터 예측의 프레임 의존성(frame dependency)을 이용하기 위해, 본

실시예에 따른 SR 집적 복호화기는 앵커 프레임에 SR을 적용하여 고해상도 프레임을 복원한 후 캐시하고, 인터 예측을 이용하여 캐시된 프레임을 전이(transfer)함으로써 나머지 비앵커 프레임들을 복원한다. SR이 적용되는 프레임에 대한 연산량이 인터 예측을 이용하는 비앵커 프레임들에 대한 연산량보다 수백 배 수준으로 월등하게 많으므로, 앵커 프레임에 대한 SR 적용은 연산 지연을 비앵커 프레임 측으로 분산시킬 수 있다. 이러한 연산량의 분산에 따라 모바일 디바이스에서의 실시간 스트리밍이 가능해질 수 있다.

- [87] 앵커 프레임의 선별을 위하여 도 4에 도시된 바와 같은, 영상 부호화기가 생성한 프레임 의존성이 참조될 수 있다. 도 4의 도시는 상용 영상 부호화기(비특허문헌 4 참조)가 생성한 프레임 의존성을 나타내며, 특정 비디오 내 하나의 GOP에 포함된 15 개의 프레임에 대한 의존성 그래프(dependency graph)를 표현한다. 상용 부호화기는 참조도가 높은 세 가지의 특이 프레임(special frame)으로서, 키 프레임(key frame), 대안 참조 프레임(alternative reference frame), 및 골든 프레임(golden frame)을 생성한다. 여기서 키 프레임은 GOP의 첫 번째 프레임이고, 대안 참조 프레임은 인터 예측만을 보조하기 위해 삽입된 보이지 않는(invisible) 프레임이며, 골든 프레임은 복수로 참조되는 프레임을 나타낸다. 이들을 제외한 나머지 프레임들은 전체 프레임의 80 % 이상을 차지함에도 기껏해야 하나의 의존 프레임을 갖는 것으로 알려져 있다.
- [88] 모바일 디바이스의 컴퓨팅 용량을 절약한다는 측면에서, 앵커 프레임의 개수는 적을수록 유리하다. 따라서, 이러한 특이 프레임들을 주로 앵커 프레임으로 선별하여, SR을 적용한 결과가 재사용됨으로써, 나머지 다수의 프레임에 대한 품질 개선이 가능해질 수 있다.
- [89] 전술한 바와 같이 캐시된 고해상도 프레임을 이용하여 다른 프레임의 해상도를 높이는 경우, 품질 저하가 필연적으로 발생하는데, 이러한 품질 저하를 캐시 손실(cache erosion)로 표현한다. 캐시 손실의 크기는 비디오 콘텐츠 의존적(content-dependent)이다. 이하, 캐시 손실은 매 프레임에 SR을 적용한 경우와 본 실시예와 같이 앵커 프레임에만 SR을 적용한 경우 간의 PSNR(Peak Signal to Noise Ratio) 차이로 정의된다. 랜덤 또는 균등(uniform)하게 앵커 프레임을 선정하는 대신, 본 실시예에서는 캐시 손실이 기 설정된 품질 마진 이내로 유지되도록, 최적의 앵커 프레임이 선별된다.
- [90] 최적의 앵커 프레임을 선별하는 것은, 모바일 디바이스의 컴퓨팅 용량을 절약한다는 측면에서 선별된 앵커 프레임의 개수를 최소화하는 것이다. 최소의 앵커 프레임을 선별하기 위한 최적화 목표(optimization goal)는 수학적 식 1과 같이 나타낼 수 있다.

[91] [수식1]

$$\min_{\{AF\}} |\{AF\}| \text{ where } \{AF\} \subset \{F\}$$

$$s. t \ VQ(DNN(\{F\})) - VQ(DNN(\{AF\})) \leq VQ_T$$

[92] 여기서, $VQ(DNN(\{ \cdot \}))$ 은, SR이 적용되지 않은 경우에 대비하여, $\{ \cdot \}$ 프레임에 SR DNN을 적용하는 경우의 품질 개선으로서, 두 경우 간 PSNR 차이로 표현될 수 있다. $\{F\}$ 는 비디오 프레임의 전체 집합이고, $\{AF\}$ 는 앵커 프레임의 집합이며, $| \cdot |$ 는 대상의 크기, 즉 개수를 의미한다. VQ_T 는 기 설정된 품질 마진이다. 수학적 식 1에 나타낸 바는, 매 프레임에 SR DNN을 적용하는 경우와 앵커 프레임에만 SR을 적용한 경우 간의 품질 개선 차이, 즉 캐시 손실이 품질 마진을 만족하도록, 최소의 AF를 선별한다는 의미이다.

[93] 수학적 식 1에 나타낸 바와 같이 앵커 프레임을 선별하기 위하여, 서버는 $2^{|frame|}$ 수준의 검색 공간을 검색하여야 한다. 여기서, $|frame|$ 은 비디오(또는 비디오 문치)를 형성하는 전체 프레임의 개수이다. 이는 서버의 컴퓨팅 용량으로도 실현 가능하기 어려운 수준이다. 따라서, 검색 공간을 줄이기 위해, 앵커 프레임이 드문드문한(sparse) 경우, 나머지 프레임들에 대한 품질 개선은 가장 영향력이 큰 앵커 프레임에 의존한다고 가정한다.

[94] 이러한 전제 하에, 앵커 프레임의 집합 $\{AF\}$ 가 주어진 경우 임의의 프레임에 대한 품질 이득은, 수학적 식 2에 나타낸 바와 같이 $\{AF\}$ 중 하나의 앵커 프레임이 주어진 경우 하나의 프레임에 대한 품질 이득 중 최대값으로 근사될 수 있다.

[95] [수식2]

$$FQ(i|DNN(\{AF\})) = \max_{f \in \{AF\}} FQ(i|DNN(\{f\}))$$

[96] 여기서, $FQ(i|DNN(\{ \cdot \}))$ 은, 앵커 프레임들에 SR DNN을 적용하는 경우, i 번째 프레임에 대한 품질 개선을 나타낸다. 또한 $DNN(\{f\})$ 은 하나의 앵커 프레임 f 에만 SR DNN을 적용하는 것을 나타낸다.

[97] 수학적 식 2에 나타낸 바와 같은 하나의 프레임에 대한 이득을 이용하여, 수학적 식 3에 나타낸 바와 같이 전체 프레임에 대한 평균을 계산함으로써, 비디오 품질이 근사될 수 있다.

[98] [수식3]

$$VQ(DNN(\{AF\})) \approx \sum_{i=1}^{|frame|} \max_{f \in \{AF\}} \frac{FQ(i|DNN(\{f\}))}{|frame|}$$

[99] 수학적 식 3에 나타낸 바와 같이, 크기가 1인 앵커 프레임들의 모든 가능한 조합에 대한 품질 측정에 근거하여, 서버는 임의의 앵커 프레임 집합에 대한 품질 이득을 산정함으로써, 검색 공간의 크기를 $2^{|frame|}$ 에서 $|frame|$ 수준으로 감소시킬 수 있다.

[100] 도 5는 본 개시의 일 실시예에 따른 앵커 프레임을 선별하는 방법의 순서도이다.

[101] 서버는 변수를 초기화한다(S500). 하나의 앵커 프레임 f 에만 SR DNN을 적용했을 때의 i 번째 프레임의 품질 이득 $FQ = FQ(i|DNN(\{f\}))$ 의 집합 $\{FQ\}$ 가

공집합으로 설정된다. 또한 선별된 앵커 프레임의 집합 $\{AF'\}$ 도 공집합으로 설정된다. 그리고, 앵커 프레임의 집합 $\{AF'\}$ 에 SR이 적용되는 경우의 비디오 뭉치 전체에 대한 품질 이득 $VQ = VQ(DNN(\{AF'\}))$ 이 영(zero)으로 설정된다.

[102] 서버는 비디오 뭉치를 획득하여 프레임의 집합 $\{F\}$ 를 생성한다(S502). 프레임의존성을 고려하여, 비디오 뭉치는 하나 이상의 GOP의 그룹을 포함하는 것으로 가정한다. 비디오 뭉치에 포함된 전체 프레임의 개수는 $lframe$ 이다. 따라서 $1 \leq i, f \leq lframe$ 이다.

[103] 서버는 $\{F\}$ 에 포함되는 프레임 f 각각이 앵커 프레임인 경우에 대하여 $FQ(i|DNN(\{f\}))$ 를 산정하여 $\{FQ\}$ 를 생성한다(S504). $lframe$ 개의 프레임에 대하여, 프레임 f 만이 앵커 프레임인 경우, i 번째 프레임의 품질 이득 $FQ(i|DNN(\{f\}))$ 를 산정하여, 도 6에 나타난 바와 같이 $\{FQ\}$ 를 생성할 수 있다. 이때, FQ 를 산정하기 위해 SR 집적 복호화기가 이용될 수 있다.

[104] 서버는 모든 프레임에 SR이 적용되는 경우의 비디오 품질 이득 $VQ(DNN(\{F\}))$ 과 VQ 의 간의 차이를 기 설정된 품질 마진과 비교한다(S506).

[105] 서버는, 기 설정된 품질 마진보다 큰 경우, 하나의 신규 앵커 프레임 AF 를 선별하여 $\{AF'\}$ 에 포함시킨다(S508).

[106] 먼저, $\{AF'\}$ 이 공집합인 경우, 도 6에 예시된 하나의 행(예컨대, f 프레임이 앵커 프레임인 행)의 FQ 의 합(또는 평균)이 최대로 판정되면, 해당되는 프레임을 첫 앵커 프레임으로 선별한다. 따라서 $\{AF'\}=\{f\}$ 가 된다. 다음, 두 번째 앵커 프레임은 수학적 4에 나타난 바와 같이 선별될 수 있다.

[107] [수식4]

$$\max_{AF^c \in \{F\}} VQ(DNN(AF^c \cup \{AF'\})) \approx \max_{AF^c \in \{F\}} \left(\sum_{i=1}^{|frame|} \max_{f \in \{AF^c \cup \{AF'\}\}} \frac{FQ(i|DNN(\{f\}))}{|frame|} \right)$$

[108] 서버는 $\{AF'\}$ 에 후보 프레임 AF^c 하나를 추가하고, 먼저 수학적 3을 이용하여 $VQ(DNN(\{AF'\} \cup AF^c))$ 를 계산한다. 예컨대, 도 6에 예시된 바와 같이 $\{AF'\}=\{f\}$ 이고 두 번째 프레임이 AF^c 인 경우, $\{f\} \cup AF^c$ 에 대하여 수학적 3을 이용하여 $VQ(DNN(\{f\} \cup AF^c))$ 가 계산될 수 있다.

[109] AF^c 는 $\{AF'\}$ 에 포함되지 않은 어느 프레임이든 될 수 있으므로, 수학적 4에 나타난 바와 같이 $VQ(DNN(\{AF'\} \cup AF^c))$ 을 최대로 하는 후보 프레임을 신규 앵커 프레임 AF 로 선별하고 $\{AF'\}$ 에 포함시킨다. 예컨대, 도 6의 예시에서, 2 번째 프레임이 AF^c 이고, $VQ(DNN(\{f\} \cup AF^c))$ 가 최대로 판정된 경우, AF^c 가 신규 앵커 프레임 AF 로 선별될 수 있다.

[110] 서버는 전술한 바와 같은 과정을 이용하여, 세 번째 이후의 앵커 프레임도 선별될 수 있다.

[111] 서버는 $\{AF'\}$ 에 대하여 $VQ(DNN(\{AF'\}))$ 을 산정함으로써 VQ 를 업데이트한 후(S510), 이를 비디오 품질 이득 $VQ(DNN(\{F\}))$ 와 비교하는 단계(S506)를

반복한다. 이때, $VQ(DNN(\{AF'\}))$ 를 산정하기 위해 SR 집적 복호화기가 이용될 수 있다.

- [112] 서버는, 기 설정된 품질 마진 이내인 경우, 현재까지 선별된 $\{AF'\}$ 을 앵커 프레임으로 확정한다(S512).
- [113] 이상에서 설명된 앵커 프레임을 선별하는 방법은 비디오 고유의 특성에 적응하여 동작할 수 있다. 즉, 수학적 4의 근사식의 우측 항에 포함된 최대 선택 기능($\max_i$) 때문에, 영향력이 크거나 참조가 빈번하게 되는 프레임일수록 앵커 프레임으로 선별될 가능성이 높다. 또한 급속한 캐시 손실을 보이는 비디오일수록 기 설정된 품질 마진을 만족시키기 위해 더 많은 앵커 프레임이 선별될 수 있다.
- [114] 한편, 품질 마진은 비디오의 매 프레임에 SR을 적용한 경우의 품질에 대한 마진으로서, 설정 가능한 파라미터이고 설정치에 따라 상이한 캐시 프로파일이 생성될 수 있다. 품질 마진은 통상 0.5 dB로 설정되나, 배터리 소모와 관련하여서도 복수의 설정을 가질 수 있는데, 이에 대하여는 추후에 설명한다.
- [115] 이하, 도 7 내지 도 8의 도식을 이용하여 SR 집적 복호화기가 실행하는 영상 복호화 방법(S226 단계)에 대하여 설명한다.
- [116] 도 7은 본 개시의 일 실시예에 따른 SR 집적 복호화기의 동작을 개념적으로 설명한 블록도이다.
- [117] SR 집적 복호화기는 캐시 프로파일 및 SR DNN을 이용하여 압축된 저해상도 프레임을 고해상도 프레임으로 변환한다. 압축된 저해상도 프레임을 획득한 후, SR 집적 복호화기는 캐시 프로파일을 이용하여 획득한 프레임이 앵커 프레임인지 여부를 확인한다. 앵커 프레임인 경우, SR 집적 복호화기는 현재 프레임으로부터 저해상도 프레임을 복호화한 후, SR DNN을 적용하여 해상도를 증가시킨 고해상도 프레임을 생성하고, 추후 이용을 위해 캐시한다. 반면, 비앵커 프레임(non-anchor frame)인 경우, 프레임 의존성에 대한 정보 및 캐시된 고해상도 프레임들을 이용하여, SR 집적 복호화기는 현재 프레임에 대한 고해상도 프레임을 생성하고, 추후 이용을 위해 캐시한다.
- [118] 한편 현재 프레임은 상호 겹치지 않는 인트라 예측 블록 및 인터 예측 블록을 포함하고, 이러한 블록을 기준으로 복호화된다.
- [119] 도 8은 본 개시의 일 실시예에 따른 비앵커 프레임에 대한 인터 예측의 개념적인 예시도이다.
- [120] 현재 프레임이 비앵커 프레임인 경우, 캐시된 고해상도 프레임을 참조 프레임으로 이용하여, SR 집적 복호화기는 인터 예측 블록을 복원할 수 있다. 인터 예측을 실행하기 위하여 SR 집적 복호화기는 영상 부호화기로부터 인터 예측 블록에 대한 참조 인덱스, 움직임벡터를 전달받는다. 참조 인덱스를 이용하여 SR 집적 복호화기는 캐시되어 있는 이전에 복원된 고해상도 프레임들로부터 참조 프레임을 선택한다. 또한 움직임벡터의 크기가 조절되는데, 예를 들어, 360p로부터 1080p로 해상도를 증가시키는 경우,

움직임벡터도 3 배만큼 확대된다.

- [121] SR 집적 복호화기는 조절된 움직임벡터를 이용하여 참조 프레임으로부터 대상 블록(target block)을 예측하는 움직임 보상(motion compensation)을 실행함으로써 현재 프레임의 인터 예측 블록을 생성한다. 적절한 보간 방법을 이용하여(예컨대, 이중선형 보간(bilinear interpolation) 방법), SR 집적 복호화기는 영상 부호화기로부터 전달된 잔차 신호의 해상도를 증가시킨 후, 인터 예측 블록과 가산함으로써 고해상도 복원 블록을 생성할 수 있다. 보간 방법을 이용하여 해상도를 증가시키는 것은 고주파 대역에서의 손실을 유발하고, 결론적으로 캐시 손실을 발생시킬 수 있다. 이러한 캐시 손실의 일부는 전술한 바와 같은 앵커 프레임의 선별 방법을 이용하여 보상될 수 있다.
- [122] SR 집적 복호화기는 동일 프레임 내의 주변 픽셀을 이용하여 인트라 예측 블록을 생성한다. 캐시된 프레임을 이용하지 못하므로, 인트라 예측은 쉽지 않은 문제이다. 다만, 대부분의 인트라 예측 블록이, 높은 참조 의존도로 인하여 앵커 프레임으로 자주 선별되는 키 프레임 또는 대안 참조 프레임에 포함된다. 따라서, 앵커 프레임에 포함되는 인트라 예측 블록에 대해서는 SR DNN을 적용하고, 비앵커 프레임에 포함되는 인트라 예측 블록에 대해서는 보간을 적용함으로써 인트라 예측이 해결될 수 있다.
- [123] 이하, 비디오 및 모바일 디바이스 특화적인 제약에 대하여, 서버가 다양한 성능 옵션을 준비하는 상세한 과정을 설명한다.
- [124] 온라인 비디오 스트리밍을 실행하기 위해, 비디오 스트리밍은 실시간으로(예컨대, 30 fps(frame per second)) 처리되어야 한다. 그러나, 기 설정된 품질 마진을 만족시키기 위하여, 서버는 앵커 프레임을 선별하는 방법을 이용하여 각 비디오 뭉치 별로 상이한 수의 앵커 프레임을 생성할 수 있다. 또한, 다양한 컴퓨팅 용량을 보유한다는 측면에서 모바일 디바이스들은 비균질적(heterogeneous) 특성을 갖는다.
- [125] 이러한 디바이스 및 비디오 특화적인 제약 하에서 실시간 스트리밍을 처리하기 위해 서버는 각 비디오 버전 별로 복수의 성능 옵션을 제공한다(예컨대, 'Low', 'Mid' 및 'High'). 각각의 성능 옵션에 대하여, 서버는 다양한(varying) 품질 및 컴퓨팅 요구사항을 갖는, 분리된 DNN을 제공할 수 있다. 또한 각각의 DNN에 대하여, 서버는 앵커 프레임을 선별하는 방법을 이용하여 캐시 프로파일을 생성할 수 있다.
- [126] 모바일 디바이스는 자신의 컴퓨팅 용량을 기반으로 주어진 복수의 성능 옵션 중의 하나를 선택할 수 있다. 모바일 디바이스가 옵션을 선택하기 용이하도록, 본 실시예에 따른 서버는 디바이스 풀(pool)에 의존하는 측정 결과를 이용하여 가이드라인을 제시한다. 여기서, 디바이스 풀은 각 모바일 디바이스(또는 모바일 프로세서)에서 다양한 옵션을 실행시키기 위한 설정(set up)을 의미한다.
- [127] 한편, 각 모바일 디바이스는 각각의 비디오에 맞는 옵션을 찾는 것을 필요로 한다. 그러나, 방대한 양의 비디오 숫자를 고려한다면, 서버조차도 각각의

비디오에 대하여 모든 가능한 옵션을 테스트하는 것이 용이하지 않다.

대안적으로, 앵커 프레임 및 비앵커 프레임의 처리 지연(*processing latency*)을 산정하기 위하여, 서버는 각 모바일 디바이스로 하여금 샘플 비디오에 대하여 하나의 옵션을 실행하도록 한다. 샘플 비디오에 대한 측정 결과에 근거하여, 수학적 식 5에 나타난 바와 같이, 다른 대상 비디오에 대한 대체적인 처리 지연이 추정될 수 있다.

[128] [수식5]

$$Latency = |AF| \times T(AF) + |None_AF| \times T(None_AF)$$

[129] 여기서, $|AF|$ 및 $|None_AF|$ 는 대상 비디오의 앵커 프레임 및 비앵커 프레임의 개수를 나타낸다. $T(AF)$ 는 및 $T(None_AF)$ 는 샘플 비디오로부터 측정된 앵커 프레임 및 비앵커 프레임의 처리 지연을 나타내는데, 모바일 디바이스에 의존하는 값이다. 비디오에 포함된 비디오 뭉치 각각에 대하여 수학적 식 5에 나타난 바와 같은 처리 지연 추정을 반복실행(*iteration*)함으로써, 서버는 비디오가 발생시키는 최악의 처리 지연을 추정하고, 이를 성능 옵션에 반영할 수 있다. 각 모바일 디바이스는 실시간 제약을 만족시키는 한도 내에서 가장 고품질의 옵션을 선택할 수 있다.

[130] 모바일 디바이스의 옵션 선택을 보조하기 위하여, 옵션 선택을 위한 정보를 포함한 채로 매니페스트 파일이 서버로부터 모바일 디바이스 측으로 제공된다. 각 모바일 디바이스 및 DNN 옵션에 대하여, 서버는 전술한 바와 같은, 성능 옵션을 준비하는 과정을 오프라인으로 한 차례 실행함으로써, 매니페스트 파일을 생성할 수 있다. 매니페스트 파일은 각 DNN 별 옵션, 및 옵션이 지원할 수 있는 모바일 디바이스의 리스트를 포함할 수 있다.

[131] 한편, 일부 사용자는 다른 어떤 요소보다 배터리 지속 시간에 더 민감할 수 있다. 이러한 사용자 선호도를 수용하기 위해, 서버는 복수의 품질 마진을 기반으로 배터리 성능에 대한 다양한 옵션을 제공할 수 있다. 배터리 성능 옵션 및 배터리 소모 상태를 참고하여, 사용자는 비디오 품질과 에너지 소모에 대한 동적인 절충(*dynamic trade-off*)을 시도할 수 있다.

[132] 이상에서 설명한 바와 같이 본 실시예에 따르면, 모바일 디바이스 상에서의 비디오 스트리밍에 있어서, 서버는 심층신경망에 대한 다중 옵션, 및 SR 적용 정보를 포함하는 캐시 프로파일을 전달하고, 모바일 디바이스는 컴퓨팅 용량에 적합한 옵션을 선택하는 것이 가능한 SR 가속을 위한 장치 및 방법을 제공함으로써, 모바일 디바이스의 컴퓨팅 용량 및 비디오 특화적인(*video-specific*) 제약 하에서의 실시간 비디오 스트리밍이 가능해지는 효과가 있다.

[133] 이하 본 실시예에 따른 SR 가속장치(100)의 성능을 평가한 결과를 설명한다.

[134] 성능 평가 실험을 위해, 상용 영상 부호화기(비특허문헌 4 참조)를 개조하여 서버 측의 SR 가속장치를 구현하고, 상용 영상 복호화기(비특허문헌 5 참조)를

개조하여 모바일 디바이스 측의 SR 가속장치를 구현하였다. SR DNN에 대한 모델 및 트레이닝은 NAS(비특허문헌 2 참조)의 방식을 이용하되, 본 실시예에 따른 구현에서는 메모리 사용량을 줄이기 위해 서브픽셀(sub-pixel) 컨볼루션을 디컨볼루션으로 대체하고, SR 실현의 이득은 적으나 처리 지연이 큰 마지막 컨볼루션 레이어가 제거된 SR DNN이 사용되었다.

- [135] 성능 평가를 위한 모바일 디바이스로는 고급(high-end) 스마트폰인 제1 디바이스(비특허문헌 6 참조)가 사용되었다. 입문용(entry-level) 스마트폰, 태블릿 등의 모바일 디바이스가 추가로 성능 평가에 이용되었으나, 제1 디바이스와 유사한 측정 결과를 보였으므로, 이들에 대한 언급은 생략한다.
- [136] 비디오 데이터세트로는 상용 CDN 사이트(비특허문헌 7 참조) 상의 탑텐(top-ten) 카테고리 각각으로부터 획득된 비디오가 이용되었다. 여기서, 10 개의 카테고리는 'Product review'(C1), 'Howto'(C2), 'Vlogs'(C3), 'Game play'(C4), 'Skit'(C5), 'Haul'(C6), 'Challenges'(C7), 'Favorite'(C8), 'Education'(C9), 및 'Unboxing'(C10)이다. 각 비디오는 30 fps에서 4K 해상도를 지원하고 최소 5분의 길이를 갖는다. 실험에서는 적응적 스트리밍을 위해, Wowza의 권고(비특허문헌 8 참조)에 따라, 상용 영상 부호화기를 이용하여 각 비디오는 {240, 360, 480, 720, 1080}p 해상도 각각에 대하여 {512, 1024, 1600, 2640, 4400}kbps의 비트레이트를 갖는 비디오 버전으로 부호화된다. GOP의 크기는 120 프레임(4 초에 해당함)이다. PSNR 측정을 위한 기준으로는 해상도가 1080p인 원시 비디오가 이용되었다. 각 비디오 버전에 대한 성능 측정에는 5 분의 재생 시간이 이용되었다.
- [137] 제1 베이스라인으로는 매 프레임마다 SR DNN이 적용되는 per-frame SR이 이용되는데, 비디오 버전 별로 옵션에 따라 per-frame Low, per-frame Mid, per-frame High의 세 가지 SR DNN이 이용되었다. 제2 베이스라인으로서 SR DNN이 적용되는 대신 보간 방법이 이용되었다. 한편, 제1 디바이스는 {240, 360, 480}p 비디오를 1080p로 해상도를 향상시키는 것으로 가정한다.
- [138] 표 1은 본 실시예에 따른 비디오 처리량(processing throughput)을 나타낸다.
- [139]

[표 1]

(비디오 처리량(fps))

	옵션*	본 실시예	Per-frame Low	Per-frame Medium	Per-frame High
C1	H	84.5	7.1	6.9	5.2
C2	H	72.7	7.1	7.1	5.3
C3	M	99.0	7.1	6.7	5.4
C4	H	96.2	7.1	6.9	5.2
C5	M	59.4	7.0	6.7	5.2
C6	M	95.2	7.1	6.7	5.2
C7	H	89.9	7.0	6.8	5.3
C8	M	109.3	7.1	6.4	5.4
C9	M	119.8	7.1	6.8	5.1
C10	H	94.7	10.6	6.5	5.7

*옵션은 본 실시예가 선택한 옵션임.

[140] 앵커 프레임에 대한 SR 적용 효과를 검증하기 위해, 비디오 버전의 해상도는 240p로 고정되었다. 제1 디바이스 상의 SR 가속장치(100)는 SR 집적 복호화기를 이용하여 비디오 버전의 해상도를 1080p로 향상시킨다. SR 가속장치(100)는 비디오 카테고리 별로 상이한 성능을 갖는 SR DNN을 선택하였다(H: High, M: Medium). SR 가속장치(100)는 59 내지 120 fps를 달성함으로써, per-frame SR과 비교하여, 평균 처리량을 15배 이상 개선하였다. 이러한 처리량 개선은 전체 프레임 중 일부분(3.1 내지 10.9 %)만이 앵커 프레임으로 선별됨으로써, SR 적용에 따른 지연이 비앵커 프레임 측으로 분산된 것에 기인한다.

[141] 표 2는 본 실시예에 따른 품질 이득(quality gain)을 나타낸다.

[142]

[표2]

(PSNR 이득(dB))

	옵션*	본 실시예	Per-frame Low	Per-frame Medium	Per-frame High
C1	H	1.48	1.19	1.7	1.9
C2	H	1.67	1.48	1.88	2.11
C3	M	0.98	1.22	1.44	1.6
C4	H	1.31	1.16	1.52	1.75
C5	M	2.05	1.7	2.51	2.87
C6	M	3.5	3.16	3.95	4.38
C7	H	2.15	1.86	2.31	2.56
C8	M	1.33	1.45	1.77	1.96
C9	M	3.31	2.58	3.67	4.14
C10	H	3.09	1.85	2.72	3.45

*옵션은 본 실시예가 선택한 옵션임.

- [143] 사용된 비디오 버전의 해상도는 240p이다. 전술한 바와 같이 1080p인 원시 비디오를 기준으로 본 실시예 및 제1 베이스라인(per-frame SR)의 PSNR 품질이 YUV420 색공간(color space)을 이용하여 측정되었다. 표 2에 나타난 품질 이득은, 제2 베이스라인에 따라 생성된 비디오의 PSNR에 대한 이득을 나타낸다. SR 가속장치(100)는 26.0 내지 37.8 dB의 PSNR을 달성함으로써, 제2 베이스라인에 따른 비디오의 PSNR과 비교하여, 탁월한 품질 이득(0.98 내지 3.5 dB)을 성취하였다. 또한, per-frame SR과 비교하더라도, SR 가속장치(100)는 0.5 dB 이내의 성능 차이를 보이며, 일부 카테고리에 대하여 per-frame Low 및/또는 per-frame Medium보다 개선된 품질 이득을 보인다.
- [144] 한편, 에너지 소모는 C10 카테고리를 이용하여 측정되었으며, 선택된 SR DNN은 High였다. SR 가속장치(100)는 per-frame High와 비교하여 88 % 이상 에너지 소모를 감소시킴으로써, 배터리 사용 시간을 5시간 이상 증가시켰다. 또한, 이러한 에너지 소모 감소에 따라 장시간의 테스트에서도 제1 디바이스의 온도가 적정 수준(사용자가 불편함을 느끼기 시작하는 온도는 섭씨 35 도임)보다 매우 낮게 유지되었다.
- [145] 본 실시예에 따른 SR 가속장치(100)의 실시간 스트리밍에 있어서, QoE 개선이 측정되었다. 이를 위하여 Pensieve(비특허문헌 9 참조)의 실제 3G/광대역 망 흔적(traces)을 이용하되, 적응적 스트리밍에 따른 이득을 제공하지 않는 4.4 Mbps 이상의 대역폭을 갖는 흔적은 필터링 아웃되었다. 사용된 망 흔적의 평균 대역폭은 1.3 Mbps이다. 망 흔적 상에서 적응적 스트리밍을 실행하기 위해, NAS에서 사용된 집적 ABR 알고리즘을 포함하도록 Pensieve 시뮬레이터를 확장하여, DNN 및 비디오가 동시에 스트리밍되도록 하였다.

- [146] 한편, QoE 메트릭(metric)으로는 NAS에서 사용된 메트릭을 이용했는데, 1) 각 비디오 뭉치에 대하여 선택된 비트레이트, 2) 재버퍼링 시간(rebuffering time), 및 3) 연속된 비디오 뭉치 간 품질 차이 등을 기반으로 형성된다. 비디오 뭉치에 대한 비트레이트로서 SR에 따른 품질 개선을 반영하기 위한 유효(effective) 비트레이트가 계산되었는데, 실험에서는 PSNR 품질을 유효 비트레이트로 변환하는 함수가 고안되어 이용되었다.
- [147] 제1 디바이스 상의 SR 가속장치(100)를 이용하여 10 개의 카테고리에 대하여 실시간 스트리밍의 QoE가 측정되었고, 제3 베이스라인으로는 SR이 적용되지 않는 Pensieve ABR 알고리즘에 기반하는 적응형 스트리밍이 이용되었다. 제3 베이스라인 대비하여 SR 가속장치(100)는 평균 32 %의 QoE 개선을 보였다.
- [148] 또한, SR 가속장치(100)는 QoE를 개선하는 대신, 대역폭 사용(bandwidth usage)을 절감할 수도 있다. 대역폭 절감을 측정하기 위해, 제3 베이스라인과 동일한 QoE가 될 때까지, SR 가속장치(100)가 이용하는 대역폭을 감소시켰다. 측정된 대역폭 절감은 제3 베이스라인 대비하여 평균적으로 21 % 정도였다.
- [149] 본 실시예에 따른 앵커 프레임의 선별 방법에 있어서 검색 공간의 감소에 대한 페널티(penalty)가 측정되었다. 검색 공간을 감소시키기 위해, 수학적 식 3에 나타난 바와 같은 품질 이득의 근사가 실행되었는데, 이에 따른 평균 PSNR 손실은 0.11 dB로서 감내할 수 있는 수준인 것으로 확인되었다.
- [150] 본 실시예에 따른 앵커 프레임의 선별 방법에 따른 캐시 손실이 측정되었다. 본 실시예에 따라 선별된 앵커 프레임을 이용하는 경우, 비디오 뭉치 당 평균 캐시 손실이 0.5 dB 이내로 제한되어, 랜덤하게 선별된 경우의 평균 캐시 손실 2.1 dB 및 균등하게 선별된 경우의 평균 캐시 손실 1.3 dB와 비교하여 탁월하게 적은 것으로 확인되었다.
- [151] 또한, 상용 영상 부호화기(비특허문헌 4 참조)에 본 실시예에 따른 앵커 프레임 선별 방법을 적용하는 경우, 선별된 앵커 프레임의 80 % 정도가 프레임 의존도가 높은 키 프레임 및 대안 참조 프레임으로 구성되는 것이 확인되었다.
- [152] 이상에서 설명한 바와 같이 본 실시예에 따르면, 모바일 디바이스 상에서의 실시간 비디오 스트리밍에 있어서, 심층신경망 기반의 SR을 소수의 기 선정된 프레임에 적용하고, SR이 적용된 프레임을 이용하여 나머지 프레임의 해상도를 향상시키는 SR 가속을 위한 장치 및 방법을 제공함으로써, 모바일 디바이스의 비디오 처리량(processing throughput)을 증가시키고, 에너지 소모를 감소시켜 발열을 적정 수준으로 유지하며, 대역폭 제약 하에서도 사용자 QoE를 개선시키는 것이 가능해지는 효과가 있다.
- [153] 본 실시예에 따른 각 순서도에서는 각각의 과정을 순차적으로 실행하는 것으로 기재하고 있으나, 반드시 이에 한정되는 것은 아니다. 다시 말해, 순서도에 기재된 과정을 변경하여 실행하거나 하나 이상의 과정을 병렬적으로 실행하는 것이 적용 가능할 것이므로, 순서도는 시계열적인 순서로 한정되는 것은 아니다.
- [154] 본 명세서에 설명되는 시스템들 및 기법들의 다양한 구현예들은, 디지털 전자

회로, 집적 회로, FPGA(field programmable gate array), ASIC(application specific integrated circuit), 컴퓨터 하드웨어, 펌웨어, 소프트웨어, 및/또는 이들의 조합으로 실현될 수 있다. 이러한 다양한 구현예들은 프로그래밍가능 시스템 상에서 실행가능한 하나 이상의 컴퓨터 프로그램들로 구현되는 것을 포함할 수 있다. 프로그래밍가능 시스템은, 저장 시스템, 적어도 하나의 입력 디바이스, 그리고 적어도 하나의 출력 디바이스로부터 데이터 및 명령들을 수신하고 이들에게 데이터 및 명령들을 전송하도록 결합되는 적어도 하나의 프로그래밍가능 프로세서(이것은 특수 목적 프로세서일 수 있거나 혹은 범용 프로세서일 수 있음)를 포함한다. 컴퓨터 프로그램들(이것은 또한 프로그램들, 소프트웨어, 소프트웨어 애플리케이션들 혹은 코드로서 알려져 있음)은 프로그래밍가능 프로세서에 대한 명령어들을 포함하며 "컴퓨터가 읽을 수 있는 기록매체"에 저장된다.

- [155] 컴퓨터가 읽을 수 있는 기록매체는, 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 이러한 컴퓨터가 읽을 수 있는 기록매체는 ROM, CD-ROM, 자기 테이프, 플로피디스크, 메모리 카드, 하드 디스크, 광자기 디스크, 스토리지 디바이스 등의 비휘발성(non-volatile) 또는 비일시적인(non-transitory) 매체일 수 있으며, 또한 캐리어 웨이브(예를 들어, 인터넷을 통한 전송) 및 데이터 전송 매체(data transmission medium)와 같은 일시적인(transitory) 매체를 더 포함할 수도 있다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수도 있다.
- [156] 본 명세서에 설명되는 시스템들 및 기법들의 다양한 구현예들은, 프로그램가능 컴퓨터에 의하여 구현될 수 있다. 여기서, 컴퓨터는 프로그램가능 프로세서, 데이터 저장 시스템(휘발성 메모리, 비휘발성 메모리, 또는 다른 종류의 저장 시스템이거나 이들의 조합을 포함함) 및 적어도 한 개의 커뮤니케이션 인터페이스를 포함한다. 예컨대, 프로그램가능 컴퓨터는 서버, 네트워크 기기, 셋탑 박스, 내장형 장치, 컴퓨터 확장 모듈, 개인용 컴퓨터, 랩탑, PDA(Personal Data Assistant), 클라우드 컴퓨팅 시스템 또는 모바일 장치 중 하나일 수 있다.
- [157] 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.
- [158] (부호의 설명)
- [159] 100: SR 가속을 위한 장치

[160] 102: 부호화부 104: 트레이닝부

[161] 105: 캐시 프로파일 생성부 106: 매니페스트 파일 생성부

[162] 112: 입출력부 114: 재생 버퍼

[163] 116: 복호화부

[164]

[165] CROSS-REFERENCE TO RELATED APPLICATION

[166] 본 특허출원은 2020년 7월 17일 한국에 출원한 특허출원번호

제10-2020-0088762 호에 대해 우선권을 주장하며, 그 모든 내용은 참고문헌으로
본 특허출원에 병합된다.

청구범위

- [청구항 1] 모바일 디바이스가 실행하는 SR(Super-Resolution) 가속을 위한 방법에 있어서,
비디오에 대한 매니페스트 파일(manifest file)을 서버로부터 획득하는 과정;
상기 매니페스트 파일로부터 SR DNN(Deep Neural Network)에 대한 옵션(option)과 캐시 프로파일(cache profile)에 대한 정보를 선택하는 과정;
상기 옵션과 캐시 프로파일에 대한 정보를 상기 서버 측으로 전송하여 상기 옵션에 대응하는 SR DNN, 캐시 프로파일 및 비디오 뭉치(chunk)를 상기 서버로부터 획득하는 과정; 및
상기 SR DNN을 이용하여 SR 집적 복호화기(SR-integrated decoder)를 초기화한 후, 상기 SR 집적 복호화기를 이용하여 상기 캐시 프로파일을 기반으로 상기 비디오 뭉치를 형성하는 현재 프레임에 대한 해상도를 향상시키는 과정
을 포함하는 것을 특징으로 하는 SR 가속을 위한 방법.
- [청구항 2] 제1항에 있어서,
상기 매니페스트 파일은 상기 SR DNN의 옵션이 지원하는 모바일 디바이스 리스트를 더 포함하고, 상기 리스트를 이용하여 상기 모바일 디바이스의 컴퓨팅 용량(computing capacity)에 적합한 옵션을 선택하는 것을 특징으로 하는 SR 가속을 위한 방법.
- [청구항 3] 제1항에 있어서,
상기 캐시 프로파일은 상기 현재 프레임이 상기 SR DNN을 적용하기 위한 앵커 프레임(anchor frame)인지 여부에 대한 데이터를 포함하고, 상기 캐시 프로파일에 대한 정보는 상기 서버 측에 저장되는 캐시 프로파일의 위치인 것을 특징으로 하는 SR 가속을 위한 방법.
- [청구항 4] 제3항에 있어서,
상기 해상도를 향상시키는 과정은,
상기 현재 프레임이 상기 앵커 프레임인 경우, 상기 SR 집적 복호화기가 상기 현재 프레임으로부터 저해상도 프레임을 복호화한 후, 상기 SR DNN을 적용하여 상기 저해상도 프레임에 대한 고해상도 프레임을 생성 및 캐시하는 것을 특징으로 하는 SR 가속을 위한 방법.
- [청구항 5] 제3항에 있어서,
상기 해상도를 향상시키는 과정은,
상기 현재 프레임이 상기 앵커 프레임이 아닌 경우, 프레임 의존성(frame dependency)에 대한 정보 및 캐시된 이전 고해상도 프레임들을 기반으로, SR 집적 복호화기가 인터 예측 및 인트라 예측을 이용하여 상기 현재 프레임에 대한 고해상도 프레임을 생성 및 캐시하는 것을 특징으로 하는

SR 가속을 위한 방법.

- [청구항 6] 모바일 디바이스 상에 탑재되는 SR(Super-Resolution) 가속을 위한 장치에 있어서,
비디오에 대한 매니페스트 파일(manifest file)을 서버로부터 획득하고,
상기 매니페스트 파일로부터 SR DNN(Deep Neural Network)에 대한 옵션(option)과 캐시 프로파일(cache profile)에 대한 정보를 선택하며, 상기 옵션과 캐시프로파일에 대한 정보를 상기 서버 측으로 전송하여 상기 옵션에 대응하는 SR DNN, 캐시 프로파일 및 비디오 뭉치(chunk)를 상기 서버로부터 획득하는 입출력부; 및
상기 SR DNN을 이용하여 SR 집적 복호화기(SR-integrated decoder)를 초기화한 후, 상기 SR 집적 복호화기를 이용하여 상기 캐시 프로파일을 기반으로 상기 비디오 뭉치를 형성하는 현재 프레임에 대한 해상도를 향상시키는 복호화부를 포함하는 것을 특징으로 하는 SR 가속을 위한 장치.
- [청구항 7] 제6항에 있어서,
상기 매니페스트 파일은 상기 SR DNN의 옵션이 지원하는 모바일 디바이스 리스트를 더 포함하고, 상기 리스트를 이용하여 상기 모바일 디바이스의 컴퓨팅 용량(computing capacity)에 적합한 옵션을 선택하는 것을 특징으로 하는 SR 가속을 위한 장치.
- [청구항 8] 제6항에 있어서,
상기 캐시 프로파일은 상기 현재 프레임이 상기 SR DNN을 적용하기 위한 앵커 프레임(anchor frame)인지 여부에 대한 데이터를 포함하고, 상기 캐시 프로파일에 대한 정보는 상기 서버 측에 저장되는 캐시 프로파일의 위치인 것을 특징으로 하는 SR 가속을 위한 장치.
- [청구항 9] 제8항에 있어서,
상기 복호화부는,
상기 현재 프레임이 상기 앵커 프레임인 경우, 상기 SR 집적 복호화기가 상기 현재 프레임으로부터 저해상도 프레임을 복호화한 후, 상기 SR DNN을 적용하여 상기 저해상도 프레임에 대한 고해상도 프레임을 생성 및 캐시하는 것을 특징으로 하는 SR 가속을 위한 장치.
- [청구항 10] 제8항에 있어서,
상기 복호화부는,
상기 현재 프레임이 상기 앵커 프레임이 아닌 경우, 프레임 의존성(frame dependency)에 대한 정보 및 캐시된 이전 고해상도 프레임들을 기반으로, SR 집적 복호화기가 인터 예측 및 인트라 예측을 이용하여 상기 현재 프레임에 대한 고해상도 프레임을 생성 및 캐시하는 것을 특징으로 하는 SR 가속을 위한 장치.
- [청구항 11] 서버가 실행하는 SR(Super-Resolution) 가속을 위한 방법에 있어서,

비디오를 획득하는 과정;

영상 부호화기(image encoder)를 이용하여 상기 비디오로부터 복수의 비트레이트(multiple bitrate)에 해당하는 비디오 버전(video version)을 생성하는 과정;

상기 비디오 및 상기 복수의 비디오 버전을 이용하여 상응하는 복수의 SR 심층신경망(Deep Neural Network: DNN)을 트레이닝시키는 과정;

기 설정된 품질 마진 이내에서 상기 비디오 버전으로부터 상기 각 SR DNN에 대한 앵커 프레임(anchor frame)을 선별하고, 상기 앵커 프레임에 대한 정보를 포함하는 캐시 프로파일(cache profile)을 생성하는 과정; 및 상기 캐시 프로파일에 대한 정보, 및 상기 복수의 SR DNN에 대한 정보를 포함하는 매니페스트 파일(manifest file)을 생성하는 과정을 포함하는 것을 특징으로 하는 SR 가속을 위한 방법.

[청구항 12] 제11항에 있어서,
상기 매니페스트 파일은,
상기 각 SR DNN에 대한 옵션들(options) 각각이 지원하는 모바일 디바이스 리스트를 더 포함하는 것을 특징으로 하는 SR 가속을 위한 방법.

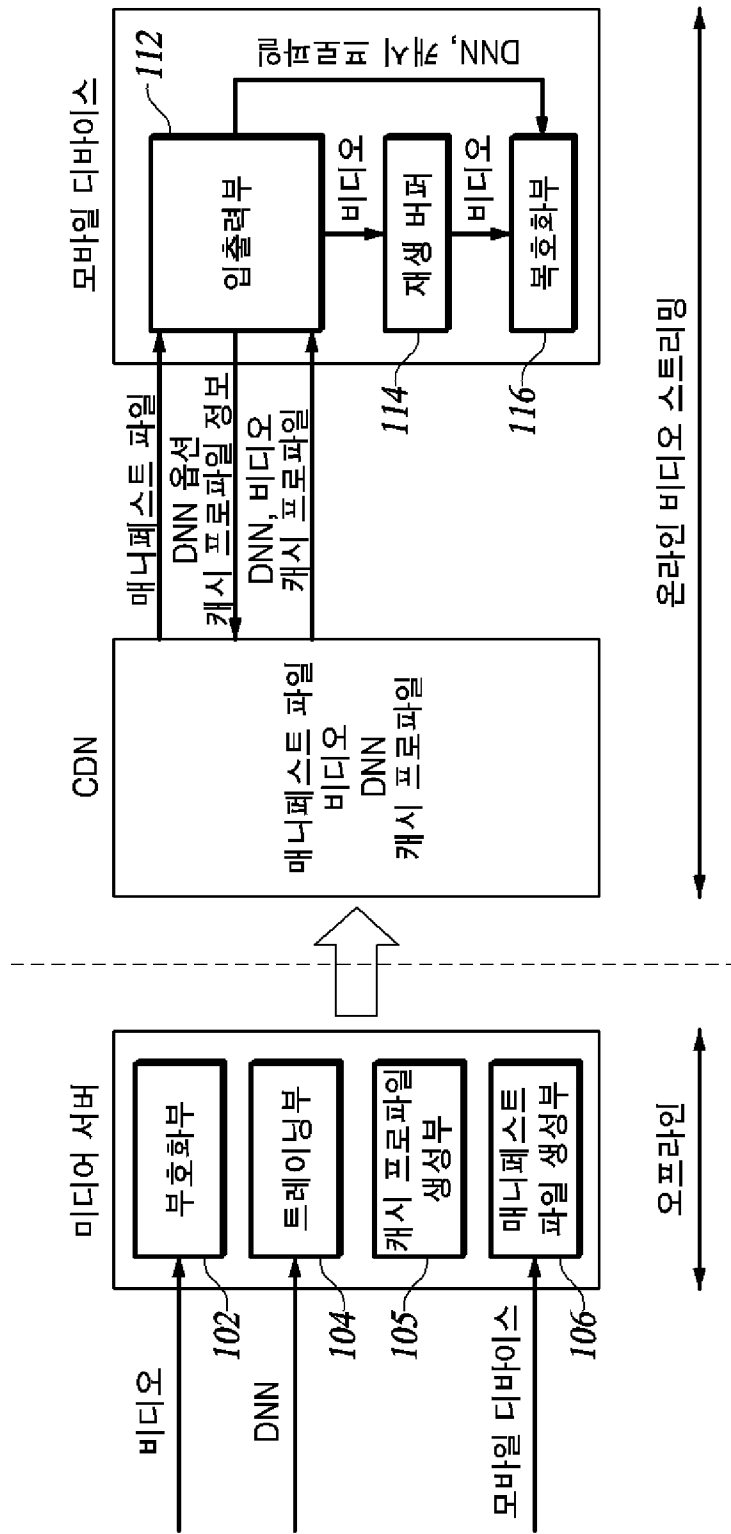
[청구항 13] 제12항에 있어서,
모바일 디바이스가 상기 비디오를 요청하는 경우, 상기 매니페스트 파일을 상기 모바일 디바이스 측으로 전송하고, 상기 매니페스트 파일로부터 선택된 SR DNN에 대한 옵션과 캐시 프로파일에 대한 정보를 상기 모바일 디바이스로부터 획득한 후, 상기 옵션에 대응하는 SR DNN, 캐시 프로파일 및 비디오 버전의 뭉치를 상기 모바일 디바이스 측으로 전송하는 것을 특징으로 하는 SR 가속을 위한 방법.

[청구항 14] 제11항에 있어서,
상기 앵커 프레임을 선별하는 과정은,
캐시 손실이 상기 품질 마진 이내로 유지되는 한도에서, 상기 앵커 프레임의 개수가 최소화되도록 선별하되,
상기 캐시 손실은,
상기 비디오 버전의 모든 프레임에 상기 SR DNN을 적용한 경우와 상기 앵커 프레임에만 상기 SR DNN을 적용한 경우 간의 품질 차이인 것을 특징으로 하는 SR 가속을 위한 방법.

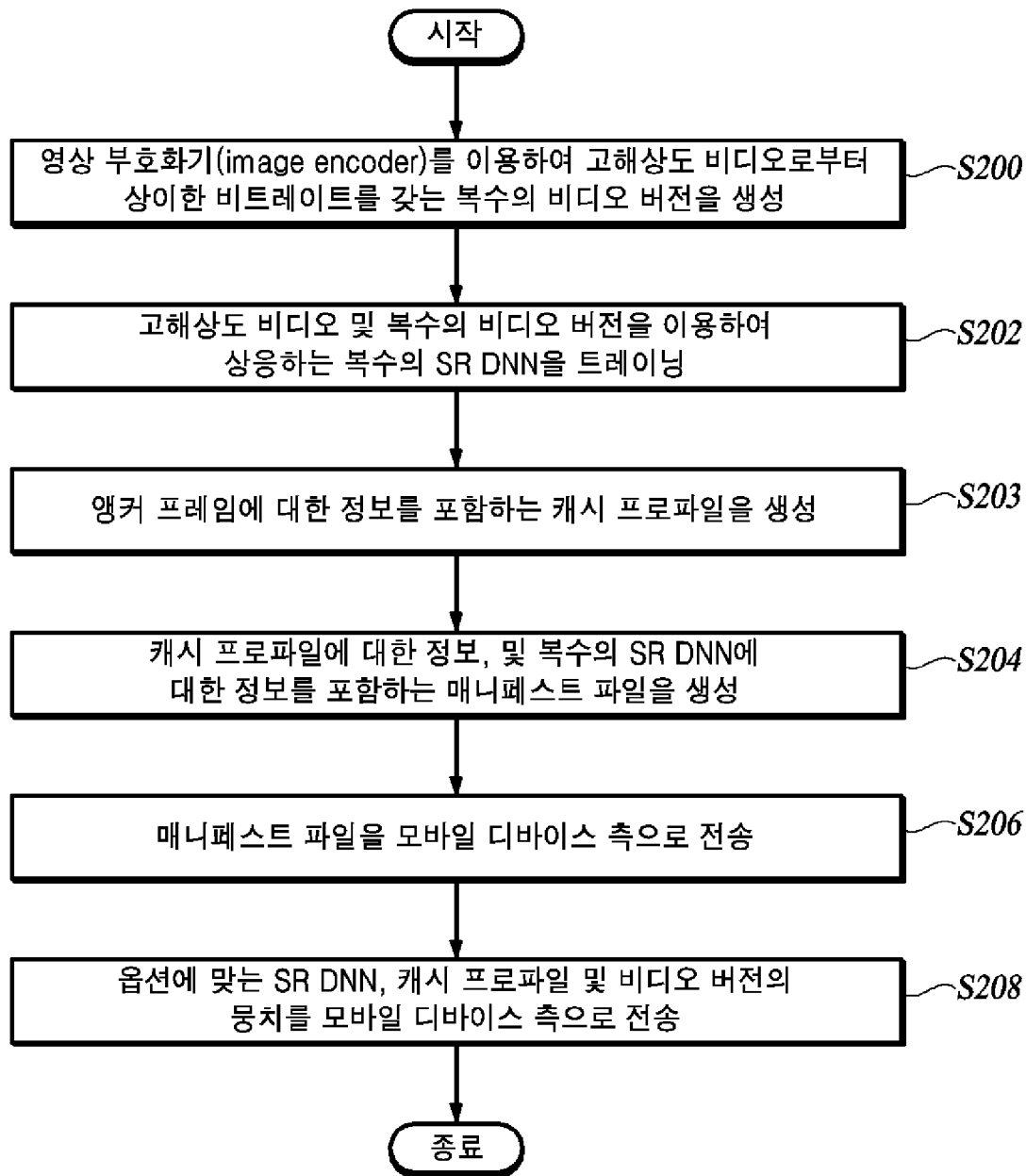
[청구항 15] 제11항에 있어서,
상기 앵커 프레임을 선별하는 과정은,
SR 집적 복호화기(SR-integrated decoder)를 이용하여 크기가 1인 앵커 프레임들의 모든 가능한 조합에 대한 품질을 측정한 결과에 근거하여, 상기 앵커 프레임의 집합에 대한 품질 이득을 산정함으로써 상기 앵커 프레임을 선별하기 위한 검색 공간의 크기를 감소시키는 것을 특징으로

- 하는 SR 가속을 위한 방법.
- [청구항 16] 제15에 있어서,
상기 SR 집적 복호화기는,
현재 프레임이 상기 앵커 프레임인 경우, 상기 현재 프레임으로부터
저해상도 프레임을 복호화한 후, 상기 SR DNN을 적용하여 상기 저해상도
프레임에 대한 고해상도 프레임을 생성 및 캐시하는 것을 특징으로 하는
SR 가속을 위한 방법.
- [청구항 17] 제15항에 있어서,
상기 SR 집적 복호화기는,
상기 현재 프레임이 상기 앵커 프레임이 아닌 경우, 프레임 의존성(frame
dependency)에 대한 정보 및 캐시된 이전 고해상도 프레임을 기반으로,
인터 예측 및 인트라 예측을 이용하여 상기 현재 프레임에 대한 고해상도
프레임을 생성 및 캐시하는 것을 특징으로 하는 SR 가속을 위한 방법.
- [청구항 18] 제11항에 있어서,
상기 캐시 프로파일은 상기 비디오 버전에 포함된 프레임이 상기 SR
DNN을 적용하기 위한 앵커 프레임인지 여부에 대한 데이터를 포함하고,
상기 캐시 프로파일에 대한 정보는 상기 서버 측에 저장되는 캐시
프로파일의 위치인 것을 특징으로 하는 SR 가속을 위한 방법.
- [청구항 19] 제1항 내지 제5항 중 어느 한 항에 따른 모바일 디바이스가 실행하는 SR
가속을 위한 방법이 포함하는 각 단계를 실행시키기 위하여 컴퓨터로
읽을 수 있는 기록매체에 저장된 컴퓨터프로그램.
- [청구항 20] 제11항 내지 제18항 중 어느 한 항에 따른 서버가 실행하는 SR 가속을
위한 방법이 포함하는 각 단계를 실행시키기 위하여 컴퓨터로 읽을 수
있는 기록매체에 저장된 컴퓨터프로그램.

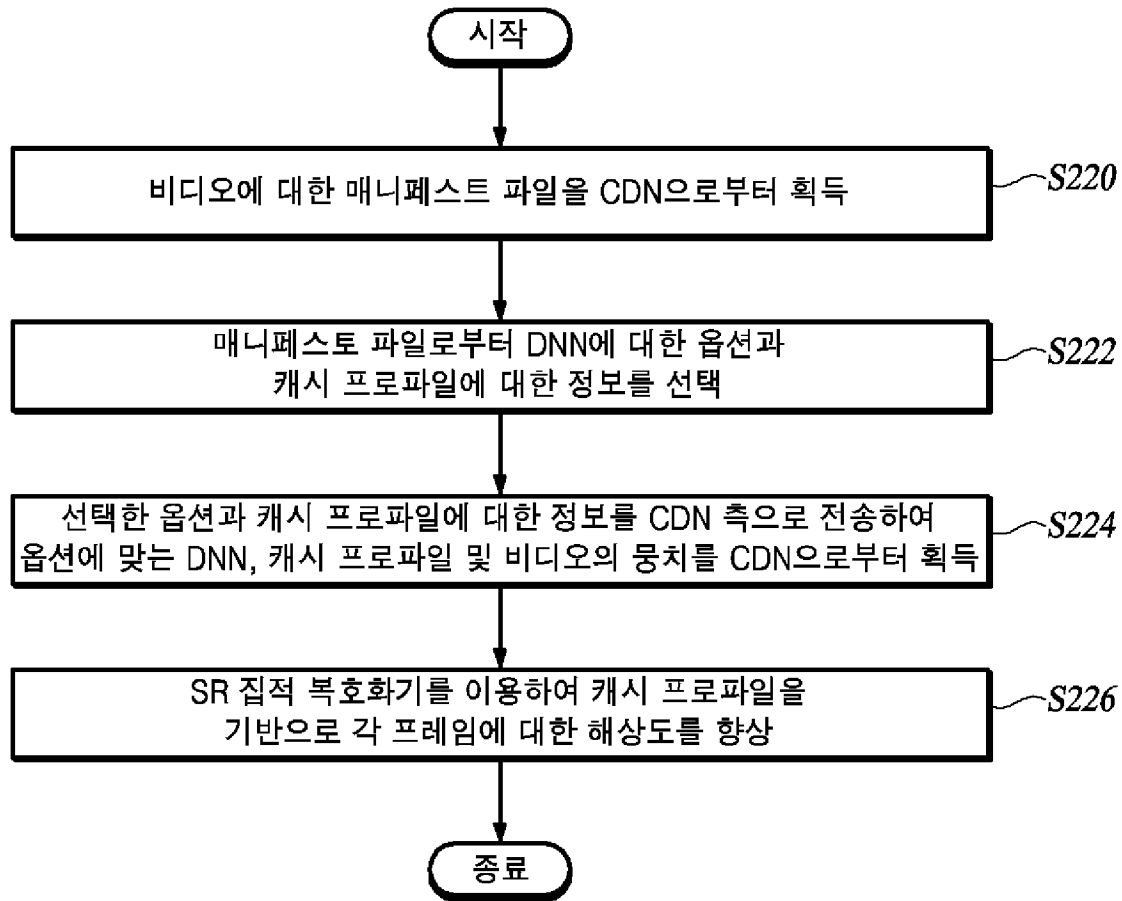
[도 1]



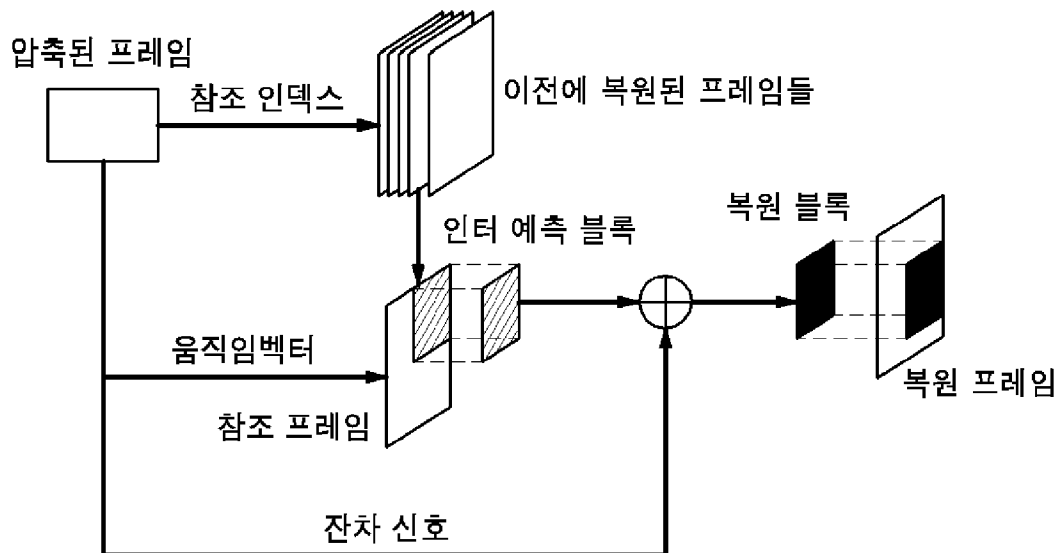
[도2]



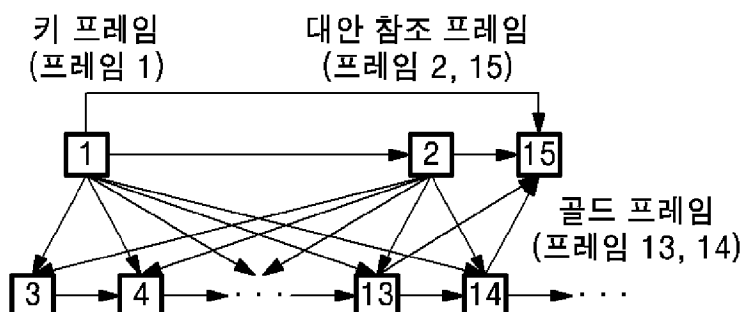
[도3]



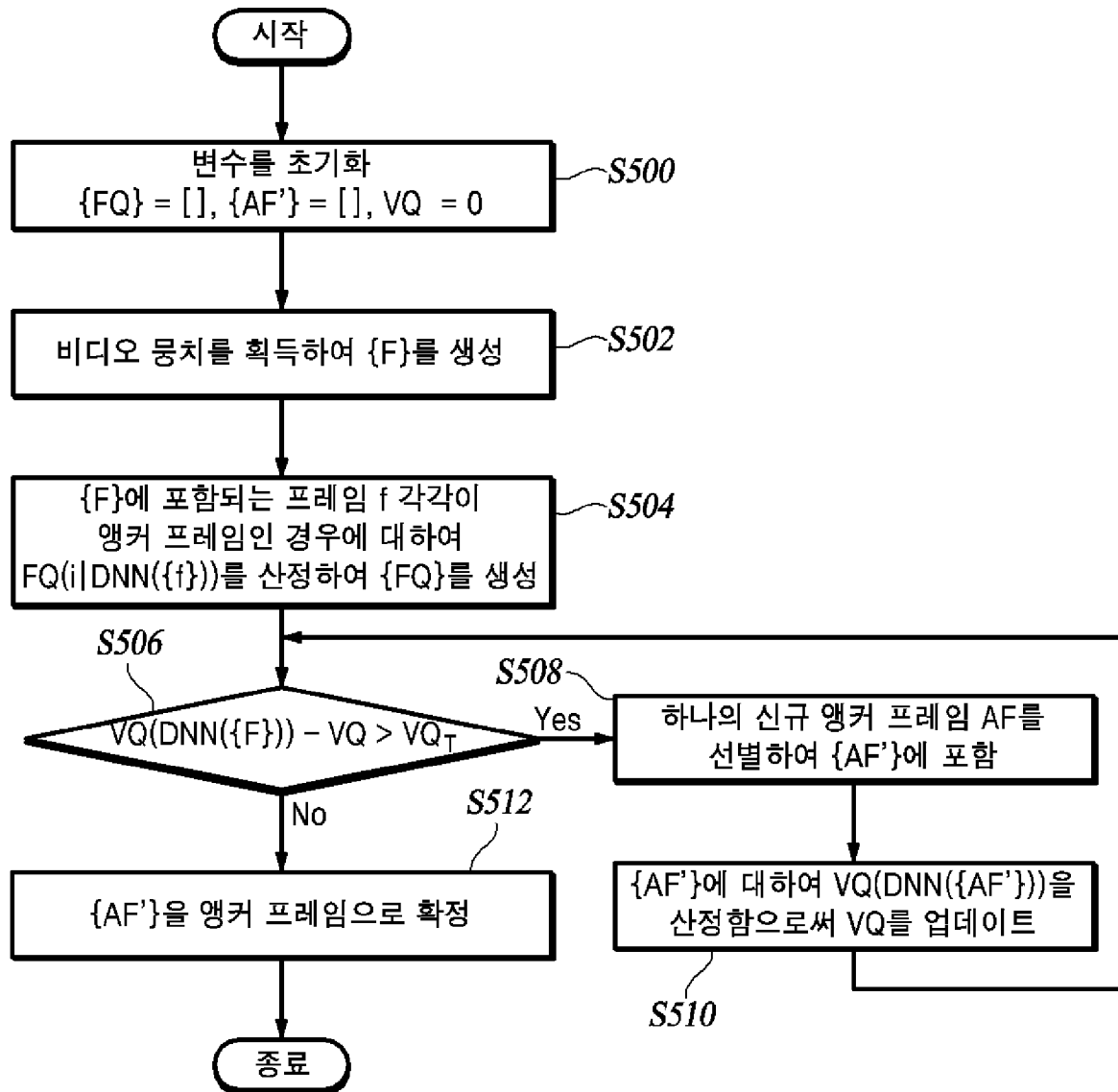
[도4]



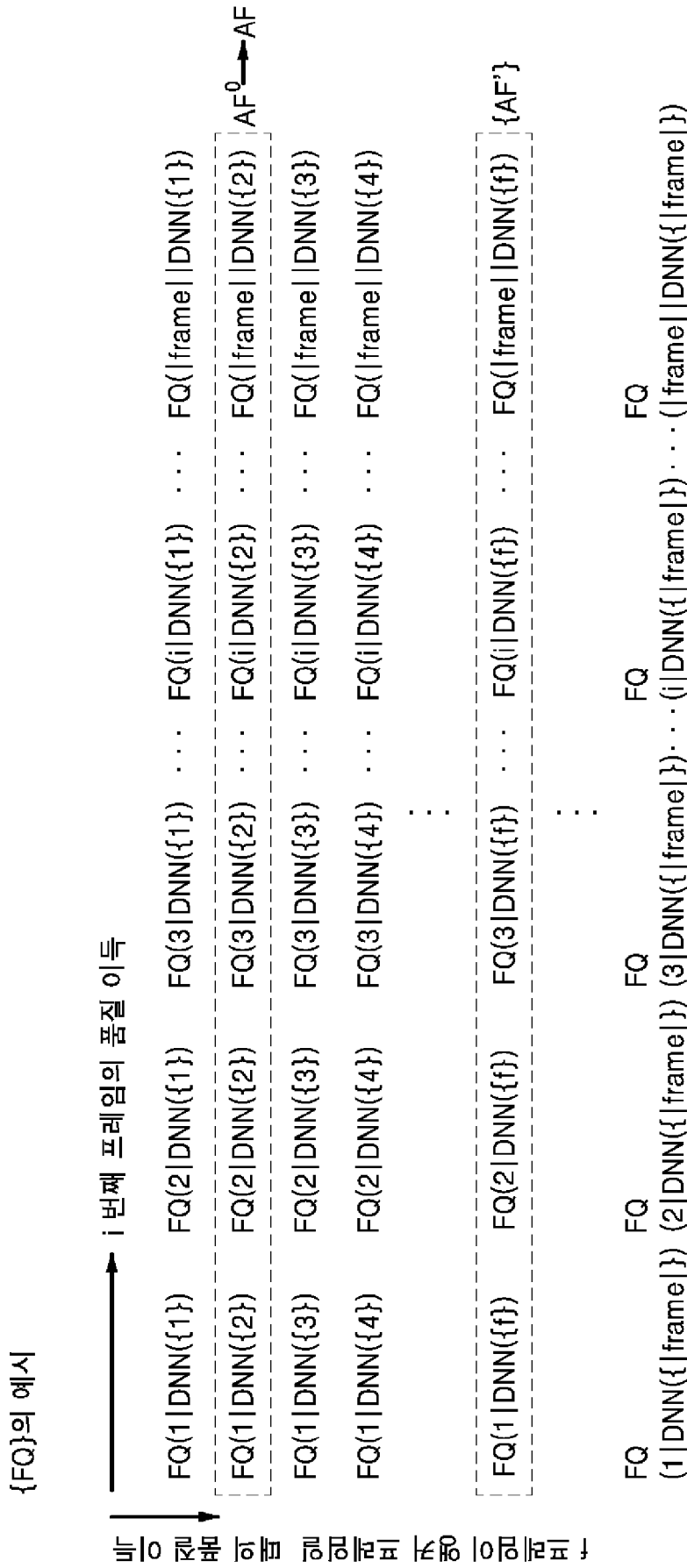
[도5]



[도6]



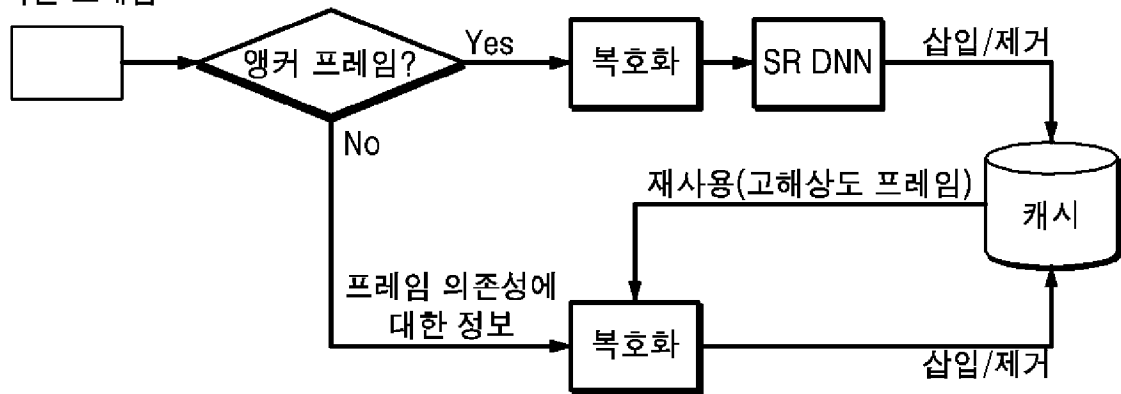
[도7]



비디오 프레임의 품질 이득

[도8]

압축된 프레임



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2020/018519

A. CLASSIFICATION OF SUBJECT MATTER H04N 21/2343(2011.01)i; H04N 19/57(2014.01)i; G06N 3/08(2006.01)i; G06T 3/40(2006.01)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 21/2343(2011.01); G06N 3/08(2006.01); G06T 7/00(2006.01); H04N 21/466(2011.01); H04N 5/21(2006.01); H04N 5/235(2006.01) 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: SR DNN(Super-Resolution Deep Neural Network), 매니페스트 파일(manifest file), 캐시 프로파일(cache profile), 비디오 뭉치(video chunk), 해상도(resolution)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	YEO, Hyunho et al. Neural Adaptive Content-aware Internet Video Delivery. 13th USENIX Symposium on Operating Systems Design and Implementation (OSDI' 18). USENIX Association. pp. 645-661, 08 October 2018. [Retrieved on 10 March 2021]. Retrieved from < https://www.usenix.org/system/files/osdi18-yeo.pdf >. See pages 645-654.	1,6,19 2-5,7-18,20
Y	WO 2020-067592 A1 (KOREA ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY) 02 April 2020 (2020-04-02) See paragraph [0033]; and figure 2.	1,6,19
A	KR 10-2017-0070715 A (SAMSUNG ELECTRONICS CO., LTD.) 22 June 2017 (2017-06-22) See paragraphs [0032]-[0038]; and figure 1.	1-20
A	JP 2015-201819 A (TOSHIBA CORP.) 12 November 2015 (2015-11-12) See paragraphs [0009]-[0055]; and figures 1-5.	1-20
☒ 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 "D" document cited by the applicant in the international application "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 19 March 2021		Date of mailing of the international search report 26 March 2021
Name and mailing address of the ISA/KR Korean Intellectual Property Office Government Complex-Daejeon Building 4, 189 Cheongsaro, Seo-gu, Daejeon 35208 Facsimile No. +82-42-481-8578		Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2020/018519**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KR 10-2020-0037015 A (KOREA ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY) 08 April 2020 (2020-04-08) See paragraphs [0036]-[0058]; and figures 6-7.	1-20

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2020/018519

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)		Publication date (day/month/year)
WO	2020-067592	A1	02 April 2020	None		
KR	10-2017-0070715	A	22 June 2017	US	10489684	B2 26 November 2019
				US	2017-0169313	A1 15 June 2017
				US	2020-0090008	A1 19 March 2020
JP	2015-201819	A	12 November 2015	JP	6396066	B2 26 September 2018
KR	10-2020-0037015	A	08 April 2020	KR	10-2129115	B1 02 July 2020

A. 발명이 속하는 기술분류(국제특허분류(IPC)) H04N 21/2343(2011.01)i; H04N 19/57(2014.01)i; G06N 3/08(2006.01)i; G06T 3/40(2006.01)i		
B. 조사된 분야 조사된 최소문헌(국제특허분류를 기재) H04N 21/2343(2011.01); G06N 3/08(2006.01); G06T 7/00(2006.01); H04N 21/466(2011.01); H04N 5/21(2006.01); H04N 5/235(2006.01) 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: SR DNN(Super-Resolution Deep Neural Network), 매니페스트 파일(manifest file), 캐시 프로파일(cache profile), 비디오 뭉치(video chunk), 해상도(resolution)		
C. 관련 문헌		
카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y A	HYUNHO YEO 등. Neural Adaptive Content-aware Internet Video Delivery. 13th USENIX Symposium on Operating Systems Design and Implementation (OSDI' 18), USENIX Association, pp. 645-661, 2018.10.08 [검색일: 2021.03.10], 출처 < https://www.usenix.org/system/files/osdi18-yeo.pdf >. 페이지 645-654	1,6,19 2-5,7-18,20
Y	WO 2020-067592 A1 (한국과학기술원) 2020.04.02 단락 [0033]; 및 도면 2	1,6,19
A	KR 10-2017-0070715 A (삼성전자주식회사) 2017.06.22 단락 [0032]-[0038]; 및 도면 1	1-20
A	JP 2015-201819 A (TOSHIBA CORP.) 2015.11.12 단락 [0009]-[0055]; 및 도면 1-5	1-20
A	KR 10-2020-0037015 A (한국과학기술원) 2020.04.08 단락 [0036]-[0058]; 및 도면 6-7	1-20
☐ 추가 문헌이 C(계속)에 기재되어 있습니다. ☒ 대응특허에 관한 별지를 참조하십시오.		
* 인용된 문헌의 특별 카테고리: “A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌 “D” 본 국제출원에서 출원인이 인용한 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 2021년03월19일(19.03.2021)		국제조사보고서 발송일 2021년03월26일(26.03.2021)
ISA/KR의 명칭 및 우편주소 대한민국 특허청 (35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사) 팩스 번호 +82-42-481-8578		심사관 양정록 전화번호 +82-42-481-5709

국 제 조 사 보 고 서
대응특허에 관한 정보

국제출원번호

PCT/KR2020/018519

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
WO 2020-067592 A1	2020/04/02	없음	
KR 10-2017-0070715 A	2017/06/22	US 10489684 B2	2019/11/26
		US 2017-0169313 A1	2017/06/15
		US 2020-0090008 A1	2020/03/19
JP 2015-201819 A	2015/11/12	JP 6396066 B2	2018/09/26
KR 10-2020-0037015 A	2020/04/08	KR 10-2129115 B1	2020/07/02