

신뢰성있는 화상회의를 위한 CTI System 설계

이종열, 정현우, 박원배

경북대학교 정보통신학과

전화 : (053) 940-8897 / 팩스 : (053) 950-5508

The Design of a CTI System for reliable video-conference

Jong Youl Lee, Hyun Woo Jung, Won Bae Park

Dept. of Information & Communication, Kyungpook National University

e-mail : theoslee@inc.knu.ac.kr, carry@inc.knu.ac.kr, wbpark@ee.knu.ac.kr

Abstract

In this paper, a design of the reliable video-conference system using CTI(Computer Telephony Integration) technology is proposed. When video-conference is run on the current existing Internet, the transmission delay problem for voice data traffic can be frequently occurred. In order to transmit the real-time voice data through the Internet efficiently, some complicated algorithms such as CODEC(Code/Decode) should be applied. It can cause further excessive processing delay which can affect the overall performance.

The voice traffic is usually transmitted through the reliable PSTN(Public Switched Telephone Network) in the CTI system. In this paper a new architecture, in which PSTN for voice traffic and Internet for video traffic are used at the same time instead of using Internet by itself, is proposed to relieve the problems on a video conference.

I. 서론

요즘 음성/화상/문자 등 다양한 형태의 데이터를 응용하는 멀티미디어(multimedia)기술의 도입으로 인해 광역화된 시스템의 기능과 고도화된 기술로 급속한 사

용자 요구 수준 향상 등에 부응하여 효율적인 시스템 통합 관리가 필요하게 되었다.

근래의 초고속 정보통신망으로 대표되는 정보 관련 사회 간접자본의 지속적인 확충 및 컴퓨터 업계에서 일고있는 멀티미디어 응용 기술의 급속한 발전은 멀티미디어 기술 적용을 가속화하고 있다. 작업환경 개선 요구의 증대 및 널리 분포하고 있는 네트워크 자원을 더욱 효율적, 경제적으로 운용해야 할 필요성이 증대함에 따라 멀티미디어 응용시스템의 연구, 도입이 전 세계적인 추세가 되어가고 있다.

이러한 시대 흐름에 맞춰 CTI(Comuter Telephony Integration) 응용시스템은 멀티미디어 정보의 응용에 의해 현장 또는 원격지 사용자의 정확하고 신속한 의사결정을 지원하여 시스템의 효율성 및 네트워크 자원의 가용성을 높일 수 있으며 여러 경로로 유입된 정보를 신속하게 처리하는 기능을 제공함으로써 시스템의 전반적인 성능을 향상시키는 데 결정적인 역할을 하게 된다[1]. 더불어 인터넷(Internet)으로 대표되는 외부망 접속 기능과 일반화된 전화망의 장점을 이용해 원격지에서 여러 사람이 마치 같은 장소에서 회의를 하듯 서로의 의사를 교환할 수 있다.

CTI라는 것은 컴퓨터의 데이터 처리능력 및 라우팅 기능 과 전화망의 호제어 및 전화제어의 장점을 결합한 기술로써 대표적인 응용분야로는 정보검색과 처리(Auditex, Fax-on-Demand, Interactive Voice Response)분야와 Messaging(Voice mail, Fax Server, Electronic Mail Reader)분야 Connectivity(Call Center,

Tele-marketing, Call Center)분야가 있으며, 대표적인 표준화 단체로는 CSTA(Computer Supported Telephony Application)나 VERSIT(diVERSITY)이 있다[2].

기존의 널리 분포하는 PSTN(Public Switched Telephone Network)망에서의 데이터 전송에 있어의

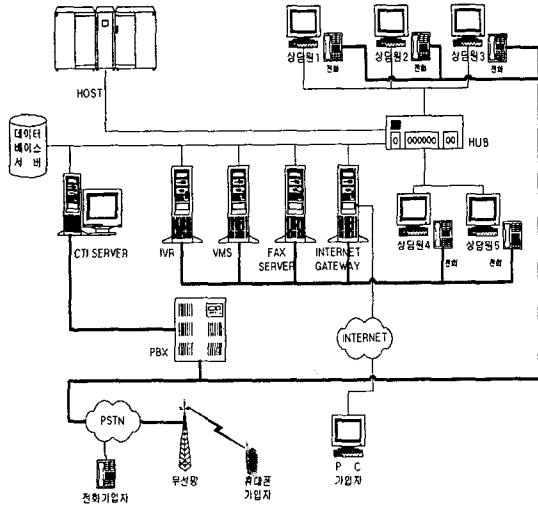


그림 1. <CTI System의 일반적인 구성도>

장점인 신뢰성 및 고정적인 대역폭을 음성의 전송을 위한 매체로 이용하고 컴퓨터에서의 다양한 데이터 처리능력을 이용하여 거대한 인터넷망을 통한 화상전송 구현에 대해 본 논문에서 기술하고자 한다.

인간의 의사 소통 의존도가 시각적인 면보다는 음성 에 그 비중이 크다는 것을 감안하여 음성을 좀더 신뢰성 있고 전송속도가 일정한 회선교환 방식인 전화망을 이용하여 전송하고, 이에 반에 의사 소통에 기여도가 상대적으로 낮은 화상을 인터넷이라는 패킷 스위칭 방식을 이용하여 전송한다.

실제 화상회의 시스템에서는 프리젠테이션이나, 화상강의의 경우, H.323 프로토콜 스택의 T.120 protocol 을 이용한 화이트보드와 같은 것을 데이터 전송 프로토콜로 구현함으로써 청자의 이해를 도울 수도 있다.

본 논문의 II장에서는 CTI 화상회의의 구현 알고리즘, III장에서는 기존의 화상회의의 시스템과의 비교, IV장에서는 본 시스템을 응용할 수 있는 분야, 그리고 끝으로 V장에서는 결론 및 추후 연구에 관하여 기술한다.

II. CTI 화상회의의 구현 알고리즘

다음장의 CTI시스템 구성도에서 CTI Server가 하는

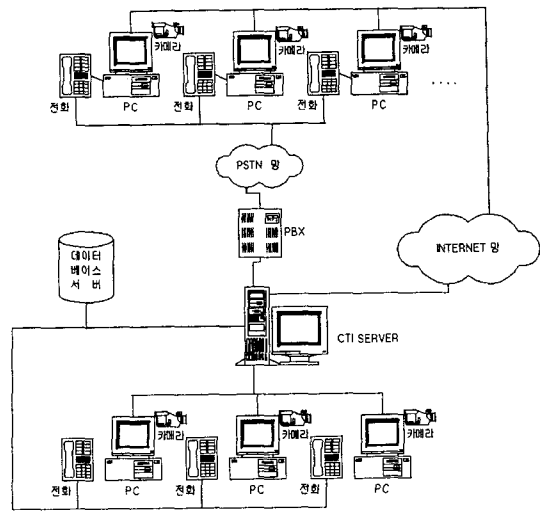


그림 2. <화상회의를 위한 CTI 시스템 구성도>

역할은 전화에 관한 것과 화상에 전송에 관한 것 둘로 나눌 수 있다. 전화에 관해서는 CTI Server내에 내장된 Telephony card가 전화를 받고 그 받은 전화신호를 역시 내장된 스위칭 카드에 연결함으로써 호 전환을 시킨다[3-4]. 가령 예를 들어 4라인으로부터 들어온 전화를 각각 독립된 스레드(Thread)로 처리하고 이를 스위칭 카드에서 호 전환을 함으로써 외부 전화를 쓰는 4명의 참가자와 1명의 주관자가 서로 통화할 수 있게 할 수 있다.

화상 압축 방식은 화상통신 표준인 H.320의 LAN환경에서 사용되는 H.261을 사용하는데, 이 프로토콜의 특징은 압축률은 낮지만 압축속도는 빠르다는 것이다. 화상포맷은 미·일 방식의 NTSC나 PAL사양의 카메라 신호가 가진 해상도를 미리 간축한 공통 중간 포맷인 CIF(Common Intermediate Format)방식을 이용하는데 이 포맷은 프레임 스킵(skip)을 허용한다[5]. 화상을 멀티캐스트하기 위해서는 MCU(Multi-point Control Unit)가 필요하다. 이 MCU는 point-to-point 방식의 멀티캐스트로서 셋 또는 그 이상의 중단점 사이에 회의를 지원하기 위한 것으로 MC(Multi-point Controller)와 MP(Multi-point Processor)로 구성된다.

본 논문에서 화상전송에 관해서는 기존의 H.323 프로토콜에서 쓰는 방식 즉, 중앙 집중식(Centralized) MCU방식을 그대로 도입한다. 즉 모든 터미널(컴퓨터)에서 전송된 모든 화상 정보는 CTI 서버가 내부 MCU를 이용해 처리한 다음 각 터미널에게 그 화상 정보를 분배를 해준다.

MCU는 MC와 MP로 나눌 수 있는데 MC가 하는 역할은 H.323프로토콜 스택중 H.245(Control Protocol

for Multi-media Communication)을 이용하여 오디오

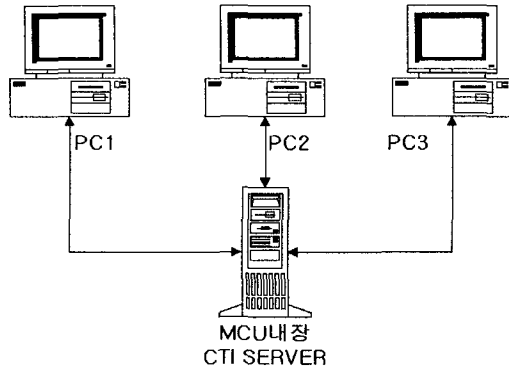


그림 3. <중앙 집중식 MCU를 내장한 CTI server>

및 비디오 처리 시에 호환이 터미널간에 가능한지 결정하는 것으로[6-7] 본 논문에서는 비디오 데이터에 관해서만 호환성을 조사하게된다. 즉, 어떠한 데이터 스트림을 직접처리를 하지 않고 MP에게 넘긴다. MP가 하는 역할은 MC로 부터 넘겨받은 비디오와 오디오 데이터들의 비트를 처리하고 선택 및 혼합한다.[6-8]

화상의 실시간 전송을 위해서는 RTP / RTCP (Realtime Transfer Protocol / Realtime Transfer Control Protocol) 프로토콜을 사용하는데 RTP는 실시간 데이터를 멀티캐스트 또는 유니캐스트 네트워크를 이용해서 전송할 수 있고, 필요에 따라 헤더를 변경 혹은 추가하여 응용에 맞는 프로토콜로 이용이 가능하다. 이에 부가적으로 RTCP는 RTP 제어 프로토콜로서, 주기적인 제어 패킷을 보내어 데이터의 전달 상황을 감시하며, 최소한의 제어기능과 매체 식별 기능을 제공한다[9-10].

음성 데이터와 화상 데이터의 동기화 문제는 음성 데이터의 전달 시간은 일정하다는 가정으로 화상 데이터에 시간을 기록하고 음성 데이터의 전송시간에 맞추어 동기화 시킨다. 만약 네트워크의 과부하로 데이터 전송이 지연되면, 적당한 범위 내에서 데이터를 삭제하고 필요한 부분만 전송하도록 한다. 또한 화상회의에 참가한 화자가 전화를 통해 음성데이터를 전송 시 그에 맞춰 화상데이터를 전송하고 화자의 침묵 시에는 화상을 전송하지 않고 정지 화상을 보여주는 방식으로 진행이 된다.

에러 처리에 관해서는 어플리케이션 계층에서 화상 패킷의 손실정도를 파악하여 실시간 전송 및 전송지연에 크게 영향을 주지 않는 범위 내에서 에러 복구 과정을 수행한다. 그리고 인터넷 네트워크 사정이 부득이한 경우에는 화상 전송을 중단하고 음성 데이터만을

전송하여 회의가 계속 진행되도록 한다.

III. 기존의 화상회의 시스템과의 비교

본 논문에서 기존의 화상회의 시스템, 즉 H.323 Protocol을 이용한 화상회의 시스템과 비교해 보았을 때 그 차이점은 다음과 같다.

첫째, QoS(Quality of Service)가 보장되지 않는 인터넷상에서 H.323 프로토콜의 음성전송의 경우 압축 코덱(CODEC)구현과 연결설정(Call Setup) 및 Internet telephony gateway를 구현[6]해야 되는 복잡한 문제가 있으나 본 논문에서의 시스템은 기존의 PSTN망을 이용하므로 추가적인 구현문제를 피할 수 있다.

둘째, H.323프로토콜에서는 음성채널과 화상채널을 인터넷이라는 동일한 망에서 구현하는 반면 본 논문에서의 구현은 음성채널을 PSTN망을 그리고 화상채널은 인터넷망을 이용하므로 그 연결 생존성이 높다 할 수 있다.

셋째, H.323프로토콜에서의 다자간 통화를 위한 MCU(Multipoint Control Unit) 구현 및 음성 믹싱[6] 기술이 필요한 반면 본 논문의 시스템은 기존의 전화 교환기의 스위치를 이용하므로 그 구현이 간단하다.

넷째, H.323프로토콜에서는 음성까지 인터넷으로 전송하여 인터넷상에서의 트래픽을 가중시키는 반면 본 논문의 시스템은 인터넷과는 별도의 음성전송망(PSTN)을 사용하므로 인터넷에서의 부하를 줄이며, 보다 나은 고품질의 음성을 전송할 수 있다.

다섯째, 인터넷으로의 음성 전송시 음성반향(Echo) 문제와 화자간의 대화시 원활한 대화를 방해하는 음성 전송지연은 본 논문에서 제시하는 시스템의 경우엔 문제가 되지 않는다.

IV. 본 시스템을 응용할 수 있는 분야

시간적 및 공간적인 제약을 초월해야 하는 현대 업무에서 원격회의는 필수적이며 높은 신뢰성과 회의의 원활한 진행이 요구된다. 위와 같은 요구사항을 충족시켜 주기 위해 본 논문의 시스템으로 응용할 수 있는 모델을 다음과 같이 제시한다.

1. 사정상 학교 학부형회의에 참석할 수 없는 학부모를 위한 학부형회의

이 모델에서는 주관자가 각 회의 참가자에게 회의의 개최시간을 미리 통보함으로써 참가자가 CTI 서버로 전화 및 컴퓨터를 통해 접속하게된다.

2. 실시간 화원 화상강의

이 모델에서는 참가자가 실시간 화원강의의 시간 내에 접속하게 되면 강의록 들을 수 있고 학생의 질문 및 강사의 답변이 가능하다. 단 참가자의 화면에는 참가자들의 화상이 디스플레이 되는 것이 아니라 강사의 화상과, 강의 진행에 필요한 화이트보드 또는 프리젠테이션 자료가 나타나게된다.

3. 기업 본사와 지사간의 화상회의

원거리 인터넷 사용이 필요한 모델로 인터넷상의 다양한 변수 요인(Network congestion, Physical error 등)으로 연결설정이 끊어지는 경우가 있을 수가 있으나, 인터넷과는 독립된 전송 매체인 PSTN망을 통해 음성데이터를 전송함으로써 회의는 계속 진행될 수 있다.

4. 컴퓨터 부재시 전화만을 이용한 전화상의

음성회의

이 모델에서는 화상이 선택사항이 되어 꼭 화상이 필요하지 않거나, 컴퓨터 사용이 불가능(부재나, 고장)한 참가자들이 컴퓨터의 유무에 관계없이 전화만으로 회의에 참가할 수가 있다.

[2] 권경협, "CTI산업의 표준화 동향, CTI 응용분야"
<http://www.naininfo.co.kr/CTI>

[3] "The Features of D/41ESC",
http://www.dialogic.com/products/d_sheets/2550-web.htm

[4] "The Features of MSI/SC-Global",
http://www.dialogic.com/products/d_sheets/4549-web.htm

[5] "디지털 영상 압축 기술의 분류",
<http://www.sjnc.com/tech/tech01.html>

[6] ITU-T Recommendation H.323,
 "Visual Telephone Over Non-Guaranteed Qos LANs"

[7] The International Engineering Consortium,
 "H.323", <http://www.iec.org>

[8] "H.323 white paper",
http://www.elemedia.com/Main/h323center/H323_Version2.htm

[9] "RTP : Overview",
<http://www.cs.columbia.edu/~hgs/rtp/>

[10] "실시간프로토콜(RTP)",
<http://cherry.soongsil.ac.kr/~rtsp/>

V. 결론 및 추후 연구

본 논문에서는 기존의 화상회의 시스템과 비교해서 음성데이터 전송에 좀 더 비중을 둔 화상회의 시스템을 CTI응용으로서 설계하는 것에 대해 제시하였다. 인터넷이라는 QoS(Quality of Service)가 보장되지 않는 네트워크환경에서 화상회의 시스템을 구성하는 것 보다 음성데이터와 화상 데이터를 서로 다른 매체(PSTN 망, Internet 망)를 통해 전송함으로써 연결설정의 신뢰성을 높이고, 음성데이터를 확실하게 전송할 수가 있다는 것이 본 논문의 요지이다.

화상전송에 있어 음성 데이터와의 좀 더 정확한 동기화를 위해 전화 신호의 일부를 컴퓨터 내부로 입력해 그 음성 신호의 데이터 양에 따라 초당 전송할 화상 프레임의 수를 가변 시키는 방법과 정지 화상이 입력으로 들어 올 경우 그 프레임을 전송하지 않는 방법이 요구된다. 또한 CTI Server의 데이터 처리 과부하를 막기 위해 분산식 MCU나 MCU를 독립형으로 구성하는 방법이 필요하다.

참고문헌(또는 Reference)

[1] <http://camars.kaist.ac.kr/~mrg/first-proj/index.html>