

# 인터넷 정보가전용 멀티미디어 서비스 미들웨어 기술

특 집

김 두 현\*, 마 평 수\*\*, 김 채 규\*\*\*

● 목 차 ●

1. 서 론
2. 스트리밍 서비스
3. 대화형 멀티미디어 서비스
4. 결 론

## 1. 서 론

인터넷 정보가전이란 TV를 비롯한 냉장고, 전자레인지 등의 가전기기가 디지털화되고 유무선으로 네트워크되어 상호 데이터 통신이 가능한 인터넷 단말기를 말한다. 거실에서 사용되는 디지털 TV, 주방에 사용되는 디지털 냉장고와 디지털 전자레인지, 침실과 욕실에서 사용되는 디지털 의료기기, 디지털 세탁기 등이 대표적인 인터넷 정보가전이라 할 수 있다. 최근에는 PC외에 인터넷에 접속할 수 있는 장비를 통칭하는 말로 확장되어 마이크로소프트의 Tablet PC, 노키아의 미디어스크린이 이에 해당된다[1].

한편 인터넷과 가전의 결합이라는 패러다임 속에서의 정보가전 응용 서비스와 아울러 셋탑박스나 디지털 TV를 통한 멀티미디어 서비스에 대한 요구가 증대하고 있다. 인터넷 정보가전에서의 멀티미디어 서비스는 크게 Intra-Home 멀티미디어 서비스와 Extra-Home 멀티미디어 서비스로 나눌 수 있다. 전자는 홈 씨어터와 같이 가정내의 고속 통

신망, 예를 들어 IEEE 1394 등을 통하여 네트워크화된 AV 시스템을 만들어 주는 데에 목적이 있는 반면, 후자는 집밖의 인터넷에 접속하여 집밖과 집안을 오가는 멀티미디어 서비스를 제공하는 데에 목적이 있다. 본 고에서는 Extra-Home 멀티미디어 서비스에 대하여 보다 자세히 소개하고 있다.

Extra-Home 멀티미디어 서비스는 다시 여러가지 유형으로 나누어 질 수 있겠지만 이중에서 가장 각광 받는 서비스는 대화형 서비스와 스트리밍 서비스이다. 대화형 서비스는 영상전화와 같이 집안과 집밖 사이에 양방향의 실시간성 멀티미디어 통신이 이루어지는 것이고, 스트리밍 서비스는 인터넷 방송이나 VOD 등과 같이 단방향의 멀티미디어 통신이 이루어지는 것으로 서버에서 전송되는 데이터를 클라이언트에서 실시간으로 재생함으로써 로컬로 재생하던 Intra-Home 멀티미디어 서비스 방식과 달리 저장장치를 사용하지 않아도 되는 장점이 있다.

이상과 같은 Extra-Home 멀티미디어 서비스는 인터넷과 접속되어 통신을 통하여 이루어지는 서비스이기 때문에 서비스의 기본이 되는 프로토콜이 필요하며 멀티미디어 정보를 압축/복원에 필요한 기본 함수들도 필요하다. 그러나 이러한 요소들

\* 한국전자통신연구원 멀티미디어그룹웨어팀 팀장  
 \*\* 한국전자통신연구원 책임연구원  
 \*\*\* 한국전자통신연구원 인터넷정보가전 연구부장

을 각종의 서비스에서 나름대로 정의하여 사용할 수는 없기 때문에 당연히 핵심 함수들을 모아 미들웨어로 제공함으로써 응용 프로그램을 단기간에 호환성있게 만들 수 있도록하는 기술이 필요하다. 본 고에서는 이러한 멀티미디어 서비스 미들웨어에 핵심적으로 사용되는 MPEG-4 기반의 스트리밍 기술과 대화형 서비스를 위한 SIP 프로토콜에 대하여 소개하도록 한다.

## 2. 스트리밍 서비스

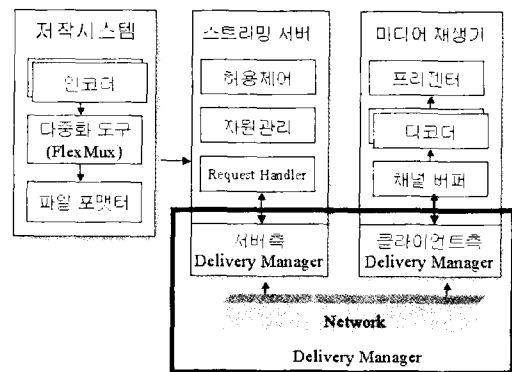
정보가전 기기를 이용한 실시간 멀티미디어 스트리밍 서비스를 제공하기 위해서는 크게 두 가지 사항을 고려해야 한다. 첫째, 정보가전 기기를 위한 통신 인프라가 모바일 환경이나 저속의 인터넷을 대상으로 하고 있기 때문에 전송률을 고려한 시스템 설계가 필요하다. 둘째, 대량의 데이터를 처리하기 위해서는 고속의 CPU 성능을 요구하게 되고 이는 기기의 가격을 결정하는 중요한 요소가 된다. 네트워크 대역폭과 정보가전 기기의 경제적인 면은 정보가전 기기의 상품화에 있어서 가장 중요한 요소이며 시장에서의 성공과 직결되는 문제이다.

현재의 네트워크 대역폭을 고려하여 스트리밍 가능한 미디어로는 음악용으로 MP3, 동영상으로는 MPEG-4가 있다. MPEG-4 재생기는 상대적으로 높은 CPU 시간을 요구하기 때문에 비슷한 전송률을 요구하고 CPU 요구량이 적은 H.263 비디오, G.723 오디오도 대안으로 많이 사용한다. 그러나 H.263은 상대적으로 화질이 떨어지고, G.723이 음성 신호가 아닌 음악 파일에 부적합하기 때문에 특정 응용에 제한된 서비스로 사용될 수 있다. 또한 MPEG-4는 기존의 MPEG 시스템과는 다르게 사용자와의 상호작용을 지원하는 것을 큰 특징으로 하고 있다.

현재 VOD 등의 스트리밍 서비스를 위해 많이 사용되는 미디어의 종류에는 국제 표준으로 정의된 MP3[2], MPEG-1[2,3,4], MPEG-4[5,6,7,8]와 상용

화되어 많이 보급된 파일 포맷인 마이크로소프트의 WMF(Windows Media File)[9], RealAudio의 RealAudio 등이 있다. 상용 파일의 경우에는 파일 포맷이 공개되지 않아 구현에 어려움이 있기 때문에 재생기와 서버에 대한 플랫폼 전체를 일괄 구매하여 시스템을 구성해야 한다. 이는 시스템 구성에 높은 비용을 요구하며, 시스템의 변경을 어렵게 하여 새로운 사용자 요구사항을 적용하기가 어렵다.

따라서 부득이하게 현재 상용화되어있는 제품이 아닌 새로운 제품을 자체 기술력으로 개발하려할 경우에는 스트리밍에 필요한 파일을 만드는 저작 시스템, 파일을 실시간으로 전송하는 스트리밍 서버, 전송되는 미디어 데이터를 디코딩하여 화면이나 스피커로 재생하고 사용자 상호작용을 수행하는 미디어 재생기, 스트리밍 서버와 미디어 재생기 간의 데이터와 제어의 처리를 담당하는 Delivery Manager가 필요하다. 저작시스템, 스트리밍 서버, Delivery Manager, 미디어 재생기의 구성은 <그림 1>과 같다.



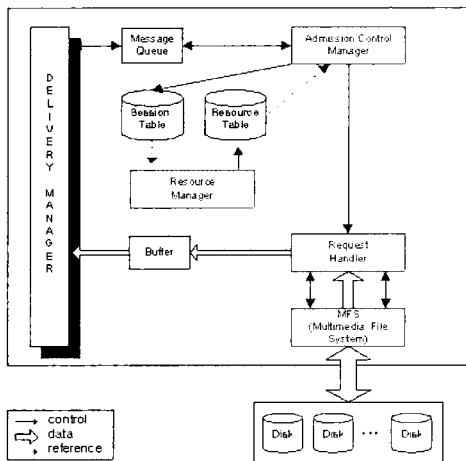
(그림 1) 스트리밍 서비스 작동 환경

### 2.1 스트리밍 서버

<그림 2>은 스트리밍 서버의 구조를 나타낸다. 스트리밍 서버는 Admission Control Manager, Request Handler, Resource Manager, 멀티미디어 파일시스템(MFS)으로 구성된다. Admission Control

Manager는 클라이언트의 요구에 대한 허용제어를 수행하며, 세션을 설정, 관리한다. 허용 제어가 성공되면 세션이 추가되고, 서비스를 요청한 파일에 대한 열기 및 네트워크 전송을 위한 버퍼를 할당한다. Request Handler는 세션에서 요구하는 데이터를 버퍼링하여 전송 시스템에 넘겨준다. Resource Manager는 세션 테이블을 주기적으로 검사하여 자원 테이블을 갱신하여 세션 정보와 자원 정보간의 일치성을 보장한다.

스트리밍 서버에서는 대량의 멀티미디어 처리를 가능케 하는 멀티미디어 파일시스템(MFS)이 요구된다. 최근의 멀티미디어 파일시스템 연구 분야에서는 한번에 읽어 들이는 비디오 데이터의 read unit 양을 크게 하거나, 파일시스템의 전송 속도를 높이기 위해 버퍼 캐쉬를 거치지 않고 디스크에서 바로 메모리로 데이터를 전송하는 Direct I/O를 지원하는 연구가 진행되고 있다. 또한 미디어 특성을 고려한 파일의 스트라이핑과 디스크 효율적인 스케줄링 기능에 대한 연구도 활발히 진행되고 있다.



(그림 2) 스트리밍 서버의 구조

## 2.2 Delivery Manager

Delivery Manager는 미디어 재생기와 스트리밍 서버간의 통신을 담당한다. Delivery Manager는 스

트리밍 서버와 클라이언트 양쪽에 하위 계층으로 존재하며, 상위 계층에 네트워크 투명성을 제공한다. Delivery Manager는 DSM-CC(Digital Storage Media-Command and Control)[10] 또는 MPEG-4의 네트워크 표준인 DMIF(Delivery Multimedia Integration Framework)[8] 등의 표준을 기반으로 많이 구현되어 진다.

DSM-CC는 저장되어진 미디어 데이터에 대한 스트리밍 기능에 대한 표준인데 여러 개의 미디어를 지원하는 MPEG-4 시스템에 적용하기는 힘들다. 따라서 MPEG-4 시스템을 구현하기 위해서는 표준에서 제시하는 DMIF를 기본 프레임워크로 사용하는 것이 유리하다. DMIF에 기반하여 인터페이스를 작성하면 MP3, MPEG-1의 경우 Channel과 TransMux 개념이 없기 때문에 MP3, MPEG-1은 각 스트림에 대해 하나의 논리적 Channel과 하나의 TransMux를 가지고 있는 것으로 취급할 수 있다. MPEG-4 스트림에 대해서는 포함하는 각 미디어 스트림에 대하여 Channel이 할당되고 전송을 위해 하나 또는 다중의 TransMux가 사용된다. 클라이언트측 Delivery Manager의 FlexDemux는 다중화된 Flexmux MPEG-4 스트림을 각각의 미디어 스트림으로 분리하여 디코딩 버퍼에 쓴다.

## 2.3 미디어 재생기

미디어 재생기는 스트리밍 서비스가 제공하는 각각의 미디어에 대한 재생을 수행한다. 미디어 재생기는 Delivery Manager를 통해 미디어 스트림을 전송 받아 버퍼에 저장한다. 버퍼에 저장된 미디어 데이터는 시스템 디코더를 거쳐 비디오, 오디오 등으로 나뉘어진다. 역다중화된 개별 비디오, 오디오 데이터는 해당 디코더에 의해 처리되고 디코딩된 결과는 프리젠테이션 시간을 검사하여 일치되면 인터페이스와 오디오 디바이스를 통해 내보낸다. 인터페이스를 통해 VCR 연산에 관한 버튼을 누르면 이벤트 관리기는 디코딩을 중단시키고 Delivery

Manager를 통해 서버에 전송 중단 및 재전송을 요청한다. 미디어 재생기는 Delivery Manager와의 통신, 네트워크 버퍼, 시스템 버퍼, 비디오/오디오 등의 버퍼 관리, 동기화 처리, VCR과 같은 사용자 이벤트 처리를 주요 기능으로 한다.

MPEG-4 재생기는 비디오, 오디오 비트열 이외에 다양한 미디어를 정의하고 각 미디어를 조합하여 장면을 구성하고 장면에 대한 사용자의 상호 작용을 처리하는 기능이 추가로 요구한다. 이를 위해 파일이 관리하는 미디어 데이터에 대한 정보를 나타내는 OD 정보[5], 각 미디어가 구성되는 장면 규칙을 기술하는 BIFS 정보[5]를 포함한다. 디코딩된 비트열 데이터를 사용하여 장면을 구성하고 재생 도중 사용자 상호작용에 따라 다양한 장면의 변환을 가능하게 한다.

## 2.4 저작 시스템

저작 시스템은 파일 전송에 사용되는 실제 파일을 만드는 과정이다. 저작시스템은 스트리밍 환경과는 오프라인으로 작업이 가능하며, 개별 미디어에 대한 인코더, 인코딩된 데이터의 다중화 도구, 그리고 스트리밍에 편리하게 파일을 생성하는 과정으로 구성된다.

MP3, MPEG-1 이나 기타 특정 업체에 종속되는 미디어 파일에 대해서는 해당 인코딩 툴이 많이 상용화되어 있다. 만약 독자적인 포맷의 미디어 파일을 생성하거나, MPEG-4와 같이 상용화된 도구가 없는 경우에는 새로운 저작 시스템을 만들어야 한다. 저작은 미디어에 대한 인코딩과 다중화를 주 기능으로 하며, 특정 포맷에 맞추는 작업 등을 부수로 요구하게 된다.

MPEG-4는 각각의 개별 미디어를 하나의 객체로 취급하여 각각을 개별 파일로 처리할 수도 있고, 네트워크 전송의 오버헤드를 줄이기 위해 미디어 데이터를 하나의 MPEG-4 파일로 통합하여 사용할 수도 있다[4]. 파일로 통합하여 사용할 때에도 개별

미디어에 대한 배치를 미디어별로 배치할 수도 있고, 동기화를 고려하여 비디오, 오디오를 다중화하여(interleave) 사용할 수도 있다. 저작 시스템은 스트리밍 서버, Delivery Manager, 미디어 재생기의 기능과 밀접하게 연관되어지며, 친숙한 사용자 인터페이스를 제공하여 손쉽게 스트리밍 서비스에 필요한 파일을 생성하여야 한다.

저작시스템은 사용자 상호작용을 지원하기 위한 BIFS 정보의 저작 기능, 미디어 데이터의 지적재산권의 보장 및 사용자의 보안 기능 등을 필요로 한다[5].

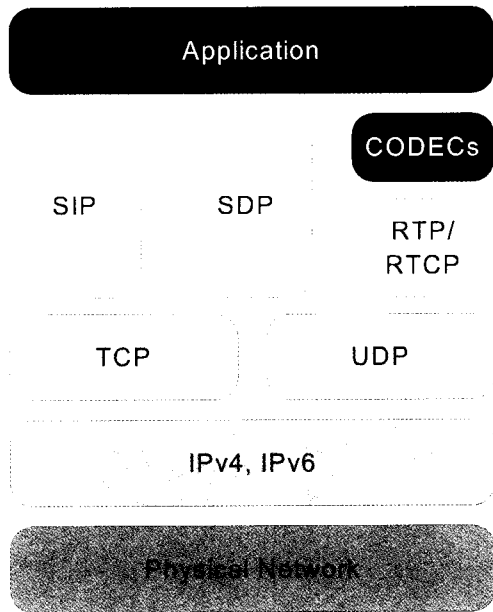
## 3. 대화형 멀티미디어 서비스

### 3.1. SIP(Session Initiation Protocol)

SIP는 네트워크 세션을 만들고 수정하고 해제하는 프로토콜이다. 현재 SIP는 멀티미디어 세션 제어와 인터넷 전화 서비스에서 사용되는 프로토콜로 많이 개발되고 있으며 그 대표적인 특성은 간결성(lightweight), 전달 망에 독립성(transport independent), 텍스트 기반(text-based)의 프로토콜이라고 할 수 있다[11]. SIP의 특성들을 살펴보면 첫째 간결성, 단지 6가지의 방법(method)으로만 구성되어 있어 간결하고 이들 방법들을 서로 조합하여 멀티미디어 세션을 완벽하게 제어할 수 있다. 둘째 전달 망에 독립성, UDP, TCP, RTP, RTCP 등 어떠한 데이터 전달 프로토콜이나 스트림 전달 프로토콜을 사용할 수 있다. 셋째 텍스트 기반, 쉽게 구문을 분석할 수 있고 확장성이 좋다.

세션을 열 때는 SIP는 signaling 프로토콜로 동작한다. 이때는 전화 signaling 프로토콜인 Q.931이나 ISUP과 유사한 서비스를 제공한다. 그러나 전화망과의 차이점은 SIP는 기본이 Internet에서 동작하고 연결 요청 시 네트워크의 자원 예약(reserve resource), 회선 설정(establish circuit)을 하지 않는다는 점이다. SIP의 기본 연결 설정은 이 메일 주소나

전화번호 등으로 구별할 수 있는 사용자나 호스트 간의 세션을 설정하는 것이다. 이때 사용자들은 접속 위치가 바뀌거나 다른 접속 장치를 사용하여도 같은 식별자를 유지할 수 있다. 식별자는 네트워크 제공자나 전화 서비스 제공자가 부여할 수 있다. <그림 3>은 SIP 구조로서 SDP, RTP등 다른 프로토콜과의 관계를 보여준다.



(그림 3) SIP 구조

### 3.1.1 SIP 구성 요소 및 연결 방식

SIP는 user agents, registrars, proxy servers, redirect servers 네 가지의 논리적 구성요소로 구분할 수 있다.

User agent는 연결을 요청하는 시스템과 최종적으로 연결 요청을 받는 시스템을 말한다. 예를 들면 인터넷 전화나 회의 소프트웨어들이 user agent 들이다. Registrar는 부여된 네트워크 도메인 내의 사용자들을 관리하는 시스템이다. Proxy server는 application-layer의 라우터로서 요청 메시지를 받아 상대방 시스템이나 다른 proxy 서버에 전달하고 요

청 메시지에 대한 응답 메시지를 보내주는 시스템이다. user agent의 request를 자신이 직접처리하고 자신이 처리하는 각 request에 대한 정보 (incoming and outgoing request)를 보관 관리하는 stateful proxy 서버일 수도 있고, outgoing request가 처리 되면 관련 정보를 관리하지 않는 stateless proxy server일 수도 있다. Redirect server는 요청 메시지를 받아 해당 메시지에 적합한 user agent 다른 위치나 해당 user agent를 찾을 수 있는 서버를 알려준다. 전달 받은 request에 대해 자신이 직접 처리하지 않고 관련 정보를 수집하여 (일반적으로 locations server를 통해 정보 수집) user agent에 응답하는 서버이다. Registrar, Proxy, Redirect 서버는 하나의 프로그램으로 구현해도 무방하다.

SIP 연결 방식은 SIP 클라이언트가 SIP 서버와 서로 통신하는 서버-클라이언트 형태를 하고 있다. User agent는 클라이언트와 서버 기능을 같이 포함하고 있으며 중간에 연결 서버들이 없이도 사용자에게이전트들간에 직접 통신을 할 수 있다. SIP 기본 통신 방식은 중간에 서버들이 존재하는 형태의 통신 방식을 취하고 있다. 이는 SIP에서 정의하고 있는 모든 서비스를 제공하기 위해서는 서버가 필요하기 때문이다.

### 3.1.2 SDP(Session Description Protocol) 개요

연결이 성립되기 위하여서는 각 user agent의 처리 능력(capability)들을 알아야 한다. 이를 위하여 SIP는 SDP를 이용한다. SDP 프로토콜은 멀티미디어 세션들을 기술하고 다양한 형식의 세션을 초기화하고 세션의 미디어 스트림에 대한 정보를 그 세션에 참석하고자 하는 사용자에게 전달하는데 사용된다. 세션 정보를 전달하는 방법으로 전자 메일과 WWW에서 사용하는 것으로, MIME 형태로 application/sdp가 사용된다. SDP는 SIP의 메시지 내용에서 사용된다.

### 3.1.3 SIP Reliability

SIP는 신뢰성을 위하여 요청 메시지를 전송할 때 다음과 같은 방식을 취한다, 하나는 INVITE를 위한 방식으로 INVITE 메시지는 최종 응답(final response)이 오기 전에 여러 가지 임시 응답 메시지를 주고 받는다 그러므로 클라이언트는 임시 응답(provisional response) 메시지가 올 때까지 타임아웃이 되면 메시지를 재전송하고 서버들은 ACK 메시지가 올 때까지 재전송한다. 그 밖의 메시지는 최종 응답 메시지가 올 때까지 최대 11번을 재전송한다.

### 3.1.4 SIP Security

SIP의 기본적인 보안 구조는 HTTP의 보안 구조를 따르고 있다. 그리고 SIP는 연결 요청 명령어 자체에 보안 설정을 해서 연결 요청을 할 수 있도록 되어있다. 또한 서버(proxy, redirect, registrar)에서 인증작업을 수행할 수 있도록 설정할 수 있다.

### 3.1.5 SIP services

SIP를 이용하여 제공할 수 있는 서비스들은 다음과 같다. 기존의 전화망에서 제공되는 서비스(Call Hold, Consultation Hold, Unattended Transfer, Attended Transfer, Unconditional Call Forwarding, Busy Call Forwarding, No Answer Call Forwarding, 3-way Call, Single-Line Extension, Find-Me, Incoming Call Screening, Outgoing Call Screening), 컨퍼런싱 서비스(멀티미디어 컨퍼런스, 멀티포인트 컨퍼런스), 이동성 지원 서비스(터미널 이동성, 사용자 이동성, 서비스 이동성), 인스턴트 메시징 서비스, 사용자 특성 선호도 제공 서비스, 인터넷 정보가전 기기 제어 서비스 등이 있다.

## 4. 결론

본 고에서는 현재 학계 및 산업계에서 관심이 고

조되고 있는 인터넷 정보가전의 Extra-Home 멀티미디어 서비스에 대한 개념 및 주요 기술에 대하여 소개 하였다. 우선 정보가전을 통한 스트리밍 서비스에 필요한 스트리밍 기술 구성으로서 스트리밍 서버, Delivery Manager, 재생기, 저장도구 등에 대하여 살펴보았다. 또한 외부 인터넷과 홈 네트워크를 연동한 대화형 멀티미디어 서비스, 즉 영상전화, 영상감시 등을 위한 방안으로서 IETF의 표준 프로토콜인 SIP(Session Initiation Protocol)에 대하여 그 개념을 설명하였다.

인터넷 정보가전 관련 기술은 홈네트워킹, 하드웨어 플랫폼, 실시간 운영체제 등 다양한 하부기술과 연계되어 있다. 그러나 이러한 기술은 인터넷 정보가전의 응용 서비스를 위하여 존재할 것이고 이러한 응용서비스의 중심에 멀티미디어 서비스가 있다고 본다. 본 고에서는 개념중심의 설명에 그쳤으나 앞으로 이와 관련한 다양한 기술이 보다 심도 있게 소개될 수 있어야 할 것이다.

## 참고문헌

- [1] "인터넷 정보가전", 한국전산원웹진-인포진 <http://www.nca.or.kr/main/ncadata/newsletter/2000/12/opinion.html>, 2000년 12월
- [2] ISO/IEC, Information Technology-Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5Mbits/s - Part 2:Audio, IS 11172-2, 1993
- [3] ISO/IEC, Information Technology-Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5Mbits/s - Part 1:System, IS 11172-1, 1993
- [4] ISO/IEC, Information Technology-Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5Mbits/s - Part 3:Video, IS 11172-3, 1993

- [5] ISO/IEC, Information Technology-Coding of Audio-Visual Objects - Part 1: Systems, IS 14496-1, July 2000
- [6] ISO/IEC, Information Technology-Coding of Audio-Visual Objects - Part 2: Visual, IS 14496-2, May 1998
- [7] ISO/IEC, Information Technology-Coding of Audio-Visual Objects - Part 3: Audio, IS 14496-3, May 1998
- [8] ISO/IEC, Information Technology-Coding of Audio-Visual Objects - Part 6: Delivery Multimedia Integration Framework, IS 144496-6, May 1998
- [9] 윤상진, "왜 임베디드 시스템인가," 마이크로소프트웨어, pp. 238-249, 2000년 12월.
- [10] ISO/IEC, Information Technology-Generic Coding of Moving Pictures and Associated Audio Information - Part 6: Extension for Digital Storage Media Command and Control, IS 13818-6, 1996
- [11] "SIP: session initiation protocol", rfc2543, Internet Engineering Task Force, March 1999.

**저자약력**



**김 두 현**

1985년 서울대학교 컴퓨터공학과(공학사)  
 1987년 한국과학기술원 전산학과(이학석사)  
 1987년-현재 한국전자통신연구원 멀티미디어그룹웨어팀 팀장, 책임연구원  
 관심분야: 인터넷 실시간 멀티미디어 서비스, 분산 멀티미디어 시스템, 멀티미디어 그룹웨어



**마 평 수**

1985년 서울대학교 식물병리학과 졸업 (학사)  
 1992년 City University of New York, USA 전산학과 (석사)  
 1995년 Wright State University, USA 전산학과 (박사)  
 1985년-1989년 시스템공학연구소 연구원  
 1989년-1990년 (주)태양고속 정보산업연구소 대리  
 1996년-현재 한국전자통신연구원 컴퓨터소프트웨어기술연구소 책임연구원  
 관심분야: 멀티미디어 저장서버, 스트리밍 기술, 멀티미디어 검색, 재생 기술 등  
 e-mail : pmah@etri.re.kr



**김 채 규**

1978년 고려대학교 수학과 (학사)  
 1977년-1998년 시스템공학연구소(연구원)  
 1983년-1990년 시스템공학연구소(선임연구원/부산사무소장, 부장)  
 1990년-1993년 Univ. of Tech.Sydney 컴퓨터과학 (석사)  
 1990년-1997년 시스템공학연구소(책임연구원)  
 1994년-1997년 Univ. of Wollongong 컴퓨터과학(호주) (박사)  
 1998년-2000년 한국전자통신연구원(책임연구원/실시간커널연구팀장)  
 2000년-현재 한국전자통신연구원(책임연구원/인터넷정보기전연구부장)  
 관심분야: 정보기전, Post-PC, 웨어러블 컴퓨팅, 유비쿼투스 컴퓨팅  
 e-mail : kyu@etri.re.kr