

# 멀티미디어 프리젠테이션 최적 스트리밍에서 최소지연 알고리즘에 관한 성능

강희조\*

## 요약

프리젠테이션 시간축 내에서 하나이상의 렌더링 간격에서 동기화 된 멀티미디어 프리젠테이션은 대상의 집합으로 구성된다. 이러한 간격은 프리젠테이션 시간축에 관련하여 물체의 시작 시간과 끝나는 시간이 명시된다. 본 논문에서는 대역제한 된 통신 네트워크상에서 서버에서 클라이언트로의 멀티미디어 프리젠테이션의 흐름을 고려한다. 각각의 정적인 대상은 계층 부호화 되었다고 가정한다. 주어진 최대 지연에 대해서는 프리젠테이션의 전체적인 질적 측정값을 최대로 하기 위해 각 대상내의 최적의 계층 수를 찾기 위한 문제를 고려한다. 몇몇 자연적 기준에서의 최적의 방법을 결정하기 위한 효율적인 알고리즘을 고려한다. 또한 이러한 것들의 시작 시간 이후의 대상들의 점차적인 렌더링 문제를 고려한다.

## Performance on the Minimum Delay Algorithm of Optimal Streaming Multimedia Presentations

Heau-Jo Kang\*

### Abstract

A synchronized multimedia presentation consists of a collection of objects, with each object having one or more rendering intervals within the presentation timeline. These intervals specify the objects' start times and end times relative to the presentation timeline. In this paper we consider the problem of streaming a multimedia presentation from a server to a client over a bandwidth-limited communication network. We suppose that each of the static objects is layered-encoded. For a given maximum delay, we consider the problem of finding the optimal number of quality of the presentation. We devise efficient algorithms for determining an optimal policy for several criteria. We also consider the problem of gradual rendering of objects after their start times.

Keywords : Multimedia, Scheduling, Internet Applications, Minimum Delay Algorithm

## 1. 서론

프리젠테이션 시간축내에서 하나 이상의 주어진 간격에서 동기화된 멀티미디어 프리젠테이션은 대상의 집합으로 구성된다. 이러한 간격은 프리젠테이션 시간축에 대해 대상의 시작 시간과

끝나는 시간을 명시한다. 예를 들어, 그 프리젠테이션은 이미지 시퀀스가 나타나는 동안 연속적인 작용을 할 수 있는 역할을 하는 오디오 스트림으로 구성된다. 또 다른 예로는 하나 이상의 비디오클립(동시 또는 연속적인 역할을 한다.), 애니메이션, 문자 형식(HTML)등을 포함한다. SMIL(Synchronized Multimedia Integration Language) [1]은 프리젠테이션 대상의 시간적, 공간적 레이아웃을 묘사하는 프리젠테이션 표현 언어의 한 예이다. 이 논문에서는 대역 제한된 통신 네트워크상에 서버에서 클라이언트로의 멀티미디어 프리젠테이션의 흐름을 고려한다. 멀티미디어 프리젠테이션으로 사용할 수 있는 대역폭이 제한되었거나 또는 이미 알려져 있다고 가

※ 제일저자(First Author) : 강희조

접수일:2006년07월15일, 심사완료:2006년09월15일

\* 목원대학교 컴퓨터공학부

[hjkang@mokwon.ac.kr](mailto:hjkang@mokwon.ac.kr)

▣ 본 연구는 산업자원부의 지역혁신 인력양성 사업의 연구결과로 수행되었음

정한다. 예를 들어, 사용 가능한 대역폭이 클라이언트의 모델에 접속한 평균 전송률이라면 이는 종종 서버와 클라이언트 사이의 병목 대역폭이 된다.

멀티미디어 프리젠테이션으로 구성된 CM (Continuous-media) 대상과 정적인 대상을 구별하는 것은 중요하다. 오디오, 비디오, 그리고 다른 대상을 포함한 CM 대상의 비트는 재생 간격에서는 연속적인 프로세스이다. 이미지, 문자 형식, 그리고 다른 대상을 포함하는 정적 대상은 비트를 이산적 시간으로 랜더링 한다. 오늘날 대중적인 흐름 상품에서 그 서버는 순간적인 소비를 하는 연속미디어 흐름에 전송을 하고 그 클라이언트 흐름을 받아서 각 연속미디어 흐름에 랜더링 한다. 이러한 생산품은 네트워크 지터를 제거하기 위한 작은 버퍼를 만들기 위해 적은 재생지연을 소개한다. 이 모델은 연속미디어 흐름에 영향을 받는데 그 소비율과 전송률이 같다고 가정한다. 왜냐하면 프리젠테이션에 그것들의 첫 번째가 출현하기 전에 정적인 대상 전반에 나타났기 때문이다. 그 클라이언트는 최종 기한이라고 불리는 예정되어 나타나기 전에 각각의 정적인 대상을 받아야 한다. 어떠한 순간에도 그 정적인 대상을 전송에서 이용할 수 있는 대역폭은 진행중인 CM 흐름에 의해 소비되는 대역폭이다.

이 모델에서 각각의 정적인 물체는 계층부호화 되었다고 가정한다. 예를 들어, 각 정적인 대상은 점차적으로 JPEG로 인코딩 될 수 있다. 계층부호화 된 대상에 대해 랜더링 된 계층은 클라이언트에서 나타나는 모든 좀 더 낮은 계층을 요구한다. 프리젠테이션이 계층된 대상을 가질 때 서버는 시작 잠재시간을 감소하기 위한 몇몇 대상들에 대한 몇몇 계층들의 전송상의 선택사항을 가지고 있다. 그러나 계층이 낮아지게 될 때 프리젠테이션의 전체적인 질은 감소한다. 이와 같이 프리젠테이션의 질과 시작 잠재시간 사이에 중대한 상충관계가 성립된다. 주어진 시작 지연에 대해 프리젠테이션의 전체적인 질의 측정을 극대화하기 위해 각 대상 내에 최적의 계층 수를 찾기 위한 문제를 고려한다. 프리젠테이션의 전체적 질에 대한 몇 가지의 자연적인 측정 있다. 자연적인 측정의 한가지는 프리젠테이션의 전체적인 질로서 관련되는 각각의 대상

들의 질을 더하는 것이다. 그러나 각각의 질에 대한 많은 자연적인 특성 함수에 대해 전체적인 질적 기준은 프리젠테이션의 끝부분 근처에 있는 대상을 지나치게 강조하는 경향을 가지게 된다. 이러한 결과를 피하기 위해 개선된 최대최소 기준을 적용하는 새로운 기법을 소개한다. 이 개선 최대최소 기준은 여분의 대역폭을 이용할 수 있을 때 일정한 방법으로 질적 개선을 시키는 동안 모든 고정적인 대상의 질적 수치값을 동일하게 하는 것이다.

## 2. MINIMIZING THE TART-UP LATENCY

$B$ =이용 가능한 대역폭(bps)

$T$ =프리젠테이션의 길이(sec)

$N$ =고정적 대상의 수

$x_i$ =고정적 대상  $i$  를 포함한 비트수

$t_i$ =고정적 대상  $i$ 의 도착 한계 (프리젠테이션 시간선 내에 가장 먼저 나타난 대상)

$M$ = 연속 미디어 대상의 수

$r_j$ = 연속적인 대상  $j$ 의 인코딩 율 (bps)

$[a_j, b_j]$ = 연속적인 대상  $j$ 가 되기 위한 간격

$\alpha$ = 지터를 제거하기 위한 버퍼 연속적 미디어 스트림의 시간(sec) 이라 하자.

연속미디어 대상  $j$ 의 비트는 마지막 비트열이 도착될 때까지 연속율  $r_j$ 로 전송될 것이다. 시간  $t$  에 전송된 연속미디어 대상의 누적된 비트 수의 함수는 다음과 같은 형태로 주어진다.

$$b_j(t) = \begin{cases} 0 & t < a_j - \alpha \\ (t - a_j + \alpha) \cdot r_j & a_j - \alpha \leq t \leq b_j - \alpha \\ (b_j - a_j) \cdot r_j & b_j - \alpha < t \leq T \end{cases}$$

연속미디어 흐름에 의해 요구된 누적 대역폭은 다음과 같이 표현되는 개개의 흐름에 요구된 누적 대역폭의 합이다.

$$B^{CM}(t) = \sum_{j=1}^M b_j(t) \tag{1}$$

랜더링 이전  $d$ 초에 프리젠테이션 데이터를 전송하기 시작한다면 시간  $t$  에 클라이언트에서 수신된 최대 비트 수는  $B(t+d)$ 이라고 할 수 있

다. 시간  $t$  까지 고정 대상에 이용 가능한 누적 대역폭은

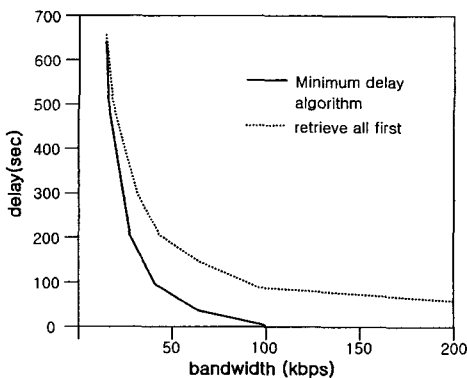
$$B_d^{static}(t) = B \cdot (t+d) - B^{CM}(t) \quad (2)$$

연속미디어 스트림이 클라이언트에 저장되지 않기 때문이다. 그들의 운반 스케줄은 프리젠테이션 시간선에 관계되어 고정된다. 그러므로 누적 연속미디어 비트의 전송에 대한 표현이  $d$  지연에 따라 좌우되지 않는다. 그러나 고정 대상은 클라이언트에 저장되고 그들의 기준선 아무 때나 전송될 수 있다. 누적 비트의 변환된 함수는  $d$ 에 영향을 받는다.

일반성을 잃지 않고 우리는  $N$  대상이 그것의 프리젠테이션 순서에 정렬된 것으로 가정할 수 있다. 부드러운 재생조건은 다음과 같다.

$$\sum_{t=1}^k \leq B_d^{static}(t_k) \quad \text{for } k=1, K, N \quad (3)$$

수식적인 예로 이 영역을 프리젠테이션에 이용 가능한 대역의 함수로 결론지었다. (그림 1)의 그래프는 시작지연을 나타낸다. 가장 윗선은 모든 고정 대상이 최초 재생을 하기 전에 첫 번째 만회할 때 생기는 지연을 보여준다. 그리고 아랫선은 최소 지연 알고리즘(특히 낮은 대역 연결에 대해)이 적용될 때 생기는 지연을 보여준다.



(그림 1) Retrieve all first 기법과 최소지연 알고리즘의 비교

### 3. 고정 대상 계층에서의 멀티미디어 프리젠테이션

이 글의 나머지 부분에 대해서는 모든 대상들이 계층 부호화 된 것으로 가정한다. 이러한 계층화된 대상들은 대상의 좀 더 많은 계층이 사용되는 것과 같은 질의 정도를 높이게 해 준다.

주어진 최대 시작 지연에 대해서, 각 대상들에서 어떤 계층을 전송 할 것인지 결정하는 문제에 있어서 프리젠테이션 질이 최적화 될 수 있도록 생각해 보자.  $d$ 를 최대 시작 지연으로 표시하자.

전과 같이,  $t_i =$  대상  $i$ 에 대한 기준을 표시하고 또한 최초 전송 시간을  $t_0$ 라 표시하자. 다시 말하여  $t_0 = -d$  라 한다.  $b_i$ 는 기준  $t_{i-1}$ 과  $t_i$  사이에 전송되는 비트 수, 다시 말하여  $b_i = B^{static}(t_i) - B^{static}(t_{i-1})$  (이후 간격  $[t_{i-1}, t_i]$ 는  $i^{th}$  간격으로,  $b_i$ 는 이 간격에 클라이언트로 전송될 수 있는 비트 수 라 표시한다.

대상  $i$ 에 대한 계층 수를 표시하기 위해서  $L_i$ 를 사용하고, 그것의  $J^{th}$  계층에서 비트 수를 나타내기 위해서  $X_{ij}$ 를 사용한다.

$y_i(j)$ 는 대상  $i$ 의 계층 1에서  $j$ 까지 보내도록 요구되는 누적 비트로 정의된다.

이를 수식으로 표현하면,

$$y_i(j) = x_{i1} + \dots + x_{ij} \quad (4)$$

기준 방침은 각각의 대상에 보내진 계층의 수를 나타내고, 이것은 클라이언트에 보내진 대상  $i$ 의 계층 수를 표시하는  $j_i$ 와  $i^{th}$ 로 구성된  $N$ -차원 벡터  $P$ 로 표시할 수 있다.

이것은,  $P = (j_i, K, j_N)$ 로 표시할 수 있으며, 여기서  $j_i$ 는 대상  $i$ 에 보내진 계층 수를 나타낸다.

$(L_i, K, L_N)$ 은 최상의 품질로 표시되는 명백한 결과이나, 이것은 주어진 대역폭과 지연의 제한 상황에서 이 데이터의 양을 보내는 것은 불가능하다. 만일 주어진 시간 전에 클라이언트에 보내진 모든 비트가 도착한다면 좋은 방법이라 할 수 있다. 명백하게 테드라인 전의 비트와 후의

비트를 함께 보낸다면 이러한 것을 부등식 N의 시스템에 따라 적당하게 표현할 수 있다.

$$\begin{aligned} y_i(j_1) &\leq b_i \\ y_i(j_i) + y_2(j_2) &\leq b_1 + b_2 \\ y_i(j_i) + K + y_N(j_N) &\leq b_1 + K + b_N \end{aligned} \quad (5)$$

지금 우리는 유연한 방책의 확정을 정의 내릴 수 있다. 만일 대상  $i$ 의 계층1에서  $j$ 까지 나타낼 수 있다면 대상  $i$ 의 품질의 값  $q_i(j)$ 을 얻을 수 있다.

그러한 품질의 값은 측정에 의해 나타난 값 (MSE) 혹은 간단한 자연 선택에 의해 주관적으로 평가된 값에 의해 정의 내릴 수 있다.

예를 들자면 자연 선택은  $q_i(j) = j/L_j$ 이며, 한 영상의 품질은 단순히 대상의 계층으로 나타내는 퍼센트를 표시하는 것이다. 또 다른 자연 방법은  $q_i(j) = y_i(j)/y_i(L_i)$ 이며, 이것은 품질의 값과 비트들의 작용으로 표현된다.

또한 대상으로 표시되는 시간의 길이의 계산과 추적에 따라 대상의 품질을 결정할 수 있다. 명백하게, 일반적인 함수  $q_i(j)$ 는 영상의 품질을 확정할 수 있게 가장 유연한 방법을 제공한다. 그리고 실질적인 값의 품질을 정의할 수 있는 모델과의 매치를 가능하게 한다.

문헌[3]의 확장된 버전은 개별 natural criteria 을 소개한다. max-min criterion 과 전체 품질의 표준, 그리고 각각의 표준에 대한 효과적인 알고리즘을 개발하여야한다. 이용 가능한 대역폭을 효율적으로 사용하는 반면에 각각의 대상의 품질을 최악의 경우로 극대화시킬 수도 있다.

전체 품질을 모든 대상에 나타내는 품질의 합으로 극대화시킬 수 있다. 또한 여섯 또는 열 개의 층을 가지는 각각의 영상과 [2]장에서 진보적인 JPEG영상을 포함한 슬라이드 쇼에 대한 알고리즘을 또한 적용한다.

#### 4. 결론

대역제한 된 통신 네트워크에서 서버에서 클라이언트로의 멀티미디어 프리젠테이션의 흐름을 고려하고, 각각의 정적인 대상은 계층 부호화 되었다고 가정한다. 주어진 최대 지연에 대해서

는 프리젠테이션의 전체적인 질적 측정값을 최대로 하기 위해 각 대상내의 최적의 계층 수를 찾기 위한 문제를 고려한다. 몇몇 자연적 기준에서의 최적의 방법을 결정하기 위한 효율적인 알고리즘을 확인 하였다.

#### 참 고 문 헌

- [1] Hoschka, P. (ed.), and the Synchronized Multimedia Working Group. Synchronized Multimedia Integration Language (SMIL) 1.0 Specification. W3C Recommendation, June 1998.
- [2] Pennebaker, W., and Mitchell, J. JPEG Still Image Data Compression Standard. Van Nostrand Reinhold Company, 1992.
- [3] Turner, D.A., and Ross, K.W. Optimal Streaming of Synchronized Multimedia Presentations with Layered-Encoded Objects, 2005.



#### 강 희 조

1994년 한국항공대학교 대학원 항공전자공학과 (공학박사)

1996년 ~1997년 : 오사카대학교 공학부 통신공학과 객원교수

1990년 ~ 2003년 : 동신대학교 전자정보 통신공학부 교수

2003년 3월 ~ 현재 : 목원대학교 컴퓨터공학부 교수

관심분야 : 멀티미디어통신, 유비쿼터스, 텔레매틱스, 무선통신, 가시광통신, 이동통신 및 위성통신, 환경전자공학, 무선광통신, 디지털콘텐츠, RFID 등