

광대역 ISDN상의 다지점 멀티미디어 통신서비스 II부 : Non-Transcoder근간의 MPEG 비디오브리지

정희원 박정호*, 황대환**, 이종형**, 구한준***, 조규섭**

Multipoint Multimedia Communication Service in Broadband ISDN Part II: MPEG Video Bridge based on Non-transcoding Mechanism

Jung Ho Park*, Dae Hwan Hwang**, Jong Hyeong Lee**, Han Joon Ku***, Kyu Seob Cho*
Regular Members

* 본 논문은 한국전자통신연구원(ETRI)의 위탁과제로 수행되었음.

요약

화소영역에서의 처리방법이 주를 이루고 있는 Multipoint Control Unit(MCU)의 화면분할(Split-Screen) 처리기능은 동영상신호의 실시간 처리에 많은 문제를 갖고 있다. 이를 해결하기 위한 방법이 일부 연구되고 있으나, 이 방법도 실제 적용하기에는 무리가 따르는 복잡한 구조를 갖거나 적용대상이 H.261으로 한정되어 있다. 본 논문은 ISO/IEC의 동영상 부호화 표준인 MPEG-1/2를 대상으로 하는 새로운 화면분할 방법을 제안한다. 새롭게 제안한 MPEG 화면분할 방법은 완전 압축영역에서 처리를 수행하며 이에 따르는 처리지연이 거의 수반되지 않으므로 실시간 다지점 멀티미디어통신 등에 응용이 가능한 효율적인 방법이다. 또한 이 방법은 MPEG 비트 스트림의 단순 해석 및 조작으로 구현이 가능하며, 압축영역의 분할방법을 사용함으로써 화질의 열화를 수반하지 않는다. 제안한 방법에 대한 소프트웨어 시뮬레이션을 통하여 기능의 검증이 완료하였으며, 기존의 방법과 구성의 용이성, 처리에 따르는 지연 그리고 분할처리된 화면의 화질 비교를 통해 제안한 방법의 효용성을 입증하였다.

ABSTRACT

The split-screen function on the multipoint control unit(MCU) which is usually using processing method based on pixel domain has many problems for manipulating the video signal in real-time. Although the researches and the developments to cope with such problems are proceeding, these have too complex architecture to implement and are limited to method for H.261 video signal. In this paper, we propose a new mechanism for split-screen that can actually apply to ISO/IEC

* 데이터주식회사 부설연구소장
** 한국전자통신연구원(선임연구원)
** 한국전자통신연구원 통신망구조연구실 근무
*** LG정밀 기술전략팀 연구원
** 성균관대학교 전기, 전자 및 컴퓨터공학부 교수
論文番號 : 98020-0113
接受日字 : 1998年 1月 13日

MPEG video standard. The new method that is proposed in this paper do the processes in complete compression domain, thus it is suitable for the application of real-time multipoint multimedia communication service. By simply interpreting and manipulating the MPEG video element stream, the split-screen functional module can be implemented easily and the result of the procedures does not accompany image degradation at all. Finally, the complexity of implementation, the aspect for processing delay and the loss of image quality as compared to that resulting in the case of applying the previous split-screen method has been investigated. And it is confirmed that the proposed mechanism has a significant advantage as a split-screen method.

I. 서 론

광대역 ISDN의 하부구조를 이루는 교환 및 전송 분야의 연구, 개발의 완성과 ITU-T, ISO/IEC를 축으로 하는 오디오, 비디오의 압축표준이 완료됨에 따라 광대역 ISDN상에서의 멀티미디어 서비스는 사용자에게 직접 적용이 가능한 형상으로 구체화되고 있다. 향후 전개될 멀티미디어 서비스의 형태로는 영상전화, 영상회의와 같은 대화형서비스와 Video on Demand(VoD)와 같은 교신, 분배형서비스를 근간으로 가시화될 것이다. 대화형 멀티미디어 통신서비스는 H.261 영상압축 코딩방식을 기반으로 다양한 네트워크에서 적용이 가능한 H.320계열의 멀티미디어 통신단말이 ITU-T에서 표준이 완료되고 있으며, Digital Audio-Visual Council(DAVIC)에서는 ISO/IEC의 MPEG표준을 이용하여 VoD형 서비스시스템의 표준을 완료한 상태이다. 멀티미디어 통신단말의 형태는 대화형서비스에는 H.261 표준 압축방식을 사용하고, VoD와 같은 교신, 분배형서비스에는 MPEG 표준 압축방식을 적용하고 있으나, ITU-T에서 ISO/IEC의 MPEG을 수용한 단말을 표준화하고, DAVIC 표준 Set-Top Box(STB)에서 대화형 양방향통신서비스를 수용하고 있는 추세에 따라[1], 기존의 표준방식을 조합한 다양한 형태의 서비스 및 멀티미디어 단말이 존재할 것이다. 특히 ISO/IEC의 MPEG-2는 CATV, 디지털 TV방송 그리고 VoD와 같은 서비스에 직접 적용되고 있으며, 이와 같은 서비스는 이미 장치의 규격, 표준이 완료된 STB를 통해 제공이 가능하다. STB는 가정에 보급되어 있는 TV나 VTR 그리고 개인용 컴퓨터와 같은 별도의 장치 또는 기존의 장치에 포함되는 형태로 제품화될 것이며, 사용자는 이를 기반으

로 광대역 ISDN의 토대 위에서 다양한 멀티미디어 통신서비스로의 접근을 시도할 것이다. 특히 영상전화 및 영상회의와 같은 대화형 통신서비스의 제공은 현재까지 해당 네트워크의 형태에 따라 H.320계열의 단말을 사용함으로써 가능하였으나, STB를 갖고 있는 사용자에게는 별도의 단말을 구입하여야 하는 부담을 주게 된다. 따라서 이와 같은 경우에는 MPEG 디코더와 망 정합기능을 갖는 STB에 MPEG 인코더를 추가하여 양방향 멀티미디어 통신서비스를 제공하는 것이 보다 효율적인 방법이라 할 수 있다[2]. 멀티미디어 통신단말은 적용되는 서비스 및 내장된 장치의 특성에 따라 다양한 유형의 미디어신호 및 그 조합이 혼재할 수 있다. 이에 따라 다지점 멀티미디어 통신서비스를 제공하는 MCU도 다양한 미디어의 처리가 가능한 구조를 가져야 한다.

본 논문은 MCU의 미디어처리부 중 MPEG기반의 동영상처리에 관련된 내용이다. 기존 MCU의 영상처리부의 화면분할 처리장치는 화소영역의 처리방법이 주류였으며, 구성의 복잡성과 실시간 미디어처리 측면에서 많은 제약을 갖고 있었다. 이와 같은 문제점을 상당 부분 해결할 수 있는 압축영역에서의 처리방법에 대한 필요성은 인지되고 있으나, 국내외에서 극히 일부방식이 제안되고 있을 뿐이다[3,4,5]. 특히 영상의 복호 및 재부호화 처리가 필요없는 완전 압축영역에서의 화면분할 방법은 H.261동영상을 대상으로 고안되어 있을 뿐, 압축방식과 부호의 구문체계가 다른 MPEG근간의 동영상에 대해서는 연구되지 못한 실정이다. 본 논문에서는 MPEG동영상을 완전 압축영역에서 화면분할 처리하는 새로운 방법을 제안하였다. 제안한 방법은 MPEG 동영상을 사용한 단말의 다지점 멀티미디어통신을 위한 MCU의 효율적인 실시간

화면분할 처리방법으로 직접 적용이 가능하다.

II. MCU내에서의 비디오브리지

2.1. MCU에서의 영상처리

다지점 접속환경에서 영상회의와 같은 서비스를 제공하기 위한 MCU는 다지점 멀티미디어 통신을 위한 서비스 노드로 볼 수 있다. 따라서, MCU는 네트워크를 통해 접속된 다수 단말의 멀티미디어 신호들을 유용하게 처리하고 서비스 특성에 맞는 세션 제어기능을 수행하므로써 각 미디어별로 적합한 브리지 기능을 제공하여야 한다. 일반적으로 MCU의 오디오 브리지는 회의 참석자들 가운데 주 화자의 음성을 검출하고, 그 음성과 주변 화자의 음성을 믹싱하여 MCU에 접속된 모든 단말에게 분배하는 기능을 제공한다. 한편, MCU의 비디오 브리지는 영상 스위칭이나 화면분할(Split Screen) 처리와 같은 다양한 화면 구성을 연출할 수 있다. 영상 스위칭 기능은 오디오처리부의 검출되는 화자 정보를 바탕으로 현 화자에게는 이전 화자의 영상을 제공하고, 나머지 회의 참석자에게는 현 화자의 영상을 제공하도록 영상을 스위칭하는 것으로써 통상 비디오 브리지의 기본 기능이 되고 있다. 화면분할 처리기능은 다수의 참석자 화면을 공간적으로 재배치하여 참석자에게 분배하여 주는 것으로 기술적 어려움을 갖고 있으나, 사용자 입장에서 이와 같은 기능을 MCU에서 제공하여 준다면 보다 현장감 있는 회의가 가능하므로 반드시 필요한 기능으로 볼 수 있다.

2.2. 기존의 화면분할 방법

2.2.1. 화소영역에서의 화면분할

화소영역에서의 화면분할 방법은 인코딩된 입력 영상신호를 완전히 디코딩한 후에 화면분할 처리를 행하고 이에 따르는 출력 영상신호를 다시 인코딩하는 방법이다. 이와 같은 방법은 별도의 디코딩, 인코딩 과정을 거침으로 인해 영상품질의 열화를 가져오는 것은 물론 전체 단-대-단 차원에서 두배 이상의 지연을 일으킨다. 그러나 화소영역에서 신호처리는 다양한 영상처리가 가능하므로 화면구성의 융통성 및 유연성을 갖는다. 또한, 실제 영상회의 환경에 이미지 및

그래픽의 기법이 가미되고, 참가자가 동적으로 이를 처리할 수 있는 기능의 구현도 가능하대[6].

2.2.2. Discrete Cosine Transform(DCT)영역에서의 화면분할

DCT영역에서의 화면분할은 가변장 역부호화(Variable Length Decoding:VLD)기능을 수행한 다음, DCT계수들을 조작해서 영상을 처리하고 다시 가변장 부호화(Variable Length Coding:VLC)하는 방법이다. 화소영역의 화면분할 처리방법을 응용하면 DCT계수를 사용한 영상의 조작이 가능하다. DCT영역에서의 처리는 화소영역에서의 처리보다 적은 계산량을 가지나 다양한 화면조작 능력은 제한된다. 또한 Motion Compensation(MC)으로 압축된 영상일 경우 영상의 조작, 처리에 필요한 데이터를 재계산하는 복잡한 과정이 필요하고, MC데이터의 처리가 수행될 때마다 영상의 질이 격감되는 문제점이 있다. MC영역에서 화면분할은 처리결과에 수반되는 MC데이터의 재계산 과정을 최소화하여 영상을 처리하는 방법이 있어야 한다. 기존의 연구는 이를 위해 MC-DCT로 인코딩된 영상 데이터를 DCT영역으로 디코딩하고, DCT영역에서 MC데이터의 처리를 수행하여 화면분할에 따르는 복잡성을 일부 개선하였다[3].

Montage는 DCT영역에서의 처리방법을 적용하여 AT&T에서 개발한 다지점 비디오브리지가대기. 이것은 풀 모션 영상의 실시간 처리가 가능하며, 인코더, 디코더를 사용하여 사용자가 독립적으로 개인적인 영상의 구성이나 조작이 가능하다. 그러나 각 단말에서 화면조작과 화면분할의 일부기능을 수행함으로써 사용자단말에 과중한 부하가 걸리며, 처리대상이 표준 압축동영상이 아닌 Moving JPEG (MJPEG)을 사용한다는 한계를 갖고 있다.

2.2.3. VLC영역에서의 화면분할

VLC영역에서의 화면분할 방법은 화소 및 DCT영역에서의 방법에 비해 하드웨어 구성이 용이하고, 처리지연이 최소화된다. 또한 영상의 열화가 발생하지 않으며 알고리즘이 간단하여 적은 계산량으로 처리가 가능하다. 반면에 이 방법은 영상스트림을 전혀 디코딩하지 않고 신택스 레벨에서 처리하기 때문에 영상

조작, 처리의 제한을 갖는다. 그러나 다지점 멀티미디어 통신과 같은 실시간서비스를 제공하기 위해 요구되는 여러 Quality of Service(QoS)조건 중 짧은 처리시간, 영상의 열화 그리고 시스템의 복잡성 등을 고려할 때 실제적으로 적용이 가능한 방법이다.

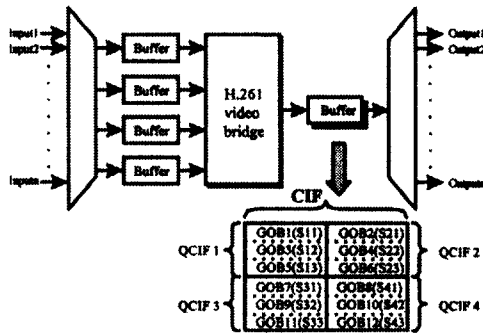


그림1. 완전 압축영역에서의 H.261 화면분할
Fig 1. Split-screen of H.261 on compressed domain

표1. 세가지 방법에 대한 비교
Table 1. Comparison of 3 methods

	픽셀 영역	DCT 영역	VLC 영역
지연	전송지연 화면분할처리 지연 폴 디코딩-인코딩 지연	전송지연 화면분할처리 지연 VLC 디코딩-인코딩 지연	전송지연 화면분할처리 지연
복잡도	높음	중간	단순
영상화질	열화가 많음	중간	열화가 없음
유연성	매우높음	중간	제한적

현재 기 제안된 방법은 H.261 표준 압축코드를 대상으로 Quarter Common Intermediate Format(QCIF)의 해상도가 CIF의 1/4인 점을 이용한다. QCIF와 CIF의 차이점은 상위 두 계층의 헤더 내용이며, 나머지 하위 계층의 구조는 같다. 따라서 4개의 QCIF를 모았을 때 생성되는 12개의 Group of Block(GOB)을 CIF에 맞도록 재구성함으로써 처리가 가능하다[4]. 그림1은 하나의 CIF에 4개의 QCIF를 GOB 단위로 배치하는 형태를 보여준다. 이 방법은 비디오브리지의 화면분할 기능이 신덱스 해석, 조작과 같은 단순한 처리로 가능하며 완전 압축코드영역에서 처리하므로 압축부호율 제어를 위한 송수신 완충버퍼링이 필요없게 되어 실시간 멀티미디어통신에 필요한 단말과 단말간의 지연

에 대한 요구사항을 만족시킬 수 있다. 그러나 현재 제안된 방법은 H.261 동영상만을 대상으로 하며 또 다른 압축표준인 MPEG에 대한 처리방법은 고안되어 있지 못하다. 표1은 지금까지 연구된 화면분할방법에 대한 비교이다. 표1에서 보인 바와 같이 VLC영역에서의 방법이 지연과 화질측면뿐만 아니라 하드웨어를 효율적으로 구성할 수 있는 방법임을 알 수 있다.

H.261은 실시간통신용 비디오 압축표준이고 MPEG-1은 저장미디어용의 비디오압축으로 표준화된 데 반해, MPEG-2는 멀티미디어 서비스의 통합, 제공을 목적으로 한 계층적코딩구조를 갖는 비디오 압축표준으로 전송 비트 에러에 높은 회복력을 가져 고품질 영상통신에 적용이 가능하다. 현재 영상전화와 영상회의와 같은 실시간 대화형통신에는 H.261을 사용한 H.320계열의 단말이 ITU-T에서 표준화되었으나, MPEG-2의 성공에 힘입어 영상전화, 영상회의 단말에 동일한 용도로 MPEG-1, MPEG-2가 사용되고 있다 [2]. 그러나 이와 같은 요구에 비해 현재 MPEG-1 또는 MPEG-2를 완전 압축영역에서 처리하는 방법은 제안되어 있지 못하다. 본 논문은 MPEG-1/2 동영상을 대상으로 하는 새로운 화면분할 처리방법을 제안한다.

III. 새로운 MPEG-1/2 화면분할 처리방법

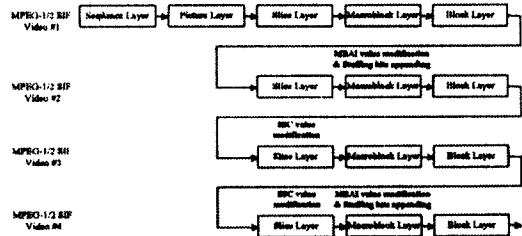


그림 2. MPEG구문의 변경처리
Fig. 2. Modification procedure of MPEG syntax

영상전화, 영상회의와 같은 대화형 통신서비스에서는 멀티미디어신호의 실시간 전달이 통신서비스의 매우 중요한 QoS 조건이 된다. 멀티미디어통신에 사용되는 기본 미디어인 오디오, 비디오 및 데이터 중 데

이터는 비교적 실시간 처리에 여유를 갖으며, 오디오의 경우는 해당되는 코딩방식에 따라 다소 차이는 있으나 수msec이내에서 처리가 가능하여 문제를 야기하지 않는다. 이에 반해 동영상 신호의 경우 압축에 따르는 복잡한 처리와 특히 가변 비트율의 특성을 갖는 출력데이터를 항등비트의 네트워크에 정합하기 위해 사용되는 송수신데이터 완충버퍼링의 결과로 발생하는 지연은 전체 멀티미디어통신의 실시간 QoS를 결정하는 주요인이 되며, 오디오가 Lip-Sync를 위해 비디오와 동기될 때의 처리지연은 반향(Line Echo)의 원인이 된다. 비디오신호의 압축 및 통신에는 최소 동영상 4프레임(4x33msec)에 해당하는 처리지연이 소요되어[8], 대화형통신의 실시간성을 만족하기 위해서는 단말의 설계와 네트워크 정합 및 경로배정에 세심한 배려를 필요로 한다. MCU의 화면분할 기능은 이와 같은 대화형통신의 QoS 특성을 보장하고 그 외 화질 및 구성의 용이성 등이 충분히 검토되어야 한다.

방법은 기존 Bell Lab에서 제안, 개발한 H.261을 대상으로 하는 화면분할 방법과 유사한 처리방식 및 구조를 가지고 있으나[4], H.261과 MPEG은 구체적인 코딩 방법 및 이에 따른 구문체계가 달라서 서로 간의 연동이 쉽지 않다. H.261을 대상으로 하는 기존방법은 QCIF(176x144)를 구성하는 GOB구문의 일부 내용을 변경하여 CIF(352x288)형태의 프레임의 생성하는 방법이다. 이에 반해 본 논문에서 제안한 방법은 MPEG동영상의 화면분할을 위해 슬라이스 단위의 처리를 수행한다. 현재 MCU에서의 화면분할 처리는 그 대상을 CIF급 이상으로 정하고 있다[9,10]. 이는 영상 처리에 따르는 화질저하와 사용자 인터페이스 측면이 고려된 것으로 다수의 영상을 분할처리 후, 동일한 화면에 표현하기 위해서는 최소한 CIF 및 SIF(352x240)급이상의 화질이 필요함을 의미한다.

표 2. MPEG구문상에서 변경되어야 할 부분
Table 2. Modification on MPEG syntax

	MPEG-1/2
Sequenc Header	Horizontal_size Vertical_size Bit rate VBV_buffer_size Constrained_parameters_flag Profile_and_Level_indication(MPEG-2)
Picture Header	Temporal_reference Picture_coding_type
Slice Header	Slice_start_code(#3,#4)
Macroblock Layer	Macroblock_address_increment(#2,#4)

본 논문은 이와 같은 대화형통신의 서비스 조건을 만족할 수 있는 새로운 화면분할 방법을 제안한다. 제안한 방법은 MPEG-1/2 동영상을 대상으로하며 Source Input Format(SIF)과 ITU-R BT601규격 사이의 해상도 차이와 MPEG-2 동영상표준의 후방향 호환(Backward Compatibility) 특성을 이용한다. 본 논문에서 제안한 MPEG 비디오브리지에서 처리하는 영상은 352x240의 해상도를 갖는 MPEG-1 SIF, MPEG-2 MP,SP/LL을 대상으로 하며 입력되는 다수의 동영상 중 4개를 선택하여 ITU-R BT.601규격의 704x480해상도를 갖는 MPEG-1/2동영상을 생성한다. 이 같은

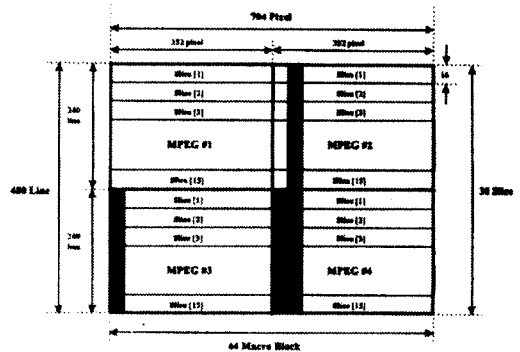


그림3. 분할처리된 MPEG 프레임 구성
Fig.3. Frame configuration for MPEG split-screen

MPEG 동영상을 사용한 멀티미디어 통신단말을 대상으로하는 MPEG 비디오브리지의 개략적인 동작은 다음과 같다. 다지점 영상회의를 위해 단말에서 출력되는 MPEG-1 SIF, MPEG-2 SP,MP/LL신호는 MPEG비디오브리지에 입력되며, 선택된 4개의 압축 동영상은 그림2에 보인 바와 같이 첫번째 MPEG 스트림의 시퀀스 및 픽처체층의 헤더 정보만을 이용하여 4개의 압축 동영상을 하나의 동영상으로 처리한다. 처음 입력된 MPEG스트림은 변화없이 출력되고 이후 입력되는 스트림의 Macroblock Address Increment (MBAI), Slice Start Code(SSC) 값을 변경하

므로써 MPEG-1/2 디코더가 연속된 스트림으로 인식하여 디코딩할 수 있도록 한다. 표2에서 MPEG-1/2 비디오스트림을 분할처리하기 위한 변경내용을 보였고, 그림3은 변경해야 할 해당 슬라이스와 매크로블럭을 포함한 분할처리된 MPEG의 프레임 구성을 나타낸다. MPEG비디오신호의 분할처리를 위해 계층별로 변경할 부분의 상세한 내용은 다음과 같다.

• 시퀀스 계층

352x240으로 지정된 MPEG-1/2 비디오스트림의 해상도를 ITU-R BT.601에 따르는 수평, 수직 해상도인 704x480으로 변경해야 한다. MPEG 수평, 수직 해상도는 변경 전, 후 동일하게 3바이트를 차지한다. 4개의 MPEG-1/2 비디오스트림을 처리하기 위한 비트율과 Video Buffer Verifier (VBV) 버퍼크기는 합성, 처리되는 각각의 MPEG-1/2의 비트율과 VBV버퍼크기의 합이 된다. Constrained Parameter Flag(CPF)는 MPEG-1에서만 사용되며 MPEG-2 비디오스트림에서는 항상 0으로 설정한다.

• 슬라이스 계층

슬라이스 시작코드는 슬라이스의 수직위치를 나타내므로 그림3에서의 #3(화면분할된 화면의 좌측 하단)과 #4(화면분할된 화면의 우측 하단)의 슬라이스 시작코드(SSC)를 15만큼씩 증가시킨다.

• 매크로블럭 계층

MBAI는 슬라이스의 스킵되지 않은 첫번째 매크로블럭의 수평위치를 나타내므로 그림3에서의 #2(화면분할된 화면의 우측 상단)와 #4(화면분할된 화면의 우측 하단)의 각 슬라이스의 첫번째 매크로블럭의 MBAI값을 22만큼씩 증가시킨다. MBAI값은 가변장 부호로 되어 있으므로 22만큼의 값을 증가시킬 경우 비트 수의 변동이 발생할 수 있다. 따라서 바이트 단위의 구조복원을 위해 해당 슬라이스의 마지막 매크로블럭에서 비트 스테핑처리를 수행한다.

그림4에서 MPEG-1/2형태의 4개 입력비디오를 화면분할하는 순서도를 보였고, 그림5는 처리결과로 생성된 분할화면을 표준 MPEG-2 소프트웨어 비디오디코더로 디코딩한 것을 보여준다.

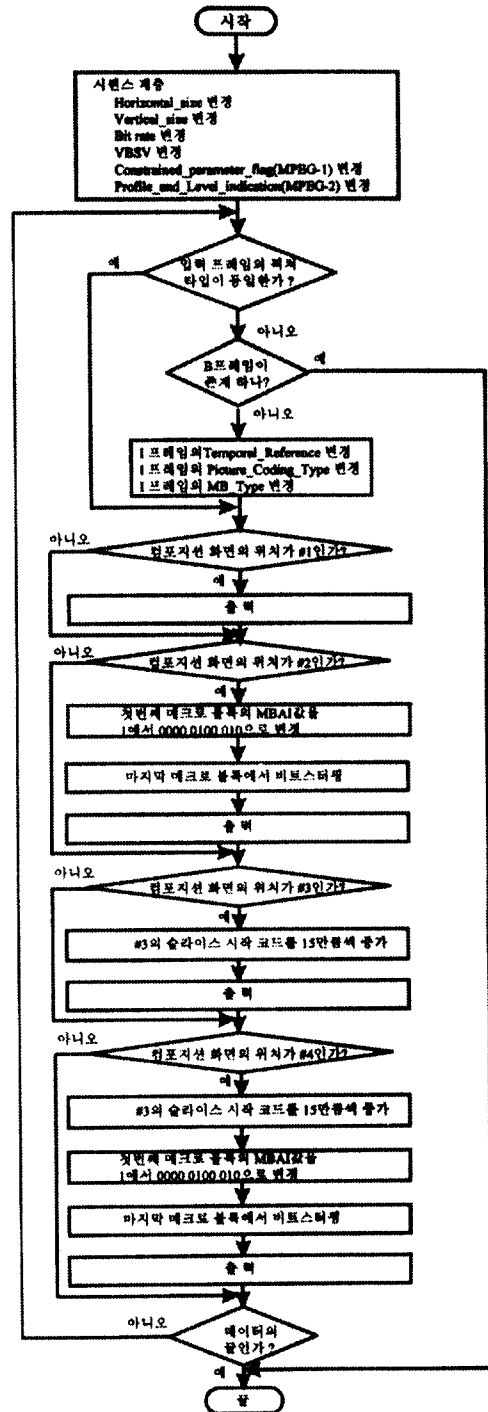


그림4. MPEG-1/2 화면분할 방법의 순서도
Fig.4 Flowchart for MPEG-1/2 split-screen method



그림5. 본 방법에 따라 분할처리된 MPEG화면
Fig.5. Split-screen according to the proposed method

본 논문에서 제안한 새로운 MPEG동영상 화면분할 방법은 앞에서 설명한 바와 같이 완전 압축영역에서 MPEG동영상의 일부 구문내용을 변경함으로써 처리가 가능하다. 이에 따라 동일한 기능을 구현하기 위해 디코더 및 인코더를 다수 사용하였던 기존 방식에 비해 간단한 형태로 구현이 가능하다. 또한 압축영역에서 처리됨에 따라 분할처리에 수반되는 화질의 열화가 전혀 발생하지 않으며, 대화형 통신의 품질을 결정하는 단-대-단 지연조건을 만족시켜줄 수 있다. 그러나 Bell Lab에서 제안된 H.261을 대상으로 하는 기존 방법과 본 방법은 공히 처리 가능한 영상이 4개로 제한된다는 한계를 갖는다. 현재 MCU의 비디오브리지는 최대16지점의 동영상을 분할처리할 수 있도록 권고하고 있다[9]. 그러나 화면분할 처리에 대한 기존의 연구 및 개발결과는 단일화면에 다수의 화면을 표현하는 것이 사용자 정합측면에서 단점이 될 수 있음을 보고하고 있다[11]. 화면분할 처리기능을 이용한 다지점 영상회의는 동일한 화면상에 토론에 참가한 참석자의 영상을 보여, 실제와 유사한 현장감을 부여하기 위한 목적으로 사용된다. 그러나 단일 화면에 일정 수 이상의 참가자를 표시하는 것은 화자의 인지용저하 및 혼동을 야기할 수 있으므로 대략 4-6인의 화면을 표현하는 것이 보편적이다[11,12,13,14,15,16]. 물론 화면의 표시로 인한 문제를 해결하기 위해 화소 영역에서의 융통성 있는 미디어 조작기능을 이용하

여, 회의 주도자를 포함한 특정화자의 영상을 부각시켜 표현하는 것도 가능하다. 그러나 이를 위해 MCU는 특정영상을 선택, 처리하는 기능을 단말의 접속점마다 별도로 가져야 함으로서 실제 구현하기에는 어려움을 갖고 구현이 비교적 간단한 단말의 영상출력부에서 이와 같은 기능을 갖는 것이 일반적이다 [13,14,15].

현재 소프트웨어로 기능의 검증이 완료된 본 방법이 실제 적용되기 위해서는 MPEG스트림의 양자화행렬값과 Group of Picture(GOP)구조가 동일해야 하고, 중간에 시퀀스 헤더가 포함되지 않아야 하며, MPEG-1의 경우는 스트림을 구성하는 슬라이스가 행을 넘지않도록 조절하여야 한다는 기본 전제조건이 있다. 특히 화면분할의 실시간 처리를 위해서는 MPEG스트림의 배치순서에 따르는 입력스트림간의 동기가 확립되어야 한다. 입력스트림의 동기를 제외한 나머지 전제조건은 ITU-T H.320계열의 단말과 STB가 사용하는 H.245 및 Digital Storage Media Command and Control(DSM-CC) 등과 같은 신호절차를 통한 능력교섭에 의해 쉽게 해결이 가능하다. 입력스트림의 동기는 화면분할 처리부의 실시간 특성을 결정하는 주요한 요소로서 정교한 메카니즘이 필요하며 단-대-망, 단-대-단 신호절차를 사용한 동기화 방법을 고려하고 있다. 또한 본 방법은 화면분할의 결과 생성되는 데이터의 양이 입력된 4개의 MPEG스트림을 합한 대략 4배의 값이 되므로 비대칭적인 통신대역을 필요로 한다. 그러나 ATM을 근간으로 하는 광대역 ISDN은 송신과 수신에 통신대역을 비대칭적으로 할당하는 것이 가능하므로 본 방법이 전혀 문제없이 적용될 수 있다.

새롭게 제안한 화면분할 방법은 다지점 멀티미디어통신을 위한 응용 이외에 원격감시시스템 및 CATV 및 VoD등에 사용되는 Preview화면을 구성하는데 사용될 수 있다. 원격감시는 현재 VoD용으로 개발이 완료된 MPEG-2 디코더를 내장한 STB에 MPEG화면분할 처리기능을 부가하고 MPEG-1 인코딩기능을 갖는 감시지역의 단말과 접속을 통해, 4지점의 상황을 모니터링하는 시스템을 효율적으로 구성할 수 있다. VoD용의 Preview화면의 구성은 이 용도로 스트림을 특정하게 인코딩하거나, 기존 인코딩된

동영상을 화소영역으로 변환한 후 제작할 수 있으나, 본 방식을 적용하면 MPEG-1 SIF 및 MPEG-2 SP,MP/LL로 인코딩된 압축 동영상의 직접 처리가 가능하므로 저장용 이외의 Preview화면의 실시간 편집이 가능하다.

IV. 실험 및 고찰

기존의 방법과 본 연구에서 제안한 방법의 성능 비교 평가를 위해 실시간 다지점 영상회의를 제공하기 위한 비디오브리지의 구성방법을 단순화시킨 그림 7과 같은 4가지 형태를 고려하였다. 첫번째는 본 논문에서 제안한 방식으로 비디오브리지에서 데이터의 단순한 조작을 통해 화면분할을 행하는 경우이다. 두번째는 화소영역에서 352x240의 해상도를 가지는 비디오소스를 분할처리하여 704x480의 해상도를 가지는 비디오로 만들기 위해 MCU내의 비디오브리지에서 디코딩과 인코딩 과정을 수행하는 경우이다. 세번째는 704x480의 해상도를 가지는 비디오소스를 화소영역에서 스케일링하는 형태이고, 네번째는 704x480의 해상도를 가지는 비디오소스를 DCT 영역에서 스케일링하는 경우이다. 제안한 방식이 실시간 다지점 영상회의를 제공하기 위한 방법으로 합당하다는 것을 보이기 위해 4가지 방법의 화질저하와 하드웨어의 단순

성, 화면분할 처리에 수반되는 지연 등을 평가한다. 입력력은 704x480의 해상도를 갖는 MPEG-2 테스트용 스트림을 용도에 맞게 처리하였다.

4.1. 지연

그림6에서 보였듯이 본 논문에서 제안한 방법1의 화면분할 처리는 입력 비디오스트림을 슬라이스 단위로 처리한 후 바로 출력시키기 때문에 매우 작은 지연만이 발생한다. 반면에 방법2,3은 화면분할 처리를 위해 MPEG 디코더와 인코더를 사용한 경우를 나타낸 것으로 이 방법은 인코딩에 수반되는 처리지연과 압축동영상을 고정된 비트율로 전송하기 위한 송신완충 버퍼링 지연을 포함한 인코더부의 지연이 누적되며, 영상 스케일링 및 배치처리에 최소한 한 필드(1/59.94Hz)의 지연이 소요된다. 그 외 디코더부의 지연이 포함되는데 디코더부는 인코더에서의 VBV값에 따라 디코더가 입력스트림을 완충하는 vbv_delay값과 순수 디코딩 처리 지연을 포함한다. 방식4는 부분적인 디코딩, 인코딩 처리과정을 수행함으로써 방식2,3보다 지연을 감소시킬 수 있지만, 여전히 방식1보다 많은 지연을 초래한다. 이와 같은 결과는 화면분할 장치의 구현방법과 사용되는 인코더, 디코더의 성능에 따라 다소 다른 결과를 낼 수 있으나, 본 논문에서 제안한 슬라이스 단위의 처리방법은 입력되는 스트림간의 동기가 확립되어 있을 경우, 최대 하나의 슬라이스 정보를 완충하기 위한 수신 버퍼링 지연이 발생할 뿐으로 실시간 통신에 적합한 방법임을 알 수 있다.

4.2. 구현의 용이성

제안한 방식은 스트림 해석 및 처리기능을 포함한 간단한 로직으로 구현이 가능하나 DCT영역에서의 처리방법을 제외한 다른 방식의 화면분할은 완전한 디코딩과 인코딩 처리를 필요로 한다. 화소영역에서의 방법은 인코딩, 스케일링, 및 디코딩 과정이 필요하다. 동영상의 디코딩, 인코딩은 각각 VLD, 역양자화, IDCT 및 움직임 보상 그리고 움직임 추정, DCT, 양자화, VLC등을 포함하는 복잡한 신호처리 과정을 필요로 하며 영상의 스케일링과 재배치 과정 역시 이에 포함된다. 이때 전체 처리과정 중 가장 복잡한 절차는

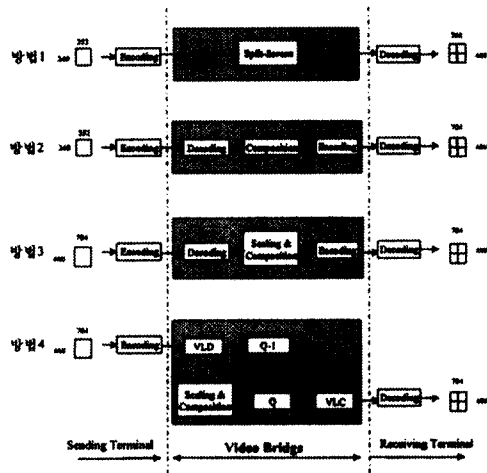


그림6. 비디오 처리부의 4가지 방식
Fig.6. 4 methods of video processing module

움직임 추정 및 보상처리부로 전체 처리과정의 70% 이상을 차지한다. 본 논문에서 제안한 방법은 화소영역의 영상 스케일링 및 재배치에 필요한 로직보다 단순한 구성으로 구현이 가능하여 실제 화소영역에서의 방법과는 비교 대상이 될 수 없으나 화소영역 및 DCT영역에서 동일하게 사용되는 일부 처리기능에 해당하는 DCT, IDCT와 MC-DCT에 소요되는 과정을 기술하여 본 방법이 갖는 단순성을 설명한다.

인코딩 및 디코딩에 필요한 DCT와 IDCT 과정의 픽셀당 곱셈수와 픽셀당 덧셈수는 다음 식1-1, 1-2와 같다[17].

$$2\log_2 N(3 + 8/N) \quad (1-1)$$

$$3(\log_2 N - 1) + 4/N \quad (1-2)$$

여기서 N은 이미지 블럭당 한번의 픽셀수이다. 스케일링을 하기 위한 행렬의 곱은 A(P/B로 나타낸다. 여기에서 A와 B는 NxN 행렬이고, P는 N픽셀의 이미지블럭이다. 계산과정은 일반적으로 2(N3)의 곱셈이 필요하지만 스케일링을 하기 위한 행렬계산은 A와B 행렬 중 8개의 성분을 빼고, 나머지는 0이기 때문에 블럭당 128개의 곱셈이 필요하다. 704x480의 해상도를 가지는 비디오소스는 최대 한 프레임 당 총 5280개의 블럭이 있으므로 프레임 당 675840의 곱셈이 필요하다. DCT영역에서의 화면분할 방법은 행렬값만 다를 뿐 스케일링방법이 화소영역에서의 처리방법과 동일하며 계산의 복잡도가 낮고, 인코딩, 디코딩 과정에서 DCT와 역DCT과정이 없다. 그러나 MC-DCT(DCT영역에서 MC)와 역MC-DCT과정이 별도로 필요하다. 이러한 과정은 다음에 나타낸 식2-1, 2-2와 같은 픽셀 당 곱셈과 덧셈이 필요하다[6].

$$(4/\beta + \sqrt{\beta}) \cdot N \cdot \alpha_2 + (2/\beta) \cdot \alpha_1 \quad (2-1)$$

$$(4/\beta + 2/\sqrt{\beta}) \cdot N \cdot \alpha_2 + (2/\beta)N + 1) \cdot \alpha_1 + 1 \quad (2-2)$$

여기서 β 는 입력 이미지의 압축률이고 α_1 은 단방향 블럭경계 조정을 필요로 하는 이미지 블럭을 나타내며, α_2 는 양방향 블럭경계 조정을 필요로 하는 이미지 블럭의 비율이다.

이와 같이 화면분할 처리를 위해 복잡한 처리과정

이 수반되는 기존방식과 비교하여 제한한 압축영역에서의 MPEG화면분할 방법은 스트림의 디코딩, 인코딩과정이 완전히 배제된 구문의 파라미터 변경과 비트 조작처리만이 필요하다. 이것은 입력버퍼에 저장된 각각의 MPEG비디오 Element Stream(ES)의 구문에서 변화시킬 값을 찾아 조작, 처리하고 필요한 비트 스테핑 기능을 수행한 후, 공통버퍼에 출력하는 기능이다. 따라서 제안한 방식은 대상 스트림의 탐색과 비트처리를 위한 기능만을 갖는 단순한 로직으로 구현이 가능하다.

4.3. 화질

그림 7은 4가지 방식으로 테스트 영상을 처리한 후의 화면을 보여 준다. 화면분할은 동일한 영상을 4개의 개별 입력으로 사용하였고 그림은 이중 하나의 화면만을 나타낸 것이다. 정확한 화질평가를 위해 다음에 나타낸 식 3-1과 3-2에 의해 PSNR을 측정하고 비교해 보았다. 표3에 4가지 방법에 따르는 PSNR 측정값과 제안한 방식과 다른 3가지 방식의 평균PSNR 차이를 나타내었다.

$$PSNR = 10\log_{10} 255^2 / MSE \quad (3-1)$$

$$MSE = 1/N \sum_{i=1}^N (X_i - \hat{X}_i)^2 \quad (3-2)$$

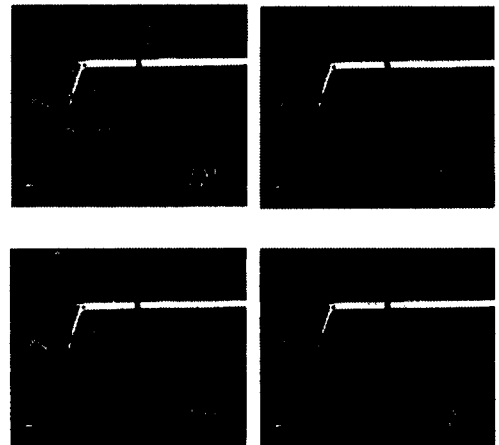


그림7. 각 방법에 따라 분할처리된 화면
Fig.7. Processed image according to the each method

표 3. 4가지 방식의 평균 PSNR값
Table 3. Mean PSNR for 4 methods

	PSNR	PSNR 차이
방법1	37.25dB	-
방법2	35.38dB	-1.87dB
방법3	34.90dB	-2.45dB
방법4	34.38dB	-2.87dB

표3과 그림8에 보였듯이 본 논문에서 제안한 화면분할 방법은 완전압축 영역의 처리이므로 영상의 화질은 화면분할 처리부를 배제한 것과 같이 송신단말의 압축영상을 수신단말에서 그대로 디코딩한 결과와 동일하게 전혀 화질저하를 유발하지 않는다. 반면에 화소영역 및 DCT영역에서의 처리는 영상의 scaling에 따르는 열화와 양자화, 역양자화 과정의 반복에 따르는 영상의 열화 누적이 발생한다. 이에 따라 영상의 화질은 본 논문에서 제안한 방법, 화소영역에서의 처리방법, DCT영역에서의 처리방법순으로 열화의 정도가 심해짐을 알 수 있다.

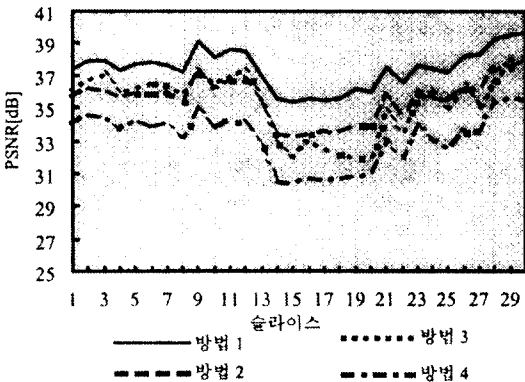


그림8. PSNR 비교
Fig.8. Comparison of PSNR

V. 결 론

회의에 참가한 참석자의 분할처리된 동영상은 분배하여, 보다 현장감 있고 실제 상황과 유사한 환경을 제공하는 MCU의 비디오브리지는 단순히 동영상을

처리하는 기능 및 방법 이외에 다양한 관점의 고려가 필요하다. 다지점 통신에 사용되는 MCU는 단말 및 네트워크에서 제공할 수 없는 유용한 서비스를 제공하는 서비스노드로 볼 수 있으며 사용자는 이를 위해 회선사용료와는 별도의 요금을 지불한다. 따라서 MCU는 사용자에게 대화형 다지점 통신서비스를 위한 최적의 환경을 제공하여 주어야 한다. 이 중 단방향 및 왕복지연을 포함하는 단-대-단 지연은 대화형 멀티미디어통신의 품질을 결정하는 매우 중요한 요소로 MCU의 미디어처리는 이와 같은 요구조건을 필히 만족하여야 한다.

본 논문은 실시간측면의 QoS를 만족하는 새로운 MPEG 화면분할 방법을 제안하였다. 제안한 방법은 화면의 처리가 4개로 국한되고, 실제 적용하기 위해서는 일부 전제조건을 만족해야 하는 제한을 갖고 있으나 MPEG 동영상을 사용한 다지점통신에 직접 적용이 가능한 효율적인 방법이다. 제안한 방법은 현재 FPGA를 사용하여 구현하고 있으며, 화면분할의 실시간 처리에 중요한 요소인 스트림간의 동기방법과 MPEG-1 및 MPEG-2의 영상이 동시에 입력될 때의 처리방법에 대한 연구도 병행하여 추진하고 있다. 현재 대화형 멀티미디어 통신서비스는 미디어의 압축표준 및 시스템을 표준화하고 있는 ITU-T와 ISO/IEC에서 권고하고 있는 내용에 따르고 있다. 따라서 사용자에게 직접 적용되는 통신단말의 형상은 요구되는 서비스 및 효용성 등이 고려된 다양한 형태가 될 것이다. ITU-T에서는 이와 같은 흐름의 결과로 발생하는 이종 멀티미디어 통신단말간의 연동을 주요항목으로 연구하고 있다. 본 논문에서는 ITU-T의 H.320계열 단말과 연계가 불가능한 MPEG을 대상을 한 MCU의 비디오브리지에 대해 기술하였다. 이것은 멀티미디어 통신서비스의 연동에 문제를 야기시킬 수 있으나 현재 이와 같은 단말 및 시스템의 개발 및 구현은 필연적이며[2], 향후 H.261, MPEG등의 이종 표준 동영상과 다양한 유형의 오디오신호 및 데이터 프로토콜이 복합적으로 사용될 멀티미디어 통신단말간의 연동 및 연계에 대한 연구가 단말과 네트워크를 포함한 전반적인 시스템 차원에서 이루어져야 할 것이다. 특히 동영상의 경우는 국제표준기구의 두 축인 ITU-T와 ISO/IEC에서 권고하고 있는 H.261, H.263, 그리고

MPEG-1/2간의 효율적인 변환(Transcoding) 방법에 대한 연구가 이루어져야 할 것이다.

참 고 문 헌

1. Digital Audio-Visual Council, "DAVIC 1.2 Specifications Baseline Document", New York 13th Meeting, June 1996.
2. 황대환 외, "광대역ISDN상의 다지점 멀티미디어통신서비스 1부: DAVIC 표준STB에서의 대화형 멀티미디어통신", 한국통신학회지, pp. 821~835, 1998. 4.
3. Shin-Fu Chang, David G. Messerschmit, "Manipulation and Compositing of MC-DCT Compressed Video", IEEE Journal on Selected Areas in Communication, Vol.13, No.1, pp. 1~11, Jan. 1995.
4. Shaw-Min Lei, Ting-Chang Chen, Ming-Ting Sun, "Video Bridging Based on H.261 Standard", IEEE Transactions on Circuits and Systems for Video Technology, Vol.4, No.4, pp. 425~437, Aug. 1994.
5. Willebeek-Lemair, D. D. Kandlur, and Z. Y. Shae, "On Multipoint Control Units for Videoconferencing", Local Computer Networks '94 Proceeding, pp. 356~364.
6. Michael E. Lukacs and David G. Boyer, "A Universal Broadband Multipoint Teleconferencing Service for the 21st Century", IEEE Communications Magazine, pp. 36~43, Nov. 1995.
7. Robert D. Gaglianella, Glenn L. Cash, "Montage : Continuous Presence Teleconferencing Utilizing Compressed Domain video Bridging", ICC' 95, pp. 573~581.
8. Okubo, "ITU-T Standardization of Audiovisual Communication Systems in ATM and LAN Environments", IEEE Journal on Selected Areas in Communications, pp. 965~982, Aug. 1997.
9. Multipoint Control Units for Audiovisual Systems using Digital Channels up to 2Mbits/s, ITU-T Rec H.231, 1996.
10. Procedures for establishing communication between three or more audiovisual terminals using digital channels up to 2Mbit/s, ITU-T Rec H.243, 1996.
11. John Robinson, "A Multimedia Interactive Conferencing Application for Personal Workstations", IEEE Transactions on Communication Vol39, No.11, pp. 1698~1708, Nov. 1991.
12. Christoph Weiss, "Desktop Video Conferencing - An Important Feature of Future Visual Communications", ICC' 90, pp. 134~139.
13. Kazutoshi MAENO, "Distributed Desktop Conferencing System (MERMAID) Based on Group Communication Architecture", ICC' 91, pp. 520~525.
14. Hiroya TANIGAWA, "Personal Multimedia Multipoint Teleconference System", INFOCOM' 91, pp. 1127~1134.
15. Kazuo Watabe, "Distributed Desktop Conferencing System with Multiuser Multimedia Interface", IEEE Journal on Selected Areas in Communications, Vol9, No.4 pp. 531~539, May 1991.
16. Mon Song Chen "A Multimedia Desktop Collaboration System", GLOBECOM' 92, pp. 739~746.
17. Shin-Fu Chang, David G. Messerschmit, "A New Approach to Decoding and Compositing Motion-Compensated DCT-Based Images", IEEE ICASSP' 93, pp. 421~424.
18. Information Technology - Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5Mbit/s- Part2: Video, ISO/IEC 11172-2, 1993.
19. Information Technology - Generic Coding of Moving pictures and Associated Audio Information - Video, ITU-T Rec. H.262/ISO/IEC 13818-2, 1995.



박 정 호(Jung Ho Park) 정회원

1985년 2월 : 성균관대학교 전자
공학과(공학사)

1987년 2월 : 성균관대학교 전자
공학과(공학석사)

1997년 2월 : 성균관대학교 전자
공학과 박사과정
수료

1987년 1월~1992년 6월 : 삼성종합기술원 전자기기
연구소(주임연구원)

1992년 7월~1994년 1월 : (주)인성정보 부설연구소
근무

1997년 3월~현재: 데이텍주식회사 부설연구소장

*주관심분야: 멀티미디어 통신, 분산 멀티미디어 시
스템

황 대 환(Dae Hwan Hwang) 정회원

1986년 2월 : 성균관대학교 전자공학과(공학사)

1988년 2월 : 성균관대학교 전자공학과(공학석사)

1995년 3월~현재 : 성균관대학교 전자공학과
(박사 과정)

1991년 2월~현재 : 한국전자통신연구원(선임연구원)

*주관심분야: 디지털 통신시스템, 멀티미디어 통신

이 중 형(Jong Hyeong Lee) 정회원

1981년 2월 : 충남대학교 전자교육공학과(공학사)

1987년 2월 : 숭실대학교 전자공학과(공학석사)

1996년 8월 : 충남대학교 컴퓨터공학과(공학박사)

1988년 2월~현재 : 한국전자통신연구원 통신망구조연
구실 근무

*주관심분야: 디지털 통신시스템, 멀티미디어 통신



구 한 준(Han-Jun Ku) 정회원

1996년 2월 성균관대학교 전자
공학과(공학사)

1998년 2월 성균관대학교 전자
공학과(공학석사)

1998년 3월 ~ 현재: LG정밀 기술
전략팀 연구원

*주관심분야: 멀티미디어 통신, 이동통신

조 규 섭(Kyu Seob Cho)

정회원

1974년 2월 : 성균관대학교 전자공학과(공학사)

1976년 2월 : 성균관대학교 전자공학과(공학석사)

1989년 2월 : 성균관대학교 전자공학과(공학박사)

1977년 3월~1992년 2월 : 한국전자통신연구원(책임연
구원)

1992년 3월~현재: 성균관대학교 전기, 전자 및 컴퓨터
공학부 교수