

H.323과 SIP를 지원하는 영상회의 시스템의 설계 및 구현

성 동 수[†]

요 약

고속통신망과 컴퓨터 기술의 발전과 함께 다양한 멀티미디어 응용 서비스가 개발되고 있다. 그 중 인터넷망에서의 영상회의 시스템은 활용도가 매우 뛰어나고 많은 관심을 받고 있으며, 이를 위하여 ITU-T H.323 및 IETF SIP 표준안이 제정되고 있다. 현재는 H.323 표준안이 많이 사용되고 있으나, 2개의 표준안이 상당기간 공존하리라 예상된다. 따라서, 두 표준안 사이에 상호운용이 필요하며, 이를 위하여 H.323-SIP 게이트웨이의 개발이 진행되고 있다. 이와는 별도로 2개의 표준안이 모두 지원되는 영상회의 시스템의 개발도 필요하다. 본 논문에서는 인터넷 환경에서 영상회의 표준안인 ITU-T H.323 및 IETF SIP를 지원하는 영상회의 터미널을 구현한다. 시스템 구현시 2개의 표준안 사이의 공통부분을 최대한 사용하도록 설계하였으며, 실험을 통하여 다른 시스템과 상호운용을 만족함을 알 수 있었다.

Design and Implementation of a Video-conference System supporting H.323 and SIP

Dong-Su Seong[†]

ABSTRACT

Various multimedia application services have been developed with techniques of high speed networks and computer. Among these services, a video-conference system over internet is useful and important, and its standardization has been shown in ITU-T H.323 and IETF SIP. H.323 standardization is currently used in many products, but two standardizations will coexist for a long time. Since the interoperability is needed between two standardizations, H.323-SIP gateway is developed for it. In addition, a video-conference system supporting two standardizations is necessary to be developed. We have implemented the video-conference terminal supporting both ITU-T H.323 and IETF SIP standardizations over internet. The system implementation has been designed using the common modules of two standardizations maximally, and the interoperability among different systems is satisfied through experiments.

키워드 : 영상회의(Video-conference), H.323, SIP, RTP/RTCP, 멀티미디어(Multimedia)

1. 서 론

고속통신망 기술과 컴퓨터 기술의 발전과 함께 다양한 정보통신 응용 서비스가 개발되고 있으며, 일반 사용자들이 이를 어디에서나 쉽게 이용할 수 있게 되었다. 이러한 변화로 텍스트 위주의 데이터 통신뿐 아니라 음성이나 동화상등 멀티미디어 서비스를 위한 다양한 통신 프로토콜이 응용 발전되고 있다. 특히, 지역적으로 떨어져 있는 사용자들이 대화형식으로 회의를 진행하기 위한 영상회의 프로토콜은 매우 중요하고 활용도가 높은 분야이다[1, 2]. 그 중 이미 기업체나 학교 및 가정 등에서 대부분 이용하고 있는 인터넷

망에서의 영상회의 시스템은 활용도가 매우 뛰어나고 관심을 받고 있으며, 이를 위하여 ITU-T(International Telecommunication Union-Telecommunication) 및 IETF(Internet Engineering Task Force)에서는 표준안을 내고 있다. H.323은 인터넷에서의 영상회의를 할 수 있도록 하기 위한 ITU-T의 표준안이며[3], SIP(Session Initiation Protocol)는 같은 목적을 위한 IETF의 표준안이다[4]. 현재는 H.323 표준안이 많이 사용되고 있으나, 상당 기간 2개의 표준안이 공존하리라 예상된다[1, 2]. 따라서, 두 표준안 사이에 영상회의가 수행되어야 하며, 이를 위하여 H.323-SIP 게이트웨이의 개발이 진행되고 있다[5, 6]. 현재 개발된 모든 영상회의 터미널은 H.323 또는 SIP 만을 지원하고 있으며, 두 표준안 사이의 상호운용을 위하여 H.323-SIP 게이트웨이를 이용해야만 한다. 이와는 별도로 2개의 표준안이 모두 지원되는

※ 이 논문은 2001학년도 경기대학교 해외파견 연구비지원에 의해 연구되었음.

† 종신회원 : 경기대학교 전자공학부 교수

논문접수 : 2002년 4월 26일, 심사완료 : 2003년 2월 10일

터미널의 개발도 필요하며, 이 터미널의 장점은 H.323-SIP 게이트웨이 없이도 이미 개발된 H.323 및 SIP 터미널과의 통화가 직접 이루어진다는 점이다. 두 표준안을 터미널의 관점에서 분석하여 보면 같은 부분이 상당히 많다. 두 표준안 이 서로 다른 점은 단지 호를 설정하는 부분과 능력교환을 수행하는 부분이며, 영상 및 음성 코덱, RTP(Real-time Transport Protocol)/RTCP(RTP Control Protocol) 등 상당한 부분이 양쪽의 표준안에서 모두 사용되고 있다. 즉, 2개의 표준안이 모두 지원되는 터미널의 개발이 필요하며, 개발에 있어 두 표준안 사이에 상당히 많은 부분이 중복되기 때문에 2개의 터미널을 각각 지니고 있는 경우보다 편리성 및 자원의 효율성 측면에서 유리함을 알 수 있었고, 이를 근거로 두 표준안이 모두 지원되는 터미널을 개발하게 되었다. 또한, 현재 표준화되고 있는 H.332 다자간 화상회의의 경우 H.323을 그대로 이용하면서 SDP(Session Description Protocol)를 추가적으로 이용함을 알 수 있다. 따라서 본 시스템을 이용할 경우 약간의 코드 추가와 함께 H.332도 지원하는 복합 영상회의 시스템을 설계할 수도 있음을 알 수 있다.

따라서, 본 논문에서는 인터넷 환경에서의 영상회의의 표준안인 ITU-T H.323 및 SIP을 소개하고, 2개의 표준안을 모두 지원하는 멀티미디어 터미널을 설계하고 구현한다. 시스템의 구현시 2개의 표준안 사이의 공통부분을 최대한 사용하도록 설계하였으며, 실험을 통하여 다른 시스템과 상호운용을 만족함을 알 수 있었다.

2. H.323 및 SIP 터미널 구성

2.1 H.323 터미널 구성

H.323은 TCP(Transmission Control Protocol)와 UDP(User Datagram Protocol)를 근간으로 하는 인터넷과 같이 QoS(Quality of Service)가 보장되지 않는 통신 환경에서 영상회의를 할 수 있도록 하기 위한 ITU-T 표준안이다[3]. H.323 터미널은 그 내부에 H.225.0[7], H.245[8] 및 각종 코덱들을 포함하고 있다. 이 중 H.225.0에는 영상회의의 호 신호(Call Signalling)를 위한 표준 규격과 인터넷 환경에서 오디오, 비디오 데이터를 실시간에 전송하기 위한 RTP 및 이를 제어하기 위한 RTCP 등을 포함하고 있다[9, 10]. 한편, H.245는 회의에 참여하는 터미널의 구성능력을 파악하여 회의 모드를 설정하고 RTP 세션을 만드는 등의 회의 설정 절차에 필요한 규격을 정하고 있다.

2.2 SIP 터미널 구성

SIP는 세션을 설정하기 위한 절차와 방법을 정하고 있는 프로토콜이다. SIP 메시지는 HTTP(Hyper Text Transfer Protocol)와 유사하게 텍스트 기반의 헤더와 바디로 구성되

어 있으며, 따라서, 상대방에게 자신에 대한 정보 즉, 자기의 SIP 주소, 포트 정보나 호에 대한 정보를 메시지 헤더에서 기술한다[4, 11]. 또한, 멀티미디어를 교환하기 위해 세션 정보를 나타내는 부분은 SDP에 따라 메시지 바디에 기술될 수 있다. SIP 메시지는 요청과 응답으로 나뉘어지며, 송신자가 요청하면 수신한 상대방은 응답을 해야 한다. SIP 터미널의 내부 구성요소로는 SIP 모듈과 자신의 세션 정보를 알려주는 SDP 모듈로 나뉘며, 그 중 SIP 모듈은 SIP 요청(Request)과 SIP 응답(Response)으로 나뉜다. 그리고, 멀티미디어 통신에 기본적인 오디오 코덱 및 비디오 코덱으로 구성되어 있다.

2.3 H.323 및 SIP에 공통적으로 이용되는 부분들

2.3.1 RTP

RTP는 실시간 데이터 전송을 위한 전송 프로토콜로써, 오디오 및 비디오 등의 실시간 데이터를 인터넷상에 전송하기에 적합한 기능을 제공해 준다. RTP는 패킷전송의 보증이나 정해진 시간 내에 정확하게 전송시키기 위한 메커니즘을 제공하는 것은 아니다. 그러나, UDP 상에서 실시간 데이터 전송 시에 발생하는 패킷손실, 패킷지연, 비순차 패킷 등은 RTP의 패킷 순차번호(sequence number)와 타임스탬프에 의하여 부분적으로 해결할 수 있다. 또한, RTCP와 함께 이용되면 손실된 패킷수 등 다양한 정보를 통계 분석하여 송신측으로 전송할 수 있으며, 송신측은 이를 이용하여 패킷 전송률을 변경하거나 인코딩 방법의 변경 등을 통하여 UDP 상의 실시간 정보전송에서 발생하는 다양한 문제들을 해결할 수 있다.

2.3.2 RTCP

RTCP는 RTP와 더불어 인터넷상에서 실시간 정보의 전송시 발생하는 패킷손실, 패킷지연, 비순차 패킷 등의 다양한 문제를 해결하기 위하여 이용되며, 이외에도 서로 다른 미디어 사이의 동기화, 간단한 세션 제어 등의 기능을 제어한다.

2.3.3 비디오 코덱(video codec)

H.323의 경우 비디오 코덱은 선택이다. H.323의 경우 비디오 통신을 제공하는 모든 터미널은 H.261 QCIF(Quarter Common Intermediate Format)에 따라 비디오를 인코딩과 디코딩을 할 수 있어야 한다. 선택적으로 다른 비디오 포맷(예 : H.261 CIF, H.263 CIF, ...)도 인코딩, 디코딩할 수 있다. 그 외 다른 비디오 코덱이나 비디오 포맷도 H.245의 중재를 통해서 사용할 수 있다. 디코더가 디코딩할 수 있는 비디오 비트율, 비디오 포맷, 비디오 알고리즘 옵션 등은 H.245를 사용해서 특성을 교환하는 중에 정의된다. 따라서, 인코더는 디코더의 디코딩 특성 범위 내에서는 자유롭게 아무거나 송신할 수 있다. SIP의 경우도 비디오 코덱은 선

택 사항이며, 만일 비디오 기능을 제공한다면 SDP에서 사용할 코덱을 결정하여야 한다. 일반적으로 사용하는 코덱으로는 QCIF나 CIF(Common Intermediate Format) 크기의 H.261, H.263, H.263+ 등이 있다. 디코더가 취할 수 있는 비디오 비트율, 영상 포맷과 알고리즘의 선택 사항들은 SDP를 사용해 성능 교환을 하는 동안 정의되면 인코더는 디코더의 능력 한도 내에서 어떠한 것도 전송할 수 있다[13, 14].

2.3.4 오디오 코덱(audio codec)

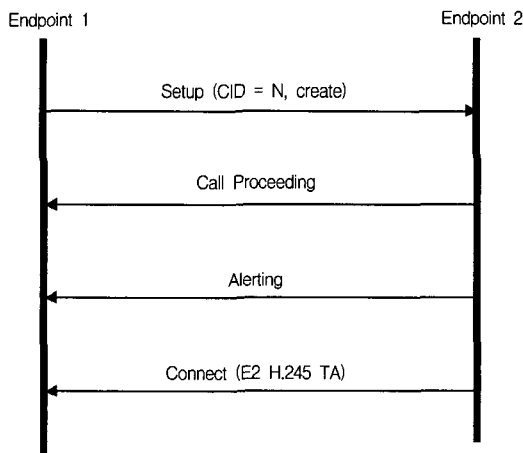
영상 회의를 하기 위해서 모든 H.323 및 SIP UA들은 오디오 코덱을 포함해야 하며, H.323의 경우 인코더가 사용하는 오디오 코덱은 H.245를 사용해서 특성을 교환하는 중에 얻을 수 있다. SIP의 경우에는 오디오 코덱들 중 엔코더로 사용되는 오디오 알고리즘은 SDP를 통해서 성능 교환되는 동안 결정된다. 일반적으로 사용하는 코덱으로는 G.711, G.722, G.728, G.729, G.723.1 등이 있다[15-17].

3. H.323 및 SIP 동작 원리

3.1 H.323 동작 원리

3.1.1 호 설정 단계

회의 생성, 회의로의 초대, 회의에의 가입 등의 프로토콜을 이용하여 회의를 시작하기 위한 신호를 주고받는다. 다자간 회의라도 동시에 여러 참여자를 호출하지 않고 우선 두 명 사이의 회의를 만든 후 다음 다른 사람들을 초대 및 가입하도록 구현되어 있다. (그림 1)은 회의 생성 과정을 나타내며, 이를 설명하면 다음과 같다.



(그림 1) H.323 회의 생성 절차

Endpoint 1(E1)이 endpoint 2(E2)에게 CID(conference ID)와 conferenceGoal = create를 포함하고 있는 Setup 메시지를 보낸다. 이때, E2는 회의에 참가할 경우 E1에게 CID = N을 포함하고 있는 Connect 메시지를 보낸다. 만일 이때, 참가하지 않을 경우 Release Complete를 보냄으로 회의를

거절한다. 사용하는 기본 메시지를 설명하면 다음과 같다. Setup 메시지는 호출하려는 터미널에게 연결 설정 요구를 지시하기 위해 호출 터미널에게 보내진다. Call Proceeding 메시지는 요청된 호 설정이 개시되었고 더 이상의 호 설정 정보는 접수하지 않겠다는 것을 지시하기 위해 피 호출 터미널에 의해 보내진다. Alerting 메시지는 피 호출 사용자 정보가 개시되었음을 지시하기 위하여 피 호출 터미널에 의하여 보내진다. Connect 메시지는 피 호출 터미널에 의해 호의 접수를 지시하기 위해 피 호출 터미널에 의해 호출 터미널로 보내진다.

3.1.2 초기 통신과 특성 교환 단계

호 설정이 이루어지면 그 결과로 H.245 프로토콜을 주고 받기 위한 H.245 세션이 설정되고 이후부터는 이를 통하여 H.245 본연의 회의 설정 절차에 관한 프로토콜이 수행되며 그 과정은 다음과 같다. 첫째, 수신된 메시지를 정확하게 분석하기 위해 사용되는 H.245의 버전 넘버를 결정하기 위해 터미널들 사이에 TerminalCapabilitySet 메시지를 교환한다. 둘째, master/slave determination 과정에 의해 Active MC(Multipoint Controller)를 결정한다. 만약 master가 MC를 갖고 있으면 이것은 Active MC가 된다. Active MC는 다른 터미널들에게 MCLocationIndication을 보낸다. MC가 지금 활성화될 수도 있고 혹은 사용자가 다자간 회의 기능을 초기화할 때 활성화될 수도 있다. 셋째, master는 터미널들에게 terminalNumberAssign 메시지를 보낸다. 터미널들은 RTP 헤더에 있는 SSRC(Synchronization Source Identifier) 필드의 하위 8비트로 할당된 16비트 수 중에서 8비트 터미널 번호를 사용해야 한다. 이 SSRC의 하위 8비트는 특별한 터미널로부터의 스트림을 구분하기 위하여 사용된다.

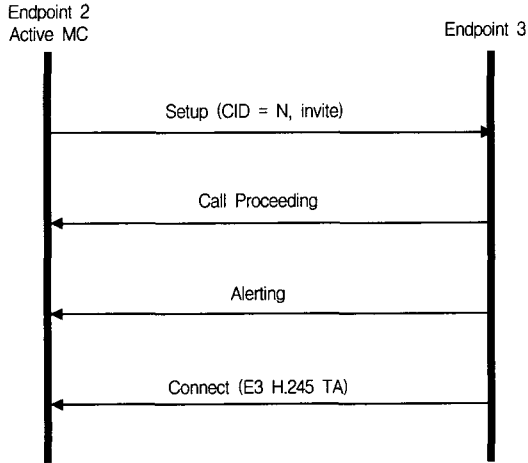
3.1.3 영상/음성 통신의 설정

H.245 프로토콜에 따라 다자간의 경우는 Active MC에 의하여 해당 회의에서 사용할 수 있는 오디오 및 비디오 채널의 특성이 정해지면 이에 맞도록 각 미디어 별로 RTP 세션을 설정한다. 일대일의 경우는 능력교환을 통하여 알게된 상대의 능력 범위 하에서 RTP 세션을 설정한다. RTP 세션은 UDP 상에 존재하는 것으로 다자간의 경우 IP Multicast를 사용할 수 있으며, QoS 제어에 필요한 RTCP 세션도 설정한다. H.323에서 실시간 영상 및 음성의 송수신 시 정확한 전송보다는 실시간 전송이 필요하며 이를 위하여 UDP를 사용한다. 반면에 파일전송이나 채팅의 경우에 해당되는 일반 데이터 전송의 경우에는 실시간 전송보다는 정확한 전송이 필요하며 이를 위하여 TCP를 이용한다.

3.1.4 호 서비스 단계

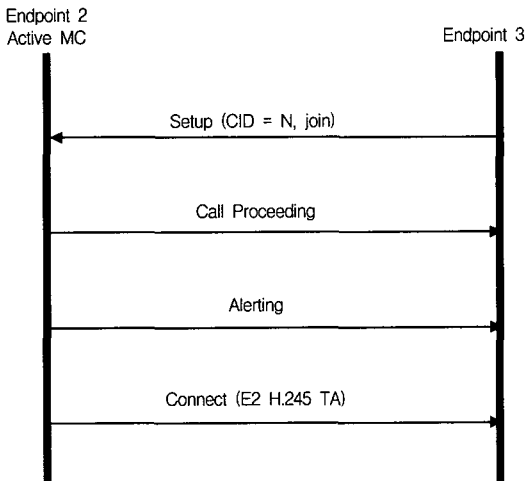
처음 두 명간의 회의가 생성되면 이를 다자간으로 확장

하기 위한 초대 및 가입 등의 서비스를 수행한다. 또한, 회의 중에 네트워크 대역폭을 증가시키기 위하여 게이트키퍼나 다른 터미널과 필요한 프로토콜을 주고받도록 한다. 초대의 경우 (그림 2)에 도시된 절차를 이용하며, 이를 설명하면 다음과 같다.



(그림 2) H.323 회의 초대 절차

먼저, E2(active MC)는 E3에게 CID = N, conferenceGoal = invite를 포함하고 있는 Setup 메시지를 보낸다. 이때, E3은 초대에 응할 경우 E2에게 Connect 메시지(with CID = N)를 보내며, 거절할 경우 Release Complete 메시지를 보낸다. 수신한 CID가 E3가 현재 참가중인 CID와 같을 경우 이미 회의가 성립되어 있으므로 Release Complete 메시지를 보내 초대를 거절해야 한다.



(그림 3) H.323 회의 가입 절차

회의에의 가입에 의하여 다자간 회의로 발전하는 과정은 (그림 3)과 같으며, 그 과정을 설명하면 다음과 같다. E3가 참가를 요청할 경우, E2에게 CID = N, conferenceGoal = join을 포함한 Setup 메시지를 보낸다. 이때 CID가 active

conference CID와 일치할 경우 E2는 E3의 참가를 허락할 경우 E2는 CID = N을 포함한 Connect 메시지를 보낸다. 만일, 참가를 허락하지 않을 경우 E2는 Release Complete 메시지를 보낸다. CID가 active conference CID와 일치하지 않을 경우 E2는 CID가 일치하지 않는다는 Release Complete 메시지를 보내야 한다.

3.1.5 호 종료

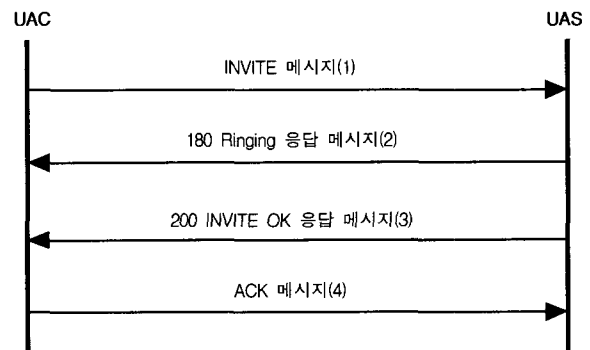
사용 중이던 오디오 및 비디오를 위한 RTP/RTCP 세션을 마감하고, H.245 세션 및 H.225 세션도 마감한다.

3.2 SIP UA의 동작원리

SIP UA(User Agent)는 UAC(User Agent Client)와 UAS(User Agent Server)로 나뉜다. 일반적으로 UAC는 요청 메시지를 만들어서 전송하는 곳을 말하며, UAS는 요청 메시지를 받아서 응답 메시지를 전송하는 곳을 말한다.

3.2.1 초기 통신과 성능 교환

상대방과 통신하기 위해서는 (그림 4)에서와 같이 INVITE 메시지를 전송해야 한다. 이때 INVITE 메시지와 200 OK 메시지, ACK 메시지를 통해 상대방과 성능 교환을 한다.



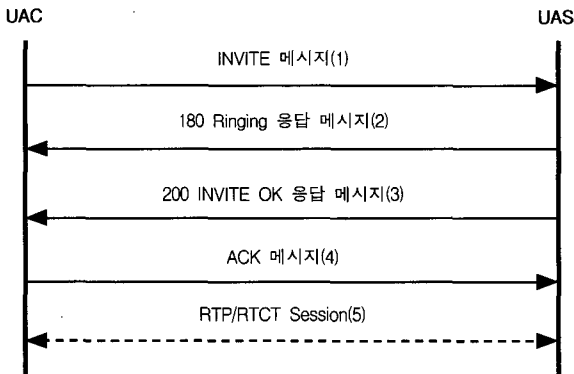
(그림 4) SIP 초기 통신 및 성능교환

성능 교환을 위한 방법은 첫 번째로는 호출자가 INVITE 시에 SDP를 보내서 상대방이 그 중 하나를 선택해서 응답 메시지를 보내서 성능을 교환한다. 두 번째로는 호출자가 INVITE시에 SDP를 보내지 않고, 상대방이 응답 메시지에 SDP를 통해 지원할 수 있는 스트림의 코덱을 호출자에게 보낸다. 호출자는 상대방에게 ACK 메소드 메시지 바디에 지원 가능한 스트림의 코덱을 SDP로 전송해서 성능을 교환한다.

3.2.2 오디오, 비디오, 데이터 채널 성립

(그림 5)와 같이 UAC가 ACK 메시지를 전송하면, 성능 교환과 함께 오디오 코덱 및 비디오 코덱이 결정된다. 뿐만 아니라, 비디오 및 오디오를 전송할 수 있는 RTP/RTCP 세션이 열린다. 이 세션을 통해서 오디오와 비디오 정보는

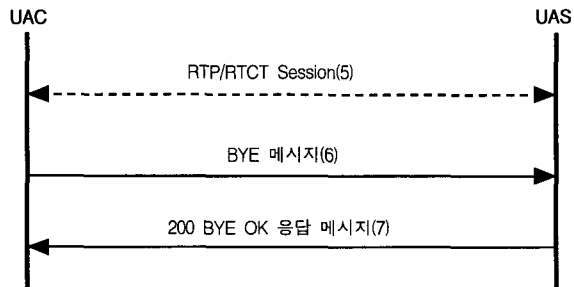
UDP 상으로 전송된다.



(그림 5) SIP RTP/RTCP 채널 성립

3.2.3 호 종료

연결된 오디오와 비디오 통신을 종료하기 위해서는 (그림 6)과 같이 BYE 메시지를 전송하면, BYE 메시지를 받은 곳에서 RTP/RTCP 세션을 종료하고, 200 BYE OK 메시지를 전송한다. 200 BYE OK 응답 메시지를 수신 받으면, 자신의 RTP/RTCP 세션을 종료한다.



(그림 6) SIP 호 종료

4. H.323-SIP 영상회의 시스템의 설계 및 구현

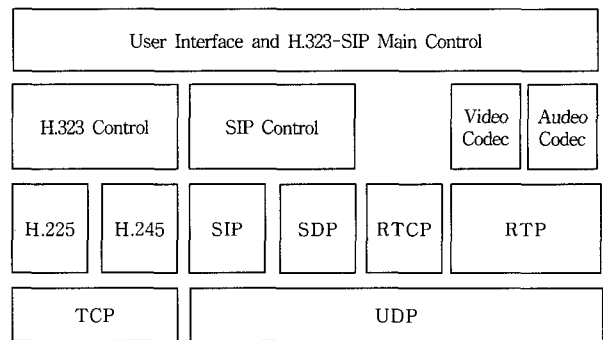
4.1 영상회의 시스템 구성

상대의 터미널이 H.323을 지원하는 터미널인 경우 H.323 제어모듈 중 H.225를 통하여 호 설정이 이루어지며, 능력교환은 H.245를 통하여 이루어진다. 그 뒤, 오디오 및 비디오 코덱을 통하여 압축된 정보를 RTP를 통하여 전송한다. SIP의 경우, SIP 제어 모듈내 SIP를 통하여 호가 설정되며 능력교환은 SDP를 통하여 이루어진다. 그 뒤, 영상 및 음성코덱 RTP, RTCP는 H.323에서 이용하는 동일한 모듈을 이용한다. 본 논문에서 구현한 시스템은 공통부분을 최대한 이용하도록 설계되어, 각각을 따로 설계하는 경우보다 복잡도가 적어짐을 알 수 있었다. 또한, 현재 표준화되고 있는 H.332 다자간 화상회의의 경우 H.323을 그대로 이용하면서 SDP를 추가적으로 이용함을 알 수 있다. 따라서, 약간의 코드 추가와 함께 H.332도 지원하는 복합 영상회의 시스템

을 설계할 수도 있음을 알 수 있다.

4.1.1 H.323 제어 모듈

H.323-SIP 터미널의 전체 구성도는 (그림 7)과 같으며, 그 중 H.323 제어 모듈은 그 내부에 H.225.0, H.245를 포함하고 있다. 이 중 H.225.0은 영상회의 호 신호를 담당하고 있으며, H.245는 회의에 참여하는 터미널의 구성 능력을 파악하여 회의 모드를 설정하고 RTP 세션을 만드는 등의 회의 설정 절차에 필요한 일을 담당하고 있다.



(그림 7) H.323-SIP 터미널의 전체 구성도

4.1.2 SIP 제어 모듈

SIP 모듈은 SIP 메시지의 생성 및 분석을 수행하는 모듈이며, SDP 모듈은 SDP 메시지의 생성 및 분석을 수행하는 모듈이다. 이 두 모듈은 SIP 제어 모듈에 의해 SIP 모듈은 SIP 메시지 헤더로, SDP 모듈은 SIP 메시지 바디로 생성되고 분석된다. SIP 제어 모듈, SIP 모듈, SDP 모듈의 사용으로 SIP 포트인 5,060 포트에 호 연결이 이루어지면, 호에서 협상된 포트와 음성 및 영상 코덱으로 세션을 생성한다.

4.1.3 RTP/RTCP 모듈

RTP/RTCP 모듈은 크게 RTP 부분과 RTCP 부분으로 나뉘어진다. RTP 부분은 코덱 부분과 미디어 데이터를 주고 받고, 또한 RTCP 부분과 제어 관련 데이터를 주고받는다. RTCP 부분은 단지 RTP 부분과의 관계만을 갖는다. RTP 모듈은 RTP 송신자/RTP 수신자, RTCP 송신자/RTCP 수신자로 구성된다. 외부 노드와의 입출력은 RTP 포트와 RTCP 포트를 통하여 이루어진다. 코덱으로부터 전달받은 정보로부터 RTP 송신자는 RTP 패킷을 만들고, 이를 RTP 포트를 통하여 상대 노드에게 전달한다. 상대 노드로부터 RTP 포트를 통하여 전달받은 RTP 패킷은 RTP 수신자로 전달되며 RTP 수신자는 전달된 정보로부터 RTP 관련 정보를 제거한 후 나머지 정보를 코덱으로 전달한다. RTCP는 RTP 송신자로부터 상대 노드에게 전달하는 패킷 및 상대 노드로부터 전달되는 RTP/RTCP 패킷 정보를 분석하고, 이 정보들을 이용하여 RTCP 패킷을 만들어 RTCP 포트를 통하여 상대 노드에 전송한다. 전달받은 RTCP 정보는 여러

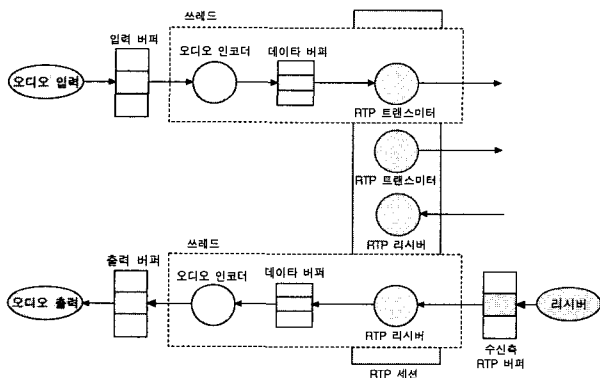
용도로 이용될 수 있으며, 주로 이를 상위 프로그램이 해석하여 RTP의 패킷량 조절이나 전송시간의 조절 등의 기능을 실행한다. RTP와 RTCP는 상호 연결되어 있어서 서로의 정보를 쉽게 이용할 수 있다.

4.1.4 코덱 모듈 구성도

코덱 모듈은 크게 인코드 모듈과 디코드 모듈로 이루어진다. 인코드 모듈은 캡처 모듈을 통해 저장된 실시간 미디어 데이터를 압축하는 모듈이다. 디코드 모듈은 압축된 미디어 데이터를 복원하는 모듈이다. 즉, 인코드 모듈을 통해 압축된 미디어 데이터는 RTP 모듈을 통해 외부로 보내지게 되고, RTP 모듈을 통해 받은 압축된 미디어 데이터는 디코드 모듈을 통해 복원된 미디어 데이터를 만든다. 현재 이용되는 오디오 코덱은 G.711, G.728, G.723.1이며 비디오 코덱은 H.261, H.263이다.

4.1.5 RTP 세션과 코덱 모듈과의 관계 구성도

(그림 8)은 RTP 세션과 오디오 코덱과의 관계 및 정보 흐름의 구성도를 나타낸 것으로서, 비디오 코덱과의 관계도 이와 동일하게 구성되어 있다.



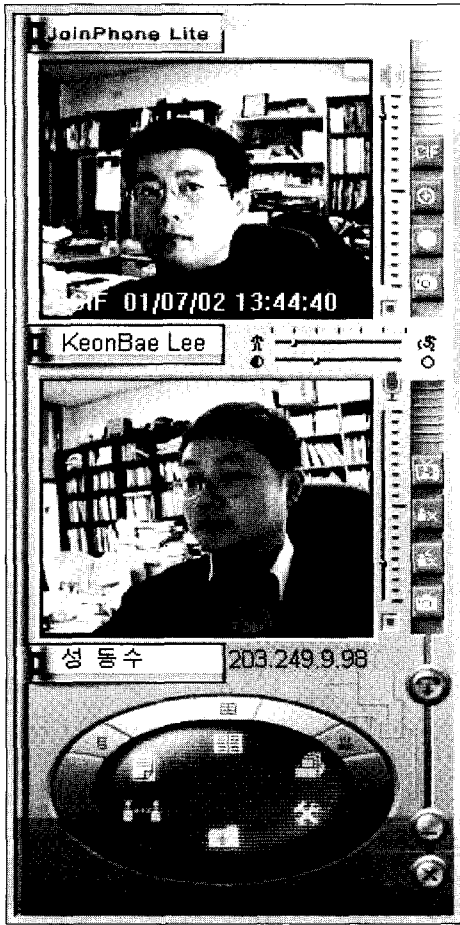
(그림 8) RTP 세션과 오디오 코덱과의 관계 구성도

RTP 세션은 크게는 RTP 송신자/RTP 수신자와 RTCP 송신자/RTCP 수신자로 구성되어 있으며, 이를 이용하는 쓰레드는 오디오 압축부와 RTP 송신자로 이루어진 전송 쓰레드와 오디오 복원부와 RTP 수신자로 이루어진 수신 쓰레드가 있다. RTCP 송신자와 수신자는 2개의 쓰레드 모두에 의해 호출된다. 정보의 흐름에 따라서 이들 사이의 관계를 살펴보면, 먼저 오디오 입력 디바이스로부터 캡처된 오디오 데이터는 입력 버퍼에 쌓이게 되고, 이렇게 쌓인 오디오 데이터를 송신 쓰레드에서 가져와 선택된 코덱으로 압축해서 압축 데이터 버퍼에 쌓아 놓는다. RTP 송신자는 압축된 오디오 데이터 버퍼로부터 하나의 패킷 양만큼 가져와 RTP 인코드를 통해 RTP 헤더를 씌워서 보내게 된다. 보내고 난 후에는 항상 RTCP 전송 간격을 검사해서 RTCP 전송 시간이 되었다면 바로 RTCP 패킷을 만들어 보낸다. 이 RTCP

패킷의 데이터는 RTP 송신자, 수신자와 RTCP 수신자로부터 얻은 정보가 된다. RTCP 수신자는 2개의 쓰레드가 작동할 때마다 항상 이용된다. 상대방으로부터 전달받은 RTP 패킷이 RTP 포트를 통하여 들어오면 수신자 RTP 버퍼에 쌓이게 되고, RTP 수신자는 이 RTP 데이터를 가져온다. 가져온 RTP 데이터는 RTP 헤더의 분석을 통해 필요한 정보를 취하고 난 후 RTP 헤더를 제거한 후, 압축된 오디오 데이터를 압축 오디오 버퍼에 쌓아둔다. 이 데이터는 다시 오디오 복원부에서 한번의 플레이 시간만큼 가져와 복원한다. 복원된 오디오 데이터를 오디오 플레이 버퍼에 쌓아두면, 오디오 출력 디바이스가 차례차례 가져가 플레이한다.

4.2 영상회의 시스템의 구현

본 논문에서 구현한 영상회의 프로그램은 현재 Window 98/ME/2000에서 운용되며, H.323 및 SIP를 지원하도록 설계되었다. 즉, H.323 표준을 지원하는 터미널이 호 연결을 신청하면 H.323으로 연결되며, SIP 표준을 지원하는 터미널이 호 연결을 지원하면, SIP 모듈이 작동되어 SIP로 연결되게 된다. (그림 9)는 QCIF 해상도로 구현된 터미널이 H.323으로 작동되는 과정을 나타낸 그림이다. 작동 과정은, 먼저 호 설정단계에서 회의 생성 프로토콜을 이용하여 회의를 시작하기 위한 신호를 주고받는다. Setup 메시지는 호출하려는 터미널에게 연결 설정 요구를 지시하기 위해 호출 터미널에게 보내진다. Call Proceeding 메시지는 요청된 호 설정이 개시되었고 더 이상의 호 설정 정보는 접수하지 않겠다는 것을 지시하기 위해 피 호출 터미널에 의해 보내진다. Alerting 메시지는 피 호출 사용자 정보가 개시되었음을 지시하기 위하여 피 호출 터미널에 의하여 보내진다. Connect 메시지는 호의 수락을 지시하기 위해 피 호출 터미널에 의해 호출 터미널로 보내진다. 둘째로, 호 설정이 이루어지면 그 결과로 H.245 프로토콜을 주고받기 위한 H.245 세션이 설정되고 이후부터는 이를 통하여 H.245 본연의 회의 설정 절차에 관한 프로토콜이 수행되며 그 과정은 수신된 메시지를 정확하게 분석하기 위해 사용되는 H.245의 버전 넘버를 결정하기 위해 터미널들 사이에 TerminalCapabilitySet 메시지를 교환하며, master/slave 결정과정에 의해 Active MC를 결정한다. 그 뒤, H.245 프로토콜에 따라 사용할 수 있는 오디오 및 비디오 채널의 특성을 정하여 이에 맞도록 각 미디어 별로 RTP 세션을 설정하고 이를 이용하면 된다. (그림 9)에서 음성은 G.723.1을 이용하고 있으며, 영상은 H.263을 이용하고 있다. 그 외에도 음성은 G.711, G.728을 지원하며 영상은 H.261을 지원한다. 음성 및 영상 코덱의 선택은 H.323의 경우에는 H.245의 능력교환을 통하여 이루어지며, SIP의 경우에는 SIP를 통하여 이루어진다.



(그림 9) QCIF 해상도로 H.323으로 작동되는 과정



(그림 10) CIF 해상도로 H.323으로 작동되는 과정

(그림 10)은 CIF 해상도로 영상회의 프로그램이 H.323 모드로 작동되는 과정을 도시한 그림이다. CIF 해상도의 경우에도 H.261, H.263이 지원되며, 이 경우에는 H.263으로 작동된다. 또한, H.323의 경우 H.245의 능력교환에 따라서 비대칭 모드로 작동하는 것도 가능하다. 즉, 비디오는 2가지 경우가 가능하다. 첫 번째는, 보내는 영상은 H.263으로 압축하고, 받는 영상은 H.261로 압축된 영상인 경우, 이 경우에는 압축부는 H.263이 작동하고, 복원부는 H.261이 작동하

는 경우이다. 두 번째는, 보내는 영상의 해상도는 CIF이고, 받는 영상의 해상도는 QCIF인 경우이거나, 그 반대의 경우이다. 음성의 경우도 유사하게 압축 코덱부와 복원 코덱부가 틀리게 작동할 수 있다. 즉, 압축 코덱부는 G.723.1로 작동하고 복원 코덱부는 G.728로 작동할 수 있다.

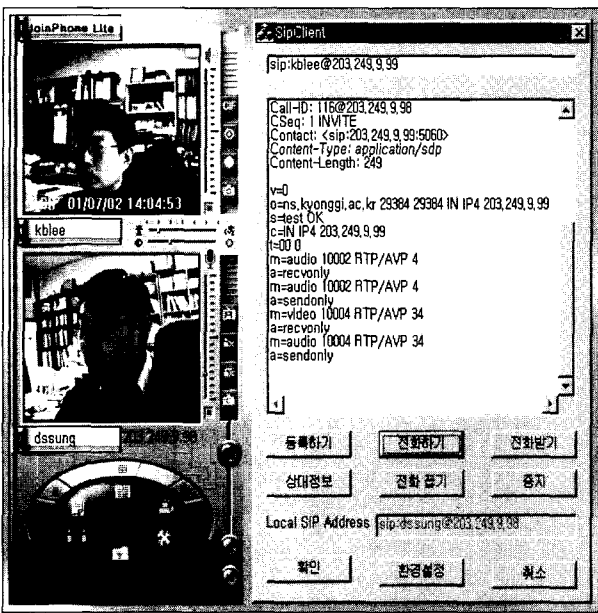
표준안으로 구현된 프로그램은 상호운용 시험을 통하여 표준안으로 적절히 구현됨을 실험해야 한다. 이를 위하여 H.323은 기존 여러 제품들과 상호운용 시험을 하였으며, SIP의 경우에는 상호운용 시험이 진행되고 있다. (그림 11)은 Microsoft사의 Netmeeting과 상호운용 작동되는 과정을 나타낸 그림이며, 이 외에도 FVC사의 CU-SeeMe Pro, SmithMicro사의 VideoLink Pro 등 H.323 프로그램과 Polycom사의 ViaVideo, ViewStation, Sorenson사의 Envision 등 하드웨어 기반의 H.323 터미널들과의 상호운용 실험하였다.



(그림 11) Netmeeting과 상호운용 작동되는 과정

(그림 12)는 QCIF 해상도로 SIP로 작동되는 과정을 나타낸 그림이다. 우측 그림은 SIP 프로토콜의 검증을 위하여 만든 부분이며, INVITE 요청 메시지를 자세히 기술하고 있다. 주소가 dssung@203.249.9.98인 UA가 음성과 영상은 각각 RTP 10,002 포트, RTP 10,004 포트를 사용하고, 오디오

오 코덱은 G.711A, G.711U, G.728, G.723.1, 비디오 코덱은 H.261, H.263을 사용하여 전송 및 수신이 가능함을 SDP에 기술하여 INVITE 메시지의 메시지 바디에 포함시킨 후 SIP 주소가 kblee@203.249.9.99인 UA에게 전송한다. INVITE 메시지를 수신한 kblee@203.249.9.99인 UA는 이에 대한 확인 메시지로 180 Ringing 응답 메시지를 자동으로 전송한다. 이때, INVITE 요청 메시지를 받은 UA가 200 OK 응답 메시지로 INVITE 메시지에 대한 확인 메시지를 전송하면, 이에 대한 ACK 메시지를 전송 받는다. 이 절차에 따라, 결과적으로 상대방과 영상 회의를 할 수 있다.



(그림 12) QCIF 해상도로 SIP로 작동되는 과정



(그림 13) VoIP를 이용하여 일반전화기와 통화하는 과정

H.323 및 SIP는 게이트웨이를 이용하여 일반전화기와의 음성통화가 가능하다. 이를 VoIP(Voice over IP)라하며 H.323 및 SIP 터미널에서 일반전화기로 연결하여 통화하는 과정을 PC-to-Phone이라 하며 일반전화기에서 H.323 및 SIP 터미널로 연결하여 통화하는 과정을 Phone-to-PC라고 한다. 현재 개발되어 판매되고 있는 게이트웨이는 대부분 H.323 게이트웨이이며 SIP 게이트웨이기도 개발되고 있다. 이와 더불어 H.323 터미널과 SIP 터미널 사이의 연결을 가능하게 하는 H.323-to-SIP 게이트웨이기도 개발되고 있다. 본 논문의 시스템은 PC-to-Phone 및 Phone-to-PC를 모두 지원하며 (그림 13)은 PC-to-Phone 방식으로 H.323 게이트웨이를 이용하여 작동되는 과정을 나타낸 그림이다.

4.3 다른 영상회의 시스템과의 비교

현재 개발되어 있는 영상회의 시스템은 표준으로 개발된 시스템들과 비표준으로 개발된 시스템들이 있으나, 상호 호환의 중요성이 대두되면서 비표준으로 새롭게 개발된 시스템들은 거의 없는 실정이다. 표준으로 개발되고 있는 영상회의 시스템은 H.323을 이용한 시스템[18]과 SIP를 이용하는 시스템[19]으로 구분할 수 있으며, 현재는 거의 대부분의 시스템이 H.323을 이용하고 있으나, 확장성 및 인터넷상의 다른 프로토콜과의 친화성의 문제로 인하여 SIP를 이용하는 시스템이 많이 개발되고 있으며, 그 비율이 점점 높아 지리라 예상된다. <표 1>에서는 본 논문에서 구현한 시스템과 기존에 개발된 시스템들과의 장단점을 비교하였다. 본 논문의 시스템의 장점은 SIP 및 H.323 프로토콜 스택을 모두 지원한다는 점이며, 따라서 현재 표준화가 진행중인 SDP를 이용하는 H.323 다자간 화상회의로의 확장도 용이하다. 또한 본 시스템은 SIP 및 H.323표준이 모두 포함되어 있기 때문에 이를 이용하여 H.323-to-SIP 게이트웨이의 개발도 용이하다. 현재 본 시스템은 게이트키퍼(Gatekeeper)

<표 1> 다른 영상회의 시스템과의 비교

	Netmeeting	Columbia-SIP	본 시스템
H.323 프로토콜 지원	○	×	○
H.225 호 지원	○	×	○
H.245 지원	○	×	○
SIP 호 지원	×	○	○
SDP 지원	×	○	○
RTP/RTCP 지원	○	○	○
집중식 다자간회의 지원	○	×	○
분산식 다자간회의 지원	×	×	×
VoIP (Phone to PC)	×	×	○
VoIP (PC to Phone)	○	×	○
Gatekeeper 등록	○	×	○
SIP proxy 등록	×	○	○

에 등록하여 ID를 이용한 단말기간 통신이 가능하며, H.323 게이트웨이 설정을 통하여 PC-to-Phone 인터넷 전화를 지원하고 있다. 또한 FVC MeetingPoint 또는 Cu-SeeMe Conference Server H.323 다지점 통신장치와의 연결을 통하여 Non-CP(Continuous Presence) 모드 회의실에 참가하여 다자간 회의를 지원하고 있다. Non-CP 모드에서는 회의에 참가한 사용자들의 음성을 기준으로 발표자의 영상을 모든 참가자들에게 전달하는 방식을 지원한다. 현재 H.323을 지원하는 다지점 통신장치를 개발하고 있으며, 추후과제로 H.323 및 SIP를 지원하는 다자간 회의 시스템을 개발할 예정이다.

5. 결 론

본 논문에서는 영상회의 표준안인 ITU-T H.323 및 SIP를 지원하는 멀티미디어 터미널을 설계하고 구현하였다. 다양한 정보통신의 발전과 함께 다양한 응용 서비스가 개발되고 있으며, 일반 사용자들이 이를 어디에서나 쉽게 이용할 수 있게 되었다. 이러한 변화로 음성이나 동화상 등 멀티미디어 서비스를 위한 다양한 대칭 및 비대칭 응용 프로토콜이 제안되고 발전되고 있다. 대칭 통신 응용 서비스 중 지역적으로 떨어져 있는 사용자들이 대화형식으로 회의를 진행하기 위한 영상회의 프로토콜은 매우 중요하고 활용도가 높은 분야이다. 그 중 인터넷망에서의 영상회의 시스템은 활용도가 매우 뛰어나며 관심을 받고 있으며, 이를 위하여 ITU-T 및 SIP에서는 표준안을 내고 있다. H.323은 인터넷에서의 영상회의를 할 수 있도록 하기 위한 ITU-T의 표준안이며, SIP는 같은 목적을 위한 IETF의 표준안이다. 현재는 H.323 표준안이 많이 사용되고 있으나, 상당 기간 2개의 표준안이 공존하리라 예상된다. 따라서, 두 표준안 사이에 영상회의가 수행되어야 하며, 이를 위하여 H.323-SIP 게이트웨이의 개발이 진행되고 있다. 현재는 모든 영상회의 터미널은 H.323 또는 SIP만을 지원하고 있으며, 두 표준안 사이의 상호운용은 게이트웨이를 이용해야만 한다. 이와는 별도로 2개의 표준안이 모두 지원되는 터미널의 개발도 필요하게 되었다.

따라서, 본 논문에서는 인터넷 환경에서 영상회의 표준안인 ITU-T H.323 및 SIP를 지원하는 멀티미디어 터미널을 구현하였다. 상대의 터미널이 H.323을 지원하는 터미널인 경우 H.323 제어 모듈 중 H.225를 통하여 호가 설정되며, 능력교환은 H.245를 통하여 이루어진다. 그 뒤, 오디오 및 비디오 코덱을 통하여 압축된 정보를 RTP를 통하여 전송한다. SIP의 경우 SIP 제어 모듈내 SIP를 통하여 호가 설정되며 능력교환은 SDP를 통하여 이루어진다. 그 뒤, 영상 및 음성 코덱, RTP, RTCP는 H.323에서 이용하는 동일한

모듈을 이용한다. 구현된 시스템은 공통부분을 최대한 이용하도록 설계되어 각각을 따로 설계하는 것 보다 복잡도가 줄어들음을 알 수 있었다. 또한, 현재 표준화되고 있는 H.323 다자간 영상회의의 경우 H.323을 그대로 이용하면서 SDP를 추가적으로 이용함을 알 수 있다. 따라서, 약간의 코드 추가와 함께 H.323도 지원하는 복합 영상회의 시스템을 설계할 수도 있음을 알 수 있었다.

향후 과제로는 첫째, 다자간 회의를 위하여 H.323 및 SIP를 지원하는 다지점 제어장치를 개발하는 것과 둘째, H.323 게이트키퍼와 SIP 서버와의 통합 서버의 개발이 필요하며, 이 경우 H.323과 SIP의 통합이 가시화 될 것으로 생각된다.

참 고 문 헌

- [1] V. Kumar, M. Korpi, S. Sengodan, IP Telephony with H.323, Wiley, 2001.
- [2] Uyless Black, Voice Over IP, Prentice-Hall, 2000.
- [3] ITU-T Recommendation H.323 : Packet-based multimedia communications, Nov., 2000.
- [4] M. Handley, H. Schulzrinne, E. Schooler and J. Rosenberg, "SIP : Session Initiation Protocol," Request for Comments 2543, Internet Engineering Task Force, Mar., 1999.
- [5] Radhika R Roy, Vipin Palawat, Alan Johnston, Charles Agboh, David Wang, Henning Schulzrinne, Kundan Singh, Joon Maeng, "SIP-H.323 Interworking," Internet Draft, Internet Engineering Task Force, Feb., 2001.
- [6] H. Schulzrinne, J. Rosenberg, "A Comparison of SIP and H.323 for Internet Telephony," Network and Operating System Support for Digital Audio and Video (NOSSDAV), Cambridge, England, Jul., 1998.
- [7] ITU-T Recommendation H.225.0, Call signalling protocols and media system packetization for packet-based multimedia communication system, Nov., 2000.
- [8] ITU-T Recommendation H.245, Control Protocol for Multimedia Communication, Feb., 2000.
- [9] RFC Recommendation 1889 RTP, A Transport Protocol for Real-Time Applications, Jan., 1996.
- [10] RFC Recommendation 1890 RTCP, A Transport Control Protocol for Real-Time Applications, Jan., 1996.
- [11] Alan Johnston, Robert Sparks, Steve Donovan, Kevin Summers, "SIP Service Examples," Internet Draft, Internet Engineering Task Force, Mar., 2001.
- [12] M. Handley, V. Jacobson, "SDP : Session Description Protocol," Request for Comments 2327, Internet Engineering Task Force, Apr., 1998.
- [13] ITU-T Recommendation H.261, Video Codec for Audio-visual Services at p×64Kbits, Mar., 1993.
- [14] ITU-T Recommendation H.263, Video Coding for Low Bit

Rate Communication, Jan., 1996.

- [15] ITU-T Recommendation G.711, Pulse Code Modulation (PCM) of voice frequencies, 1972.
- [16] ITU-T Recommendation G.728, Coding of Speech at 16kbps using Low-Delay Code Excited Linear Prediction, Sep., 1992.
- [17] ITU-T Recommendation G.723.1, Dual Rate Speech Coder for Multimedia Communications Transmitting at 5.3 and 6.3Kbps, Mar., 1996.
- [18] <http://www.microsoft.com>.
- [19] <http://www.cs.columbia.edu/~hgs/sip/>.



성 동 수

e-mail : dssung@kyonggi.ac.kr

1987년 한양대학교 전자공학과(공학사)

1989년 한국과학기술원 전기 및 전자공학과
(공학석사)

1992년 한국과학기술원 전기 및 전자공학과
(공학박사)

1992년~1993년 한국과학기술원 정보전자연구소 연구원

2002년~2003년 University of Washington 방문연구교수

1993년~현재 경기대학교 전자공학부 전자공학전공 부교수

관심분야 : MoIP, 멀티미디어통신, 병렬처리