

멀티미디어 기반의 원격 이동 로봇 제어 시스템

변재영, 문호석, 정제한, 고성제
고려대학교 전자컴퓨터공학과
전화 : 02-3290-3228 / 팩스 : 02-921-0544

Remote Mobile robot control system using multimedia data

Jae-Young Pyun, Ho-Seok Moon, Jae-Han Jung, Sung-Jea Ko
Dept. of Electronics Engineering, Korea University, Seoul, Korea
E-mail : sjko@dali.korea.ac.kr

Abstract

This paper presents a remote mobile robot system that transmits streaming video and audio over the lossy packet networks such as (Wireless) LAN. The error resilient video and audio packets are transmitted on the RTP/UDP/IP protocol stack. The mobile robot can be accessed by a certified user from the remotied area. Thus, the movement of mobile robot can be controlled by the operator observing the working surroundings.

멀티미디어 트래픽을 유·무선 환경에서 통신 채널과 무관하게 그리고 실시간적으로 지원할 수 있다.

다음절에서 이 시스템을 구성하는 로봇의 움직임 제어부와 압축 코덱인 MPEG-4^[1,2]와 G.723.1^[3] 그리고 데이터 전송부에 대해서 기술하고, III절에서 영상과 음성의 동기화를 설명한다. 마지막 절에서는 시스템의 구현에 대해 요약한다.

I. 서론

현재의 유무선 통신 산업의 급속한 발전과 멀티미디어 서비스의 증가는 이동 로봇에 의한 용도의 범위를 한층 증가시키고 있다. 이동 로봇은 home security, 산업 현장에서의 자재 운반, 유해 환경에서의 작업등을 수행할 수 있으며, 이러한 로봇의 기능을 구현하기 위해서는 작업 환경의 영상과 음성을 실시간으로 압축 및 전송할 수 있는 시스템이 필요하다.

본 논문은 무선 환경에 있는 이동 로봇의 움직임과 정밀 작업을 제어하기 위한 원격 제어 시스템 개발에 관한 내용을 기술하였다. 개발된 이동 로봇은 802.11기반의 WLAN 환경에서 멀티미디어 서버로 구현되며, 원격지의 PC기반 클라이언트는 유선 환경에 위치한다. 이러한 애플리케이션은 WLAN이 TCP와 UDP 이터넷 프로토콜을 지원할 뿐만 아니라 최대 11Mbps의 고속 전송율을 보장하고 있으므로, 넓은 대역폭을 요구하는

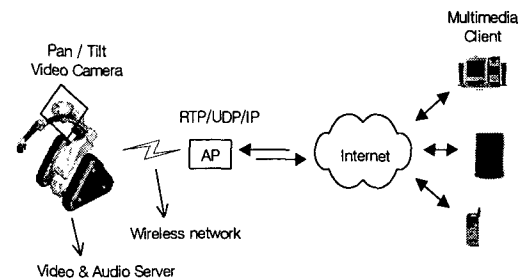


그림 1. 시스템 개요도

II. 시스템 구성

이동 로봇은 로봇의 이동을 제어하기 위한 움직임 제어부와 주위의 장면 및 소리를 캡처하고 압축하기 위한 압축 코덱, 그리고 압축된 멀티미디어 데이터를 전송하기 위한 전송부로 크게 구성된다.

2.1 이동 로봇의 움직임 제어부

(1) CPU Board

CPU board는 영상과 음성 프레임을 캡처하고 획득한 데이터를 압축하기 위한 기능을 수행하며, 전송까지 실시간으로 할 수 있어야 한다. 따라서 고속의

CPU board를 필요로 하며, 본 시스템은 INTEL pentium-4 기반의 mini-board를 사용한다.

(2) DSP Board

DSP board는 본 시스템에서 몸체를 제어하기 위한 DSP 2-Axis DC motor Controller와 Arm System의 4·5축을 제어하기 위한 동일 보드, 그리고 1·2·3축을 제어하기 위한 DSP 3-Axis DC motor Controller 보드가 포함된다.

(3) Arm-System (Zini-5)

로봇 팔이 큰 payload를 요구하지 않으며, 간단한 물체만을 이동시킨다고 가정 하에, 4W 급의 모터를 이용한다. 모터 모델은 다음과 같다.

- Motor: 2224R012SR (경격 12V, 4W)
- Gear: 22/5 (29.6:1, 161:1, 879:1 세 종류를 사용)
- Encoder: IE2-64, 16

(4) Charge System

이동체의 움직임을 위해서는 충전장치를 로봇에 탑재해야 한다. 최소한의 시간으로 최대의 충전효과를 볼 수 있도록 하기 위해 로봇의 제어부분은 저전력 회로로 설계되었다. 전용 충전 장치는 로봇이 어디서나 도킹 가능하도록 3방향이 Open된 형태를 가진다.

2.2 멀티미디어 압축 코덱

(1) MPEG-4 Video 코덱

기존의 MPEG-1과 MPEG-2 표준 방식에서는 입력 되는 영상에 담긴 내용과는 무관하게 영상의 화소 값만을 기본으로 데이터 압축을 수행한다. 그러나 영상에 포함되어 있는 내용을 고려하지 않은 기존의 압축 방법은 영상의 화소 값을 직접 처리하므로 그 성능에 한계가 있다. 따라서 MPEG-4 표준은 단순히 블록 단위의 부호화를 수행했던 기존의 표준들과는 달리, 영상 분할을 통해 객체 별 코딩을 통해 응용 범위의 확대에 초점을 두고 있다. 이러한 압축 방법과 영상 분할 및 확정성 등의 성격을 profile 별로 분류하였으며, 본 시스템은 자연영상의 압축만을 고려한 simple profile (SP)을 사용한다.

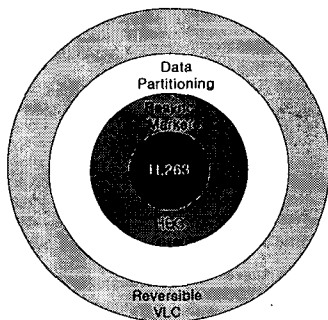


그림 2. MPEG-4 simple profile.

그림 2는 MPEG-4의 simple profile을 보이고 것으로서 H.263과 유사한 압축 방식을 갖고 있으며, 무선 환경에서의 에러 강인성과 은닉성을 위해서 data partitioning, resync markers, HEC, 그리고 RVLC를 사용하고 있다.

(2) G.723.1 Audio Codec

G.723.1은 낮은 비트율의 멀티미디어 통신에서 음성이나 다른 소리의 압축에 쓰이는 코딩된 표현을 제시하고 있다. 이 부호화기는 제한되어 있는 복잡성 내에서 높은 음질을 갖는 음성 신호를 표현할 수 있도록 최적화 되었다. 이 부호화기는 음성 신호를 위해 최적화 되어 있으나 음악이나 다른 소리들도 음성만큼은 아니지만 이 부호화기를 사용하여 압축하고 풀 수 있다.

[비트율]

이 부호화기에는 5.3kbps와 6.3kbps 두 종류의 비트율이 있으며, 6.3kbps로 압축 및 코딩된 음질이 상대적으로 5.3kbps의 경우보다 우수하다. 하지만 5.3kbps의 비트율은 낮은 비트율을 사용하므로 설계자에게 전송율 절약이라는 유연성을 준다는 장점이 있다.

2.3 멀티미디어 전송부

(Wireless) LAN 환경에서 압축된 멀티미디어 정보를 원격지에 위치한 클라이언트에게 패킷화하여 전송하기 위해서는 일반적으로 UDP/IP 프로토콜 스택을 사용한다. 그러나 UDP 프로토콜의 특성상 패킷 손실 여부를 확인할 수 가 없으며, 타임스탬프(timestamp) 정보가 없기 때문에 RTP 레이어를 추가적으로 사용해야 한다^[4]. 그림 3은 RTP/UDP/IP 프로토콜 스택 구조를 갖는 MPEG-4 비디오 헤더의 패킷타이징 구조를 나타낸다. 본 시스템에서는 MPEG-4 simple profile 을 Video 코덱으로 사용하고 디폴트 Maximum Transfer Unit (MTU) 크기인 576 바이트와 패킷 헤더 오버헤드 50 byte를 고려하여 패킷의 최대 크기를 526 bytes로 하였다.

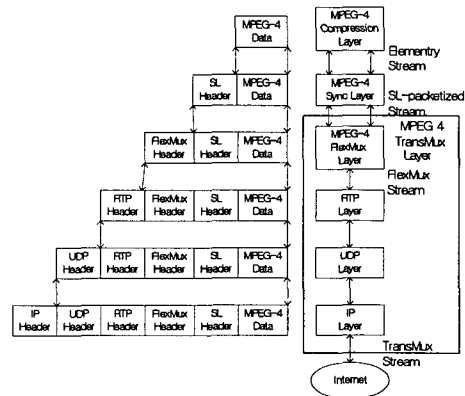


그림 5 . 각 레이어 에서의 데이터 헤더 포맷.

(1) RTP (real-time transport protocol)

RTP는 영상, 음성과 같은 지연에 민감한 데이터를 전송하는 응용서비스에 알맞으며 UDP에서 제공하지 못하는 부가적인 정보를 전송하는 기능을 갖고 있다. 특히, Sequence number (SN)와 타임스탬프는 영상과 음성의 동기화를 위해서 꼭 필요하며 패킷 손실 확인에도 필수적이다.

(2) RTCP (real-time transport control protocol)

RTP의 데이터 전송기능에 제어기능을 추가하기 위해 RTCP라는 제어 프로토콜이 데이터의 전달상황을 감시하며, 최소한의 제어 기능과 매체 식별 기능을 제공한다.

III. 영상과 음성의 동기화

개발한 시스템에서는 영상 압축을 위해 MPEG-4 SP를 그리고 음성 압축을 위해 G.723.1을 사용하고 있다. 압축 영상은 I Video object plane (VOP)와 P VOP로 구성되며, I와 P의 비율을 1:30으로 고정하여 사용한다. 또한 시스템의 영상 압축은 고정 quantization parameter (QP)의 사용으로 VBR 인코딩 특성을 갖는 반면 음성은 5.3kbps의 CBR 인코딩을 한다. 이때 영상 패킷과 음성 패킷은 전송 도중에 겪는 지연과 손실 상태가 서로 다르기 때문에 영상과 음성의 동기화는 필수적이다.

그림 4와 그림 5는 각각 server(이동 로봇)와 client(사용자)의 블록도를 나타내고 있으며, 각 step에서의 과정은 다음과 같다.

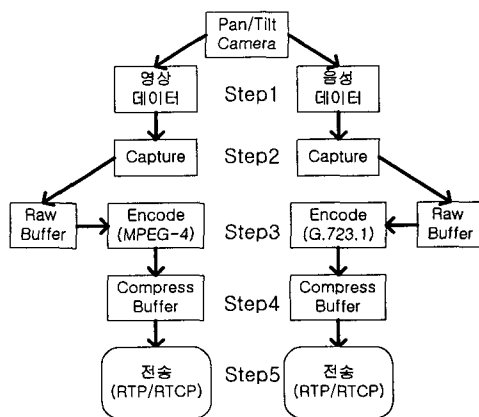


그림 6. Server의 기능별 구성.

3.1 Server

(1) Step 1

음성과 영상을 캡처하는 단계이다. 시스템에서 사용하는 카메라는 pan/tilt 카메라이며 이것은 원격지에

있는 사용자로부터 제어신호에 따라 이동 로봇뿐만 아니라 카메라를 이동함으로써 움직임의 자율성을 높일 수 있다.

(2) Step 2

Pan/tilt 카메라는 영상을 10 FPS로 캡처하고 음성을 초당 10 프레임으로 캡처하여 각각의 raw buffer에 저장 한다. 이때의 raw buffer는 linked-list 방식의 queue를 이용하여 프레임 skipping 기능을 수행할 수도 있다.

(3) Step 3

Raw buffer에 있는 영상 정보는 MPEG-4 SP를 이용하고 음성 정보는 G.723.1을 이용하며, 각각 프레임 단위로 압축을 하게 된다.

(4) Step 4

압축된 정보는 linked-list 형태로 구성된 compress buffer에 저장된다. Compress buffer는 전송부에서 패킷을 전송하기 전에 forward error correction (FEC)이나 RTP 패킷타이징을 할 수 있도록 하기 위해 필요하다.

(5) Step 5

압축된 영상과 음성 데이터에 RTP 헤더를 추가하고 RTCP 패킷을 생성한다. 또한 패킷 전송로의 종단에서 돌아온 RTCP의 패킷을 분석하여 접속자 정보와 트래픽 상태를 파악한다. 또한 클라이언트로부터 전송물 재설정 신호와 카메라 및 로봇의 제어 신호를 수신받는다.

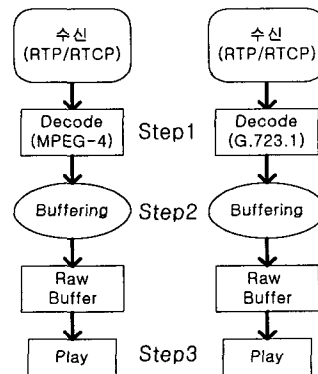


그림 7. Client 기능별 구성.

3.2 Client

(1) Step 1

전송된 음성과 영상은 모두 RTP 헤더를 갖고 있으며, RTP 헤더의 synchronization source (SSRC) ID와 payload type (PT)을 통해 서로 구분할 수 있다. 그리고 구분된 음성 패킷과 영상 패킷은 1개 프레임의 끝을 나타내는 marker(M) 필드를 통해 프레임단위의 디

코딩을 할 수 있다.

(2) Step 2

디코딩된 영상과 음성 정보는 linked-list 형태의 큐에 저장되며, RTP의 SN을 통해 패킷 손실 유무를 확인하고 패킷의 전송 지연에 의한 폐기 유무를 결정한다. 즉, 전송 지연의 한계치를 넘어서는 패킷은 폐기하며, 정상적으로 수신된 패킷은 타임 스탬프로 얻어진 시간 정보에 의해 플레이 시간이 결정되어진다.

(3) Step 3

동기화된 영상/음성 데이터를 플레이하고 사용자가 카메라와 로봇 움직임의 제어신호를 Server에 보내게 된다.

Chiang, Y.Q. Zhang, and H.J.Chao, "On end-to-end architecture for transporting MPEG-4 video over the Internet," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 10, no. 6, pp. 923-941 Sept. 2000.

[3] ITU-T Recommendation G.723.1, "Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s," Mar. 1996.

[4] H. Schulzrinne S. Casner R. Frederick V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications," *IETF RFC 1889*, Jan. 1996

IV. 시스템 구현

MPEG-4 SP 코덱은 유럽 ACTS 프로젝트로 개발된 MoMuSys을 이용하였으며, 음성 코덱으로는 ITU-T G.723.1 Speech Coder인 Ver. 5을 사용하였다. 이동 로봇에서 사용하는 영상 크기는 여러 채널 환경을 고려하여 QCIF(176×144)를 사용하였으며, 음성과 영상의 초당 발생되는 프레임 수는 10 FPS로 하였다. 그림 6은 개발된 모바일 로봇의 사진으로써 이동체의 바퀴, 로봇 팔, 그리고 pan/tilt카메라를 보이고 있다.

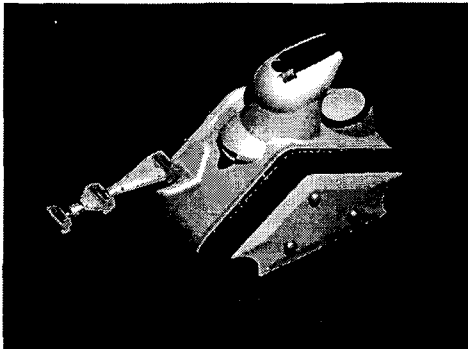


그림 8. 원격 이동 로봇.

Reference

- [1] M. Budagavi, W.R. Heinzelman, J. Webb, R. Talluri, "Wireless MPEG-4 video communication on DSP chips," *IEEE Signal Processing Magazine*, vol. 17, no. 1, pp. 36-53 Jan. 2000
- [2] D. Wu, Y.T. Hou, W. Zhu, H.J. Lee, T.