

무선 자유시점 비디오 스트리밍에서 인기도 기반 전송 기기 선택 기법

구재우*, 조영중*, 강경란^o

Transmitting Devices Selection Based on Viewpoint Popularity for Wireless Free-Viewpoint Video Streaming

Jae-Woo Koo*, Young-Jong Cho*, Kyungran Kang^o

요약

다수의 무선 기기를 통해 하나의 대상을 촬영하여, 이 영상을 통해 실제로 촬영하지 않은 위치에서 바라볼 때의 영상으로 합성하는 기술을 자유시점 비디오라고 한다. 이 기술을 무선기기를 이용한 실시간 스트리밍 서비스에 접목하면 사용자에게 새로운 서비스를 제공할 수 있으나, 모든 시점에서 촬영한 영상을 전부 전송하기에는 다수의 전송 기기가 공유하는 무선 자원의 제한과, 스트리밍 서비스가 가지는 시간적 제약으로 인한 어려움이 해결되어야 한다. 따라서, 최적의 품질을 위해서는 무선 채널의 상태와 사용자의 요청 비율을 고려하여, 서버로 전송하는 기기들을 효과적으로 선별해야 할 필요가 있다. 본 연구는 실시간 자유시점 비디오 스트리밍 서비스 환경을 가정하고, 전송 과정에서 발생할 수 있는 무선 채널 오류와 다양한 시점에 대한 사용자들의 인기도를 반영하여, 최적 품질을 보장하는 전송기기 집합을 선별하는 추단적 알고리즘을 제안한다. 성능분석을 통해 제안된 방식은 가장 인기가 높은 시점의 품질이 우선적으로 보장되는 효과를 가짐을 보였다. 또한, 이 알고리즘의 성능을 촬영 기기의 위치만을 고려한 기존 연구의 방식과 비교한다.

Key Words : Free-Viewpoint Video, Real-time Streaming, Viewpoint Popularity, Optimization

ABSTRACT

Free-viewpoint video (FVV) is a synthesization technology that generates a virtual viewpoint video using multiple videos recorded via wireless devices at heterogeneous locations. In order to introduce a new service that grafts the FVV onto the real-time streaming service using wireless devices, we need to overcome several constraints. Two main factors of those constraints are the limited wireless capacity that are shared fairly by multiple devices, and the transmission time constraint with which live streaming services have to comply. Therefore, for optimal quality of entire videos, a set of transmitting devices should be effectively selected depending on the condition of wireless channel and the required video popularity of specific viewpoint requested from users. For optimal selection, this study proposes a heuristic algorithm that takes into account the aforementioned factors from possible wireless transmission error behaviors and the requested viewpoint popularity. Through analysis and simulation, we show that with this algorithm, quality of most popular viewpoint videos is guaranteed. Furthermore, performance comparison against the existing scheme which is based only on the location of recording devices is made.

* First Author : Ajou University Department of Computer Engineering, nvmask@ajou.ac.kr 학생회원

^o Corresponding Author : Ajou University Department of Software and Computer Engineering, korykang@ajou.ac.kr, 정회원

* Ajou University Department of Software and Computer Engineering, yjcho@ajou.ac.kr, 종신회원

논문번호 : KICS2016-02-032, Received February 16, 2016; Revised May 9, 2016; Accepted May 9, 2016

I. Introduction

하나의 대상을 다각도에서 촬영하여 다양한 사용자의 요구에 맞춰 서비스하는 다시점 비디오(MVV, Multi-Viewpoint Video) 기술은 구면 디스플레이 등의 응용으로 이어졌으며^[1], 이를 기반으로 실제로 촬영하지 않은 시점(viewpoint)에서의 영상을 가상으로 생성하는 자유시점 비디오(FVV, Free-Viewpoint Video)기술 역시 개발되고 있다. 다시점 비디오와 자유시점 비디오 서비스는 전송해야 할 시점이 증가하기 때문에 전송 트래픽 역시 증가하게 되므로, 해당 분야에서는 전송 트래픽 감소에 대한 연구가 진행되고 있다.

하지만 기존의 연구들은 다수의 카메라와 서버는 유선으로 연결되어 있다고 가정된 환경에서, 서버에서 리시버로 영상이 전송될 때 발생하는 트래픽 감소만을 대상으로 한다. 이러한 환경을 현실에 반영하기에는 촬영 현장이 실내에 국한되어 있으며, 실외의 경우에도 다수의 카메라를 유선으로 연결해야 한다는 제약이 있다.

무선 카메라를 이용하여 영상을 촬영한 뒤, 무선 채널을 이용하여 서버로 전송하는 환경에서 모든 영상에 발생하는 손실을 최소화 하며 트래픽을 감소시키는 방법을 다룬 기법들이 존재하^[2,3], 전송오류 발생률과 같은 무선 채널의 특성을 고려하지 않았다. 또한, 이 연구들은 각 카메라의 전송여부 역시 그 위치만을 참고로 하기 때문에 시점 별 요청비율에 따른 유연한 결정이 불가능하다는 단점을 가지고 있다.

본 논문이 다루는 환경은 그림 1과 같다. 무선 촬영 기기를 통해 관심 대상을 다각도에서 촬영하여, 무선 채널을 통해 서버로 전송하는 취득 과정(Acquisition)을 거친다. 서버에서는 인코딩 및 디코딩 과정을 거친 뒤, 각 리시버가 요청한 시점의 영상을 전송하는 분배

과정(Dissemination)을 거친다. 취득 과정에서 각 시점의 영상에 발생하는 오류는 모두 상이하며 (heterogeneous), 다수 리시버가 요청하는 시점 역시 상이하다. 본 연구는 취득 과정에서 시스템의 성능을 보장하는 전송 기기 선택 기법을 다룬다. 이를 위해 서비스 받는 사용자 전체가 겪는 왜곡을 최소화 하는 기존의 수식에 오류 발생 확률을 반영한 모델을 제안하였으며, 실시간 스트리밍에 적용 가능한 추단적 (heuristic) 알고리즘을 제안하였다.

이후의 구성은 다음과 같다. 본문에서는 다시점 비디오 및 자유시점 비디오 관련 연구를 짚어보고, 본 연구에서 다루는 시스템 환경과 왜곡 모델을 정의한 뒤, 이 모델을 활용하여 최적의 전송 기기 집합을 찾기 위한 억지적 (brute-force) 알고리즘과 추단적 (heuristic) 알고리즘을 제안한다. 실험 결과 제안한 기법이 촬영 기기의 위치에만 기반한 기존의 연구보다 더 적은 왜곡을 발생시키는 것으로 나타났다.

II. Background and Related Works

2.1 자유시점 비디오

자유시점 비디오는 하나의 대상을 다각도에서 동시에 촬영하여, 시청자가 그 공간에 있는 것과 같이 시점을 자유롭게 변경할 수 있는 기술로^[4], 실제로 촬영하지 않은 시점의 영상은 DIBR(Depth Image Based Rendering) 을 이용해 가상으로 생성하여 제공한다^[5]. 실제로 촬영한 시점을 ‘실제 시점 (actual view)’이라고 하고, 실제 시점을 통해 생성해 내는 시점을 ‘생성 시점 (synthesized view)’이라고 한다.

동일한 해상도와 품질로 촬영한 실제 시점의 집합을 $V = \{1, 2, \dots, V\}$ 라고 하고, 연속적인 두 개의 시점 사이에 동일한 거리를 둔 가상 시점을 K 개 생성한다고 할 때, 서비스 가능한 모든 시점의 집합은 $U = \{1, 1+1/(K+1), 1+2/(K+1), \dots, (V-1)+K/(K+1)\}$

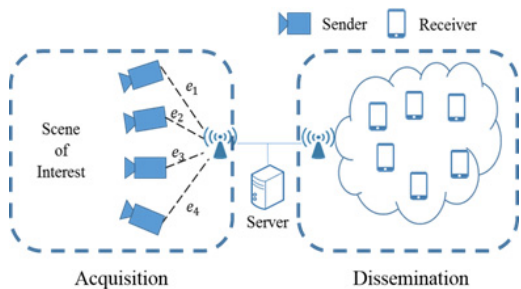


그림 1. 시스템 환경
Fig. 1. System Environment

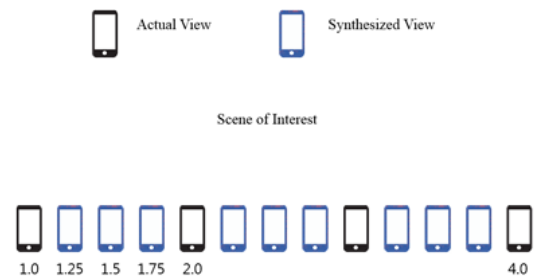


그림 2. 실제 시점과 생성 시점
Fig. 2. Actual View and Synthesized View

가 되며, $|U|=(V-1)(K+1)+1$ 가 된다. 그림 2는 $V=4, K=3$ 인 환경을 표현한 것으로, $U = \{1, 1.25, 1.5, 1.75, \dots, 3.75, 4\}$ 로, 사용자들은 총 13개의 시점을 서비스 받을 수 있다.

DIBR을 이용하여 생성 시점 u 을 생성할 때, u 의 좌측과 우측에 위치한 실제 시점 $v^l < u < v^r$, $\{v^l, v^r\} \subset V$ 를 이용하며, v^l, v^r 은 u 와 가장 가까운 실제 시점일 필요는 없다. 하지만, v^l, v^r 와 u 의 거리가 커질수록 생성시점의 왜곡이 커진다. 이는 전송하는 실제 시점이 $V' = \{1, V\}$ 일 때 왜곡이 심하지만 모든 시점의 영상을 서비스 할 수 있음을 의미하며, 모든 실제 시점을 전송할 때 서비스하는 영상의 왜곡이 최소화 됨을 뜻한다.

2.2 관련 연구 (Related Works)

자유시점 비디오에 대한 연구에 앞서, 다시점 비디오의 연구가 선행되어 왔다. 참고문헌 [6]에서는 다시점 비디오 서비스에서 분배 과정의 트래픽 감소를 위해 각 시점 간의 중복 영역 (overlapping area)에 해당하는 영상 프레임은 멀티캐스트 (multicast)를 이용하여 전송하고, 중복되지 않는 영역을 각각의 사용자에게 유니캐스트 (unicast)로 전송하는 UMSM 기법을 제안한다. 또한, 참고문헌 [7]에서는 시점 계위성을 통해 트래픽을 감소시키는 방식을 제안한다. 하지만 이러한 방식은 서버에 이미 모든 시점의 영상이 저장되어 있는 환경을 다루므로, 실시간 스트리밍 서비스에는 적용할 수 없다.

자유시점 비디오 환경을 다룬 참고문헌 [3]의 경우, 촬영 기기의 위치에 따라 중요도를 나누고, 네트워크 상태에 따른 UEP (unequal error protection) 네트워크 코딩을 적용하여 분배과정의 트래픽을 감소시켰다. 하지만 이 연구 역시 다수의 촬영 기기와 서버가 유선으로 연결되어 있다고 가정하고 있어, 촬영 환경이 실내이며 다수의 카메라를 유선으로 되어있는 환경을 가정하고 있다는 한계가 있다. 상기 연구가 분배 과정의 트래픽 감소만을 다루고 있는 것에 반해, 취득 과정의 트래픽 감소를 다룬 연구로 참고문헌 [2]가 있다. 취득과정을 다루었다는 점에서 의미가 있지만, 다른 무선 기기가 전송하는 비디오 데이터를 엿들어 (overhearing) 촬영 기기가 자신이 전송할 데이터를 줄이는 방식이므로, 기기의 배터리와 계산 능력을 고려하지 않았다는 단점이 있다. 본 논문은 참고문헌 [2]와 마찬가지로, 무선 기기를 활용한 영상 취득 과정의 트래픽을 다룬다. 참고문헌 [2]가 무선 기기가 다른 촬영 기기의 데이터를 전부 엿들어 중복 데이터를 계

산하는 과정이 필요한 것과 달리, 요청 시점의 인기도와 취득 과정의 오류 발생 확률을 이용하여 요청 받은 모든 영상에서 발생하는 손실을 최소화 하는 촬영 기기를 선택하는 기법을 제안한다.

III. Real-time Free-Viewpoint Video Streaming

그림 3은 본 논문에서 다루는 시스템 환경을 표현한 것이다. 다수의 촬영 기기가 하나의 관심 대상을 촬영하여 무선 채널을 통해 서버로 전송하고, 서버는 다수의 리시버가 요청한 영상을 전송한다. 전자를 취득 과정, 후자를 분배 과정이라고 한다.

촬영 기기와 리시버는 동일 subnet 내에 접속되어 있거나, 1-hop 내외로 통신이 가능한 근거리로 위치하며, 서비스 가능한 모든 시점의 인기도가 상이하다고 가정한다. 또한, 각각의 촬영 기기가 전송한 영상은 취득 과정에서 상이한 오류가 발생하며, 서버에서는 사전에 추측이 가능하다고 가정한다.

다수 리시버가 서버에 원하는 시점의 영상을 요청하면, 서버는 각 시점의 인기도와 오류 확률을 근거로 서버에 영상을 전송할 촬영 기기를 선택하여 전송을 요청한다. 영상 전송을 요청 받은 기기는 자신이 촬영한 영상을 서버에 전송한다. 이와 같은 취득 과정 후, 서버는 참고문헌 [8-11]의 기술을 통해 리시버가 요청한 시점이 실제 시점인 경우 단순 포워딩을, 생성 시점인 경우 해당 시점 영상을 생성한 뒤 서비스 한다. 이 과정에서 참고문헌 [12]를 통해 영상의 깊이 정보를 추정하는 과정 역시 포함된다.

실시간 스트리밍 서비스 환경에서는 촬영 기기가 영상을 촬영하고, 리시버가 영상을 전송 받을 때까지 시간지연이 존재한다. 본 논문에서는 이 지연을 Playback delay로 정의하고, 이 지연에 따라 취득 과정에 참여할 촬영 기기의 수를 결정한다. 그림 3에서

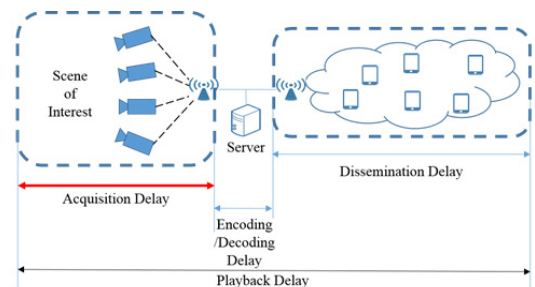


그림 3. System 및 Delay
Fig. 3. System and Delay

표현한 바와 같이 취득 과정에서 발생하는 Acquisition delay, 서버의 인코딩 및 디코딩 과정에서 발생하는 Encoding/Decoding delay, 분배 과정에서 발생하는 Dissemination delay로 이루어진다고 가정하고, Encoding/Decoding delay와 Dissemination delay는 일정하다고 가정한다. 따라서, 원활한 서비스를 위한 Playback delay를 보장할 때 조절 가능한 지연은 Acquisition delay이다. 이 제한 내에 전송 가능한 촬영 기기의 최대 수를 V_{max} 라고 정의한다.

IV. Proposed Scheme

4.1 왜곡 모델 (Distortion Model)

참고문헌 [13,14]에서는 생성 시점 u 의 왜곡이 지점을 생성하기 위해 쓰이는 실제 시점 v^l, v^r 과의 공간적 거리에만 영향을 받는다고 가정하여, 왜곡 모델 표현 수식을 다음과 같이 가정한다.

$$d_u(v^l, v^r) = \gamma e^{\alpha_u(v_u^r - v_u^l)} [e^{\beta_u(\min(u - v^l), (u - v^r))} - 1] \quad (1)$$

(1)은 v^l, v^r 을 합성하여 u 를 생성할 때의 왜곡을 추산하는 수식으로, $\gamma, \alpha_u, \beta_u$ 는 각 요소의 반영 정도를 결정하는 Multiplicative coefficient 이다. 또한, 참고문헌 [3]에서는 $V' \subseteq V$ 전송 시 서비스 가능한 모든 시점에서 발생하는 왜곡에 각 시점의 인기도를 Weight factor로 반영한 (2)의 수식을 시스템의 품질로 정의한다.

$$D(V') = \sum_{\substack{u \in U \\ u: v_u^l, v_u^r \in V'}} q_u d_u(v_u^l, v_u^r) + \sum_{\substack{u \in U \\ u: v_u^l, v_u^r \in V'}} q_u D_{max} \quad (2)$$

(2)의 q_u 는 전체 요청 중 시점 u 에 대한 요청이 차지하는 비율로, 시점 u 의 인기도를 뜻하며, D_{max} 은 시점이 생성되지 못할 경우 발생하는 최대 왜곡을 의미한다. 하지만 참고문헌 [3]에서는 취득 과정에서 발생 가능한 오류를 고려하지 않고 있다. 본 논문에서 전송과정의 오류를 반영하여 새롭게 제안하는 수식은 (3)과 같다.

$$d_u(v_u^l, v_u^r, e^l, e^r) = \begin{cases} e^u D_{min} & \text{for } u \in V \\ \frac{d_u(v_u^l, v_u^r)}{(1 - e^l)(1 - e^r)^\delta} & \text{otherwise} \end{cases} \quad (3)$$

e^l, e^r 은 u 를 생성하는데 사용하는 실제 시점 v_u^l, v_u^r 가 서버로 전송될 때 발생하는 실제 시점 영상의 오류로, 참고문헌 [15][16] 등에서 제안한 기법을 통해 사전에 추정이 가능하다. 본 논문에서는 취득 과정의 패킷 충돌 (collision)과 무선 채널의 BER (Bit Error Rate) 등의 영향으로 발생하는 전송 오류를 $e, 0 \leq e \leq 1$ 라고 정의한다. δ 는 왜곡 계산 과정에서 오류의 영향을 조절하는 계수이다. 요청 받은 시점이 실제 시점인 경우, 최소한의 왜곡으로 설정 가능한 수치인 D_{min} 에 해당 시점의 전송 과정의 오류를 반영한다.

$$D(V') = \sum_{\substack{u \in U \\ u: v_u^l, v_u^r \in V'}} q_u d_u(v_u^l, v_u^r, e^l, e^r) \quad (4)$$

서버는 (4)를 통해 서비스하는 모든 시점 V 에 발생하는 왜곡을 계산하여, 이를 최소화 하는 $V' \subset V, |V'| \leq V_{max}$ 를 선별하고, $\forall v \in V'$ 에 해당하는 기기에 영상 전송을 요청한다.

4.2 기기 선정 알고리즘 (Device Selection Algorithm)

최소한의 $D(V')$ 를 만드는 실제 시점의 집합 V'_{opt} 는 다음과 같이 표현 가능하다.

$$V'_{opt} = \operatorname{argmin}_{V' \subset V} D(V') \quad (5)$$

such that $|V'| \leq V_{max}$

이러한 집합 V'_{opt} 를 찾는 가장 단순한 방법은 Brute-force 알고리즘이다. δ 에 따라 $|V'|$ 의 증가가 $D(V')$ 의 감소를 보장하지 못하는 경우도 존재하기 때문에, $|V'| \leq V_{max}$ 인 모든 가능한 조합에 대하여 $D(V')$ 를 계산하고, 최소값을 가지게 하는 V'_{opt} 를 선택해야 한다.

Algorithm 1은 V'_{opt} 를 찾는 Brute-force 알고리즘이다. $i \leq V_{max}$ 인 경우의 조합 $\binom{V}{i}$ 를 V' 의 집합으로 설정하고, 모든 V' 에 대해 $D(V')$ 를 계산한다. $D(V')$ 의 계산에는 모든 시점 $u \in U$ 의 왜곡을 구해야 하므로, $O(VK)$ 의 연산이 필요하다. $i \leq V_{max}$ 에 해당하는 모든 조합 $\binom{V}{i}$ 에 대하여 왜곡을 계산해야 하므로, 시간복잡도는 $O(V!)$ 가 되어 Playback delay

Algorithm 1 : Brute-force Algorithm

Input: Population vector Q , Error vector E ,
 V_{max} , V , U

Output: V'_{opt}

$Distortion_{min} = \infty$
for $i = 2$ **to** V_{max}
 List of $V'_{temp} \leftarrow Combination(V, i)$
 for all $V'_{temp} \in$ List of V'_{temp}
 if $D(V'_{temp}) < Distortion_{min}$
 $V'_{opt} \leftarrow V'_{temp}$
 $Distortion_{min} \leftarrow D(V'_{temp})$
 end of if
 end of for
end of for
return V'_{opt}

내에 영상을 전송해야 하는 실시간 스트리밍 서비스에 적용하는 것은 불가능하다.

Algorithm 2는 sub-optimal V'_{opt} 를 찾는 Heuristic 알고리즘이다. 우선 요청 받은 시점 u 에서 가장 왼쪽에 위치한 시점 u^L 보다 좌측에 위치한 실제 시점 중, u^L 에 가장 가까운 v^L 과, 가장 오른쪽에 위치한 시점 u^R 보다 우측에 위치한 실제 시점 중에서 u^R 에 가장 가까운 v^R 을 선택한다. v^L 과 v^R 은 $u \in U$ 의 정의에 의해 각각 $\lfloor u^L \rfloor$ 과 $\lceil u^R \rceil$ 로 표현 가능하다. 두 개의 시점만으로도 $\forall u \in U$ 가 서비스 가능하다.

이후, $|V'| \leq V_{max}$ 범위에서 전송할 실제 시점을 하나씩 늘려가며 최소한의 $D(V')$ 를 가지는 V'_{opt} 를 찾아낸다. 이 알고리즘의 시간복잡도는 $O(V^3K)$ 로, Brute-force 알고리즘에 비교할 때 실시간 스트리밍 서비스에 적합한 수준이라고 할 수 있다.

V. Numerical Results

실험을 통해 제안한 Heuristic 알고리즘의 성능을 보이기 위해, Brute-force 알고리즘, 참고문헌 [3]에서 제안한 위치 기반의 기법과 각각 비교하였다. 실험에는 R을 이용하였으며, 성능 비교를 위한 수식으로 Achievable Minimum Distortion Rate(ADR)를 (7)과 같이 정의한다.

Algorithm 2 : Heuristic Algorithm

Input: Population vector Q , Error vector E ,
 V_{max} , V , U

Output: V'_{opt}

$v^L \leftarrow \lfloor \min_{u \in U} u \rfloor$
 $v^R \leftarrow \lceil \max_{u \in U} u \rceil$
 $V'_{opt} = v^L, v^R$
 $Distortion_{min} \leftarrow D(V'_{opt})$
 $V' = \{v | v^L \leq u \leq v^R\}$
if $V_{max} = 2$
 return V'_{opt}
for $i = 1$ **to** $V_{max} - 2$
 for $j = 1$ **to** $|V|$
 $V' \leftarrow V'_{opt} \cup \{v_j \in V\}$
 if $D(V') < Distortion_{min}$
 $v_{new} \leftarrow v_j$
 $Distortion_{min} \leftarrow D(V')$
 end of if
 end of for
 $V'_{opt} = V'_{opt} \cup \{v_{new}\}$
 $V = V - \{v_{new}\}$
end of for
return V'_{opt}

Achievable Minimum Distortion Rate

$$= \frac{Distortion \text{ when transmit } V}{Distortion_{opt} \text{ when transmit } V'} \tag{7}$$

이는, 모든 기기가 전송을 할 경우 발생하는 왜곡에 대하여, 기기를 선별하여 전송할 때 발생하는 왜곡의 비율의 역수로, 1에 가까울수록 더 좋은 성능을 보이는 것을 의미한다.

실험에 앞서, 본 실험에서 사용한 인기도 분포를 언급할 필요가 있다. 본 실험에서는 리시버들의 요청이 Uniform distribution, Binomial distribution, Binomial mixture distribution를 따르는 세 가지 유형이 있다고 가정하며, 그 분포를 나타낸 것이 그림 4이다. x축은 u 의 인덱스를, y축은 전체 요청에서 $u \in U$ 의 인기도를 의미한다. 왼쪽부터 촬영 중인 실제 시점의 수 V 가 10, 50, 100인 그래프로, 연속적인 실제 시점 사이에 생성하는 가상 시점의 수 $K=4$ 일 때 인기도 분포를 나타낸다. 앞으로의 실험에서는 언급이

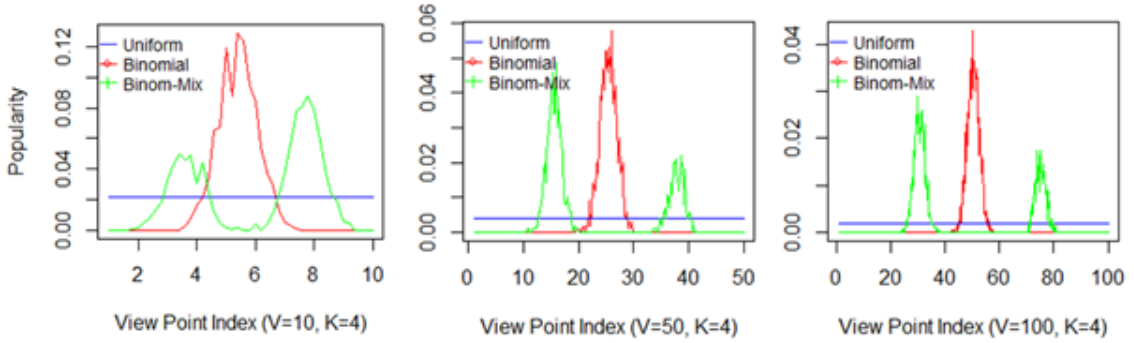


그림 4. 인기도 분포 모델
Fig. 4. Popularity Distribution Model

없는 한 $K=4$, $D_{min} = e$, $\alpha=1$, $\beta=1$, $\gamma=1$, $\delta=0.3$ 이고, 평균 오류 발생 확률이 0.3으로 설정했다.

첫 번째 실험은 $V=10$ 일 때, Brute-force 알고리즘과 Heuristic 알고리즘으로 구한 V' 를 전송할 때의 ADR을 비교한 것이다.

그림 5의 x축은 V_{max} 를, y축은 ADR을 나타낸 것으로, 2개의 촬영 기기만이 전송 가능한 경우에는 요청 모델에 무관하게 극심한 $D(V')$ 를 가진다. 하지만, 리시버의 요청이 한 곳으로 모여있는 Binomial 모델의 경우 3개의 촬영 기기가 전송을 하는 경우부터 급격히 왜곡이 감소하는 추세를 보이며, Uniform 모델이나 Binomial mixture 모델보다 적은 수의 기기가 전송을 해도 가장 낮은 $D(V')$ 를 보인다. 또한, Heuristic 알고리즘을 이용할 때에도 Brute-force 알고리즘을 사용할 때와 거의 동일한 수준의 성능을 보이는 것을 확인할 수 있다.

그림 6은 가장 인기 있는 시점의 d_{pop} 에 대한 ADR

을 나타낸 그래프로, $D(V')$ 와 비교할 때 더 적은 수의 기기가 전송하는 상황에서도 좋은 성능을 보인다. 이는 $D(V')$ 를 구하는 과정에서 Weight factor로써 시점의 인기도를 반영하는 것에 기인한다.

두 번째 실험은 V_{max} 의 변화에 따른 ADR의 변화에 대한 실험으로, 그림 7은 Heuristic 알고리즘을 통해 얻은 $D(V'_{opt})$, d_{pop} 에 대한 ADR을 촬영하고 있는 실제 시점의 수 $V=10, 50, 100$ 의 세 가지 경우에 나타낸 것이다. $D(V'_{opt})$ 보다 d_{pop} 가 더 빠르게 최적의 상태로 수렴하는 것을 알 수 있다. 이는, 취득 과정에서 전송 가능한 촬영 기기의 수가 제한적일 때, 인기도가 많은 시점의 왜곡을 최소화 하며, 인기도가 높은 시점의 품질을 우선적으로 보장한다는 주장을 뒷받침 한다.

마지막 실험은 참고문헌 [3]에서 제안한 촬영 기기 선택 기법과 우리가 제안한 Heuristic 알고리즘의 성능 비교이다. 참고문헌 [3]의 경우, 촬영 기기의 위치에 기반하여 전송하는 촬영 기기의 클래스를 나누

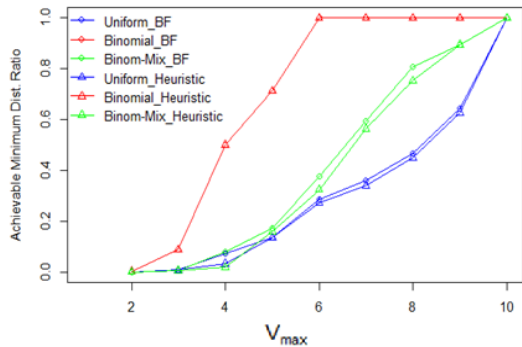


그림 5. Brute-force와 Heuristic 적용 시 $D(V')$ 비교 ($V=10$)
Fig. 5. Comparison of $D(V')$ via Brute-force and Heuristic Algorithms ($V=10$)

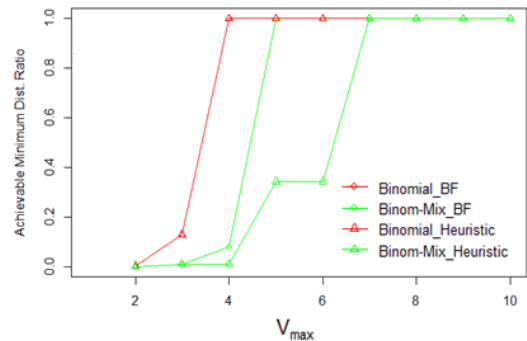


그림 6. Brute-force와 Heuristic 적용 시 d_{pop} 비교 ($V=10$)
Fig. 6. Comparison of d_{pop} via Brute-force and Heuristic Algorithms ($V=10$)

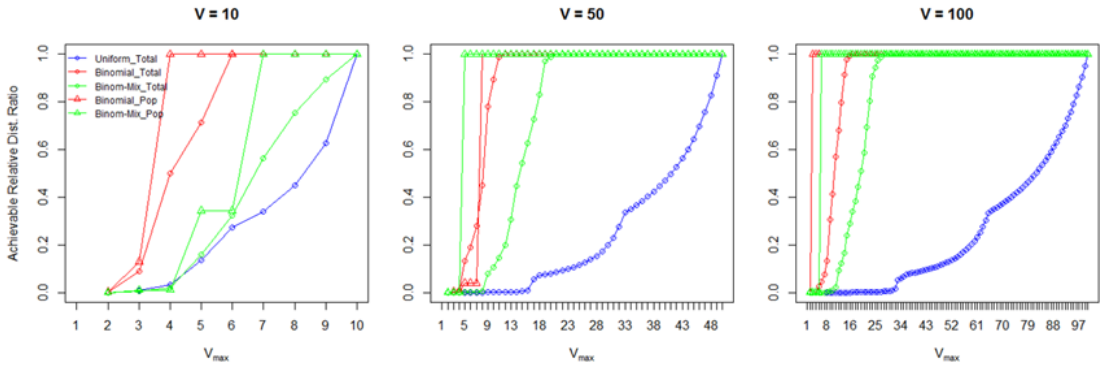


그림 7. Heuristic 알고리즘에 의한 Relative Distortion Ratio
Fig. 7. Relative Distortion Ratio via Heuristic Algorithm

고, UEP방식을 응용하여 전송 데이터에 네트워크 코딩 기법을 적용한다. 이 실험에서는 $V=17, K=4, V_{max}=3, 5, 9$ 인 세 가지 경우에 대하여 가장 인기 있는 지점을 변화에 따른 ADR의 변화를 계산하였다. 참고문헌 [3]에서는 $V_{max}=3$ 일 때, 가장 왼쪽과 가장 오른쪽, 가운데 실제 지점을 선택하여 $V' = \{1, 9, 17\}$ 를 전송한다. $V_{max}=5$ 일 때 가장 왼쪽과 가운데 위치한 실제 지점의 중앙에 위치한 지점과 가장 오른쪽과 가운데 위치한 실제 지점의 중앙에 위치한 지점을 추가하여, $V' = \{1, 5, 9, 13, 17\}$ 이 되며, $V_{max}=9$ 일 때 $V' = \{1, 3, 5, 7, 9, 11, 13, 15, 17\}$ 이 된다.

그림 8의 x축은 가장 인기 있는 지점의 인덱스를 뜻하며, y축은 해당 지점이 가장 인기 있을 때 $D(V'_{opt})$ 에 대한 ADR을 나타낸다. $V_{max}=3$ 일 때는 정 중앙에 위치한 지점이 가장 인기 있을 때 두 기법 모두 큰 왜곡이 발생하나, V_{max} 의 증가에 따라 제안 기법을 적용할 때 더 적은 왜곡이 발생하는 것을

알 수 있다. 또한, 제안한 기법은 기존 기법과 다르게 $V_{max}=9$ 일 때 모든 촬영 기기가 전송할 때와 같은 수준의 총 왜곡을 발생시키는데, 이는 위치에만 기반한 기존 기법과 다르게 인기도가 높은 지점에 따라 V'_{opt} 를 유연하게 결정하기 때문에 발생하는 차이이다. 그림 9는 x축은 가장 인기 있는 지점의 인덱스를 뜻하며, y축은 해당 지점이 가장 인기 있을 때 d_{pop} 의 ADR을 나타낸 그래프이다. $V_{max}=3$ 의 경우, 기존의 기법은 가장 인기 있는 지점이 $u_{pop} = \{1, 9, 17\}$ 일 때, $V_{max}=5$ 일 경우 $u_{pop} = \{1, 5, 9, 13, 17\}$ 일 때, $V_{max}=9$ 일 경우 $u_{pop} = \{1, 3, 5, 7, 9, 11, 13, 15, 17\}$ 로, u_{pop} 과 동일한 경우에만 최소의 왜곡을 갖는다. 반면 제안한 기법은 인기도에 기반하여 전송 기기를 선택하므로, u_{pop} 의 위치에 따라 전송 기기가 변화하며, 더 유연한 선택이 가능하기 때문에 u_{pop} 의 위치에 의한 영향이 상대적으로 미비하다.

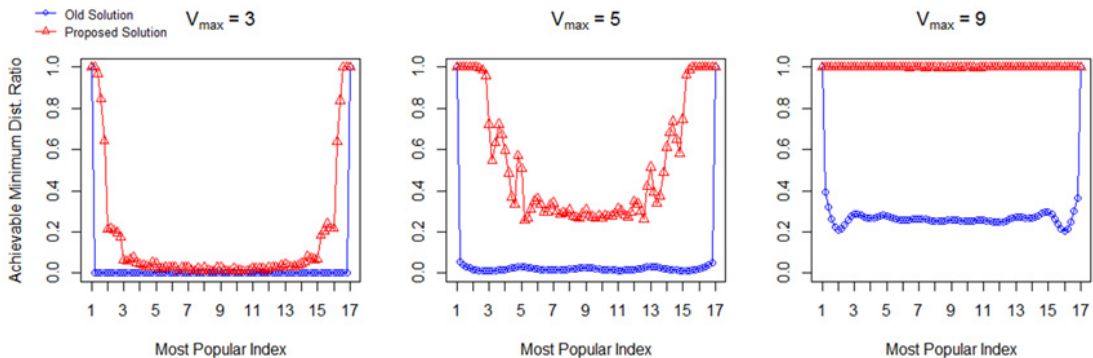


그림 8. 위치 기반 기법과 제안 기법 적용 시 $D(V')$ 비교
Fig. 8. Comparison of $D(V')$ via Location-based and Proposed Methods

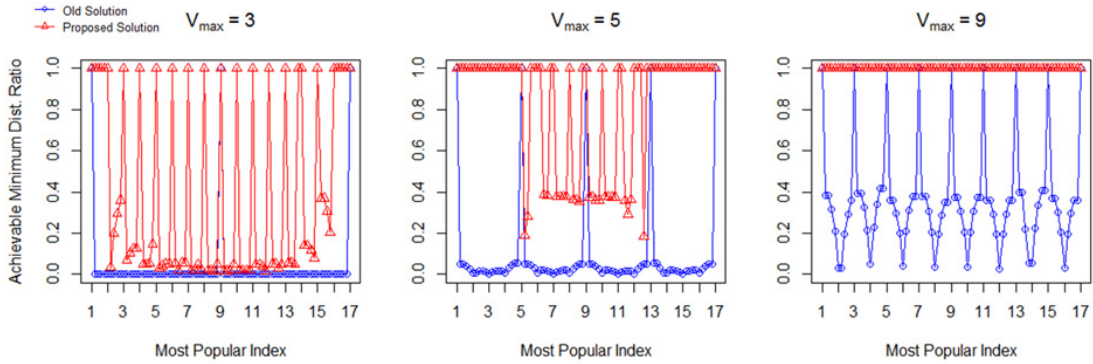


그림 9. 위치 기반 기법과 제안 기법 적용 시 d_{pop} 비교
 Fig. 9. Comparison of d_{pop} via Location-based and Proposed Methods

VI. Conclusion

본 논문에서는 실시간 자유시점 비디오 스트리밍 서비스 환경에서, 제한적 상황에서 최소한의 왜곡을 발생하는 기기를 선택하는 방식을 연구했다. 기존의 연구가 영상 처리 측면의 접근을 통해 자유시점 비디오의 트래픽 문제를 해결한 것과 달리, 본 연구에서는 전송 여부를 결정하는, 네트워크 측면의 접근을 통해 문제를 해결했다. 기기의 위치만을 고려한 기존의 방법과 다르게 각 시점의 인기도와 전송 과정에서 발생할 수 있는 오류를 감안한 새로운 방식을 제안하였으며, 실시간 스트리밍 환경에 적용 가능한 전송 기기 선정 방법을 제시했다.

자유시점 비디오는 현재 연구가 활발히 진행되고 있고, 이 기술이 실시간 스트리밍에 적용 될 경우 엔터테인먼트 분야에서 새로운 서비스는 물론, 하나의 목표에 대한 다각도의 시점이 필요한 전장 환경에서도 작전 수행 시 유용하게 사용될 수 있을 것이다. 특히, 본 논문에서 제안한 방법은 현재 스마트폰의 카메라를 이용하여 구현할 수 있으므로, 의도하지 않은 다수의 사용자들이 촬영하는 영상을 통해 새로운 정보를 창출해 낸다는 점에서 의미가 있을 것이다. 또한, 수많은 촬영 기기를 컨트롤 한다는 특징은 현재 국제 표준화 준비 작업이 진행되고 있는 초다시점 비디오^[17] 분야의 응용에도 가능 할 것이다.

하지만 본 논문에서 사용한 왜곡 수식은, 모든 촬영 기기가 영상을 전송할 때와의 상대적인 성능만을 고려하고 있다는 약점이 있다. 시점을 생성하는데 사용한 실제 시점의 왜곡이나 최대 신호 대 잡음비(Peak Signal-to-noise Ratio, PSNR)와 생성 시점의 PSNR의 관계에 대한 연구가 진행된다면, 더욱 정확한 왜곡

계산 공식을 적용할 수 있을 것이다. 또한, 우리가 아는 한, 3개 이상의 기기로 촬영한 영상 샘플이 존재하지 않는다. 차후 논문에서 다른 것과 같은 환경에 적용 가능한 영상 샘플이 등장하거나, 혹은 직접 녹화과정을 거쳐 샘플을 생성한 뒤에 시뮬레이션이 병행되어야 할 것이다.

References

- [1] J. Yang, M. Kim, J. Jeon, and G. Park, "Implement free-viewpoint image through spherical mapping of 360 degree panorama image," in *Proc. KICS Winter Conf. 2015*, pp. 1163-1164, Jung-sun, Korea, Jan. 2015.
- [2] S. Kodera, et al., "Multi-view video streaming with mobile cameras," in *Proc. GLOBECOM 2014*, pp. 1412-1417, Austin, USA, Dec. 2014.
- [3] L. Toni, et al., "Interactive free viewpoint video streaming using prioritized network coding," in *Proc. 2013 IEEE 15th Int. Workshop on Multimedia Signal Process. (MMSP)*, pp. 446-451, Pula, Italy, Oct. 2013.
- [4] M. Tanimoto, "Overview of FTV," in *Proc. 2009 IEEE Int. Conf. Multimedia and Expo*, pp. 1552-1553, New York, USA, Jul. 2009.
- [5] K. Müller, et al., "View synthesis for advanced 3D video systems," *EURASIP J. Image and Video Process*, vol. 2008, no. 4, Sept. 2008.
- [6] T. Fujihashi, et al., "UMSM: A traffic reduction method on multi-view video streaming for

multiple users,” *IEEE Trans. Multimedia*, vol. 16, no. 1, pp. 228-241, Sept. 2013.

[7] J. Kim, M. Choi, Y. Baek, D. Kim, H. Kim, and J. Yoo, “Multi-view video coding with view scalability,” *J. KICS*, vol. 32, no. 8, pp. 703-711, Aug. 2007.

[8] S. Knorr, et al., “From 2D-to stereo-to multi-view video,” in *Proc. 2007 3DTV Conf.*, pp. 1-4, Kos Island, Greece, May 2007.

[9] J. Koo, Y. Seo, H. Choi, J. Yoo, and D. Kim, “Intermediate depth image generation using disparity increment of stereo depth images,” *J. Broadcast. Eng.*, vol. 17, no. 2, pp. 363-373, Mar. 2012.

[10] J. Koo, H. Choi, Y. Seo, and D. Kim, “Fast virtual image generation using stereo matching of depth images,” in *Proc. Korean Soc. Broadcast Engineers Conf.*, pp. 126-129, Jeju Island, Korea, Jul. 2012.

[11] A. Park, D. Lee, Y. Kim, K. Lee, W. Jang, Y. Park, and J. Kim, “3D video reconstruction and rendering from multiple RGB-D cameras,” in *Proc. KICS Int. Conf. Commun.*, pp. 1075-1076, Jeju Island, Korea, Jun. 2014.

[12] K. Karsch, C. Liu, and S. Kang, “Depth extraction from video using non-parametric sampling,” in *Proc. ECCV'12*, vol. Part V, pp. 775-788, Florence, Italy, Oct. 2012.

[13] Y. Yuan, B. Hu, G. Cheung, and V. Zhao, “Optimizing peer grouping for live free viewpoint video streaming,” in *Proc. 2013 IEEE Int. Conf. Image Process. (ICIP)*, pp. 1593-1597, Melbourne, Australia, Sept. 2013.

[14] D. Ren, S. G. Chan, G. Cheung, V. Zhao, and P. Frossard, “Collaborative P2P streaming of interactive live free viewpoint video,” arXiv: 1211.4767, Nov. 2012.

[15] R. Zhang, S. L. Regunathan, and K. Rose, “Optimized video streaming over lossy networks with real-time estimation of end-to-end distortion,” in *Proc. IEEE ICME '02*, vol. 1, pp. 861-864, Lausanne, Switzerland, Aug. 2002.

[16] Z. Chen and D. Wu, “Prediction of transmission distortion for wireless video

communication: Analysis,” *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 1123-1137, Mar. 2012.

[17] Y. Ho, “A survey on international standards for super multi-view video,” *KICS Inf. & Commun. Mag.*, vol. 32, no. 3, pp. 3-10, Feb. 2015.

구재우 (Jae-Woo Koo)



2014년 2월 : 아주대학교 컴퓨터공학과 졸업
 2016년 2월 : 아주대학교 컴퓨터공학과 석사
 <관심분야> 멀티미디어, 네트워크 코딩, CDN, SDN

조영종 (Young-Jong Cho)



1985년 2월 : KAIST 석사
 1990년 2월 : KAIST 박사
 1996년 3월~현재 : 아주대학교 교수
 <관심분야> 네트워크 데이터 해석, 기계학습 활용, 트래픽 모델링, 네트워크 코딩

강경란 (Kyungran Kang)



1994년 2월 : KAIST 석사
 1999년 2월 : KAIST 박사
 2004년 3월~현재 : 아주대학교 교수
 <관심분야> 멀티캐스트, 이동 네트워크, 네트워크 코딩