

# 홈오토메이션을 위한 영상/로봇제어 시스템의 설계와 구현

## Design and Implementation of Visual/Control Communication Protocol for Home Automated Robot Interaction and Control

조 명 지\*                      김 성 환\*\*  
Myung-Ji Cho                Seong-Whan Kim

### 요 약

기존의 음성통신 위주의 공중전화망 (PSTN) 시스템은 TCP/IP 프로토콜을 사용하는 데이터통신 위주의 IP 망으로 진화하고 있으며, IP 망을 이용하여, 멀티미디어 서비스를 제공할 수 있는 VoIP (voice over IP) 서비스를 제공한다. 본 논문은 원격지의 사용자에게 해당 가정의 로봇이 제공하는 VoIP 서비스를 사용하여, 홈오토메이션 기능을 제공하기 위한 시스템을 제안한다. 본 논문에서 구현한 시스템은, VoIP 서비스를 위한 대표적인 시그널링 표준인 HTTP (hyper-text transfer protocol) 기반의 SIP (session initiation protocol) 를 이용하여, 시그널링 프로토콜을 구성하였으며, 원격사용자와 로봇 내 장착된 카메라를 통해 가정의 환경을 모니터링하기 위한 RTP (real-time transport protocol)과 홈오토메이션 신호전달을 위한 DTMF (dual tone multi-frequency) 신호 전송프로토콜을 제공한다. 본 논문에서 제안한 시스템을 실제 구현하여 로봇의 통신모듈로 구성하였으며, KT 망을 통한 영상통화 및 제어신호 전달품질측정결과, 대부분의 트래픽환경에도 원활한 로봇제어 및 평균 MOS 9.15 이상의 우수한 영상품질을 보임을 확인하였다.

### ABSTRACT

PSTN (public switched telephone network) provides voice communication service, whereas IP network provides data oriented service, and we can use IP network for multimedia transport service (e.g. voice over IP service) with economic price. In this paper, we propose RoIP (robot on IP) service scenario, signaling call flow, and implementation to provide home automation and monitoring service for remote site users. In our scheme, we used an extended SIP (session initiation protocol) for signaling protocol between remote site users and home robots. For our bearer transport control, we implemented H.263 video codec over RTP (real-time transport protocol) and additionally DTMF (dual tone multi-frequency) transport for robot actuator control. We implemented our scheme on home robots and experimented with KTF operator network, and it shows good communication quality (average MOS = 9.15) and flexible robot controls.

☞ KeyWords : 홈 오토메이션 (home automation), 영상전화 (vedio phone), VoIP (voice of IP), SIP (session initiation protocol)

## 1. 서 론

유비쿼터스 서비스, 즉 사용자가 어느 곳에서든 다양한 장비로 원하는 정보를 습득하거나 조작할 수 있게 하는, 시대는 다른 기술들과 접목되어 빠

르게 발전하고 있다. 로봇 기술의 발전은 산업에서 가정으로, 복잡한 조작에서 간단한 조작으로, 작고 친숙한 외형으로 일반인까지 사용이 확대되고 있다. 가정에 보급된 초기 로봇은 집 안에 있는 사용자의 직접적인 조작에 의해 동작하며 청소 등의 집안일을 수행하였지만, 최근에는 집 안의 전원 장치, 가스 누출 감지 장치, 방법 장치 등 각각의 장비들을 하나의 로봇에 여러 센서를 부착하고 네트워크로 연결하여 홈오토메이션을 구현한다. 인터넷이 사회 전 분야에 영향을 주면서, 모든 통

\* 준 회 원 : 서울시립대학교 대학원 컴퓨터통계학과  
석사과정 재학중 leisure7@uos.ac.kr

\*\* 정 회 원 : 서울시립대학교 컴퓨터과학부 부교수  
swkim7@uos.ac.kr

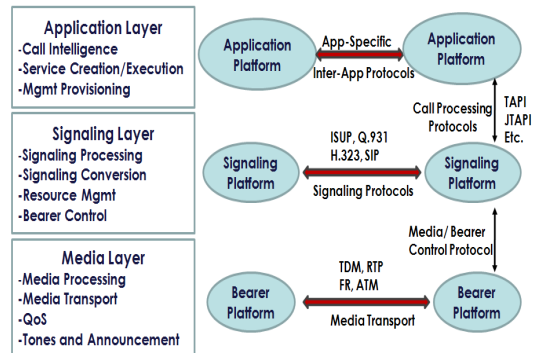
[2009/01/13 투고 - 2009/02/17 심사(2009/05/08 2차) - 2009/06/11 심사완료]

신 분야가 인터넷과 연계한 방식으로 발전되고 있다. 최근에는 초고속 인터넷서비스를 위해 사용 중인 IP망으로 데이터통신은 물론 음성 전화 서비스를 동시에 지원하는 인터넷 폰 또는 VoIP (voice over Internet protocol) 서비스 시대에 접어들고 있다. IP telephony signaling을 위한 대표적인 표준 프로토콜은 ITU-T (international telecommunication union)에서 제안한 H.323 프로토콜과, IETF (international engineering task force)에서 제안한 SIP (session initiation protocol) 이다. 현재는 두 프로토콜 중에서 기존의 HTTP (hyper text transfer protocol), SMTP (simple mail transfer protocol)의 많은 부분을 그대로 차용하여 개발된 텍스트 기반의 프로토콜로 유지 보수 및 관리가 용이하고 호환성, 확장성, 유연성 등에서 탁월한 효과를 가져서 다른 기술 규격과 이 기종 장비들을 통합적으로 제어하고 인터넷 기반 회의, 음성 메일 등 고품질의 멀티미디어 서비스를 통합적으로 지원할 수 있는 SIP의 사용이 확산되고 있다 [1].

본 논문에서는 가정용 로봇에 SIP 프로토콜을 적용하여 집의 외부에서 로봇과의 통신을 가능케 하여, 사용자가 가정용 로봇의 동작을 제어하는 영상 전화기를 설계하고 구현 결과를 보인다. 본 논문에서 구현된 시스템은, 로봇에 장착된 카메라를 통해, 집 안의 상황을 모니터링 할 수 있으며, 카메라 제어와 로봇 움직임 제어를 위해, 전화기의 DTMF (dual tone multi-frequency) 신호를 통하여 원격지에서 로봇을 제어할 수 있도록 구성한다. 또한 필요 시, 영상전화를 통해 로봇에 인접한 사용자와 영상통화를 수행할 수 있도록 호 처리 (call signaling) 프로토콜로는 SIP를 사용하며, 영상과 음성 전달을 위해서 RTP와 H.263 영상코덱을 사용한다. 2장에서는 로봇에 탑재된 영상 전화에서 사용되는 호처리 프로토콜인 SIP에 대하여 설명하고 3장에서는 로봇 영상 전화를 위한 프로토콜 설계 및 로봇 제어를 위한 DTMF 신호 설계를 제시한다. 4장에서는 구현한 영상 전화 시스템의 결과를 보이고, 5장에서 결론을 맺는다.

## 2. 관련 연구

VoIP 기술은 인터넷 응용 기술로서 IP 주소를 기반으로 종단간의 채널 설정을 통해 음성 신호를 압축하고 패킷화한 음성 데이터를 전달하는 기술이다. VoIP의 시스템 구성 요소는 크게 응용 계층 (application layer), 신호 계층(call signalling layer), 매체 계층 (bearer transport layer)으로 나뉘고 각 계층은 해당 프로토콜을 통하여 통신을 수행하게 된다.



(그림 1) VoIP 시스템의 구성 요소

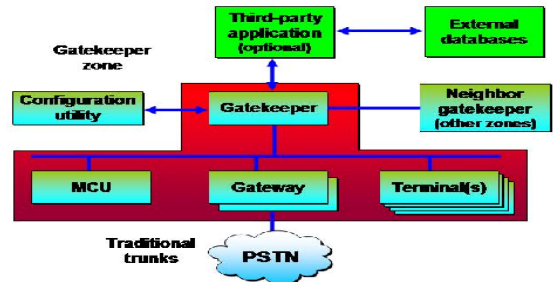
그림 1은 VoIP 시스템의 구성요소를 보여준다. 응용 계층은 서비스 생성, 수행, 관리 기능을 담당하며, 신호 계층은 호 처리, 호 변환, 자원 관리, 매체 제어를 담당하며 SIP, H.323 과 같은 프로토콜을 이용한다. 매체 계층은 실제 데이터 처리, 전달, 변형, 통화 품질 보장 등의 기능을 담당하며 RTP (real-time transport protocol), RTCP (RTP control protocol) 을 이용해 패킷으로 만들어 전송한다 [2].

신호계층의 프로토콜 중 하나인 H.323 프로토콜은, ITU-T SG 16에서 제안한 표준으로 전 세계의 주요 장비 업체나 통신사업자들이 주로 채택하여 기존 제품들과의 연계에 있어서 많은 우위에 있으나, 서비스 품질이 보장되지 않는 LAN 환경에서 실시간 음성, 데이터, 비디오 등 멀티미디어 통신을 지원하기 위하여 개발된 시그널링 프

로토콜이다. 그림 2는 H.323의 구성요소인 터미널 (terminal), 게이트웨이 (gateway), 게이트키퍼 (gatekeeper), MCU (multi-point control unit)의 구성도를 보여준다. 게이트웨이는 서로 다른 네트워크에서 사용되는 터미널 간에 통신이 되도록 하는 것으로 다른 네트워크 간의 미디어 형식의 변환, 정보전달, 호의 설정과 해제를 위한 프로토콜을 해석하는 역할을 수행하고, 게이트키퍼는 H.323의 종말점 (endpoint)들에게 호 제어 서비스를 제공하고 MCU는 3개 이상의 터미널들에 대한 멀티포인트 회의를 지원 한다 [3].

H.323 프로토콜은 대규모 네트워크로의 확장과 대규모 사용자를 지원하는데 한계가 있으며, 서비스 구현이 복잡하고 호환성을 확보하기 어렵다는 단점을 가지고 있다. SIP 표준은 IETF-MMUSIC WG에서 제안한 인터넷 상에서의 멀티미디어 서비스를 위한 표준으로, 멀티미디어 서비스 세션의 생성, 수정, 종료를 제어하는 응용계층 프로토콜로 request/response 구조로써 TCP (transmission control protocol)와 UDP (user datagram protocol)에 모두 사용할 수 있으며, 각 사용자들을 구분하기 위해 이메일 주소와 비슷한 SIP URL을 사용하여 서비스를 제공 한다 [1]. SIP를 구성하는 구성요소는 UA (user agent)와 NS (network server)로 이루어진다. UA는 소프트 폰, 인스턴트 메신저, IP 폰, 휴대폰 등 SIP 프로토콜을 지원하는 터미널 전체를 의미하며, SIP request를 보내는 UAC (user agent server), 상대방의 SIP request에 응답하는 UAS (user agent server)로 구성한다. UAC는 호를 요청하는 UA이며, UAS는 호를 받는 UA이다. NS는 SIP 네트워크를 제어하는 것으로 H.323의 게이트키퍼와 비슷한 역할을 하며, 프록시 (proxy) 서버, 리다이렉트 (redirect) 서버, 레지스타 (registrar)로 구분할 수 있다. 프록시 서버는 전화를 거는 사용자의 UAC로부터 request를 받아, 전화를 받는 사용자의 UAS로 request를 전달하는 역할을 하고, 리다이렉트 서버는 UA의 이동성을 관리한다. Registrar는 UA를 인증하여, 정당한 사용자인지를

구별하고, UA를 SIP 네트워크에 등록하는 기능을 수행한다. UA는 자신을 REGISTER 메시지를 이용해 registrar에 등록하고 다른 UA가 요청하는 경우, 자신의 등록된 정보를 상대 UA에게 제공한다 [4][5].



(그림 2) H.323의 구성요소

MEGACO (media gateway control)는 IETF-MEGACO WG에서 제안한 표준으로 외부의 CA (call agent)에 의하여 게이트웨이를 제어하기 위한 표준으로 MG (media gateway), TG (trunk gateway), SG (signaling gateway) 및 MGC (media gateway controller)간의 게이트웨이 제어 프로토콜로 사용하며, 기존의 음성망을 가지고 있는 Telco, ISP 등이 신규망과 연동하는 경우 사용한다 [2].

RTP는 오디오, 비디오와 같은 실시간 데이터를 멀티캐스트 또는 유니캐스트 네트워크를 이용해서 전송하는 end-to-end 네트워크 전송을 위한 표준이다. RTP는 RSVP (reservation protocol)와 같이 자원 예약 및 적시 전달, 순차 전달 등의 서비스 품질 보증 기능은 수행하지 않는 대신, RTCP를 사용하여, 데이터의 전달상황을 감시하며, 최소한의 제어 기능과 매체 식별 기능을 수행한다. RTP는 실시간 멀티미디어 전송을 위해 UDP 혹은 IP 상단에서 동작한다.

제한된 채널에 더 많은 사용자를 수용하고 더 적은 전송 비트수로 더 높은 서비스 품질을 내기 위해, 고압축율의 코덱 (CODEC: coder-decoder)이 필요하다 [6]. 음성을 압축하는 음성코덱을 vocoder라 하며 PCM (pulse coded modulation) 혹

은 CELP (code excited linear prediction) 방식이 대표적이다. PCM 방식은 무압축 코덱으로, 방식에 따라 64 Kbps 또는 16Kbps 로 동작할 수 있지만, 저전송률에서는 사용할 수 없다. CELP는 서로 다른 백색 가우시안 잡음 (white Gaussian noise)로 코드북 (code book)을 구성하고 입력되는 음성신호와 합성음과의 오차가 최소가 되는 최적의 백색 잡음 열에 해당하는 인덱스를 대신 전송하여 압축효과를 얻어내는 형태를 말한다. 이 방식은 음성신호를 대치할 코드북의 검색 시간이 오래 걸린다는 단점을 가지고 있지만, QCELP (Qualcomm CELP)와 같은 변형된 형태를 핸드폰의 음성코덱으로 사용한다.

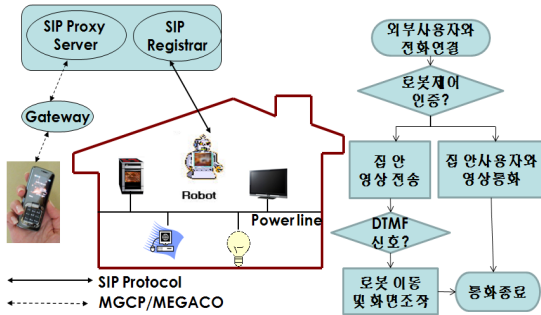
영상 코덱은 H.261, H.263, H264, MPEG-4 (motion picture experts group 4)가 있다. H.261은 ISDN 에서의 영상 전화를 위한 표준으로, 압축률이 상대적으로 낮고, 특히 빠른 화면 전환에서 상대적으로 화질 열화가 심한 단점을 가지고 있다. 반면에 H.263은 PSTN과 같은 낮은 대역폭의 네트워크를 위해, 고압축 코덱으로 낮은 비트 전송률의 응용만이 아니라 광범위한 비트 전송률에도 사용될 수 있다. MPEG-4는 초저속 전송률을 가진 네트워크에서 사용하기 위한 영상압축 표준으로, 영상을 분할하여, 다수의 음성·영상 객체들로 분할하여 전송한다. H.264는 1/4 화소단위의 움직임 예측 알고리즘을 기반으로, 서브블럭 (sub-block) 움직임 추정, 효율적인 엔트로피 인코딩 알고리즘 등을 보완하여, 높은 압축률에도 높은 화질을 기대할 수 있다.

네트워크와 단말기가 다양화되고, 서로 다른 코덱이 이용됨에 따라, 단말 간의 혹은 네트워크 구성요소 간의 코덱 통합은 불가능하며, 이를 위해, 다른 코덱 간의 변환을 수행하는 트랜스코딩 (transcoding) 기능이 필요하다. 또한, 같은 표준을 적용한 통신망이라 하더라도 멀티미디어 콘텐츠 서버와 클라이언트 단말기 간 데이터 전송률, 메모리 크기 등의 차이로 송신 비트스트림의 트랜스코딩이 필요하다 [7].

### 3. 홈오토테이션을 위한 로봇 제어 시스템 설계

최근 'SmartHomes'와 'SmartOffices' 관련하여 다양한 연구가 진행되었으며, 옥스퍼드 대학교의 IWP (Interactive Workspaces Project) [11]는 다양한 장비들 사이에서 이동성을 가지는 여러 사용자들이 여러 장비들과 커뮤니케이션을 할 수 있는 인프라구조를 만드는 것으로, 디스플레이 장비 wireless LAN을 통해서 연결되는 노트북, PDA 등의 디스플레이가 지원되는 모바일 컴퓨팅 장비에서 하드웨어와 소프트웨어 테스트베드를 개발했다. SIP 기반의 유비쿼터스 서비스 시스템 SmartSpace [12]은 멀티미디어 서비스를 제공하는 유비쿼터스 서비스를 연구한 것으로, 멀티미디어 서비스까지 확장이 가능한 SIP 프로토콜의 특성을 이용하여 기존의 X.10으로 연결된 무선 카메라, TV, 전등 등 집 안의 다양한 장비들을 개별적으로 접근하여 제어할 수 있게 하는 것이다. 유비쿼터스 서비스 매니저와 서비스 에이전트로 구성되어 집 안의 기존 네트워크 X.10로 연결되어 있는 장비들을 SIP를 통해서 제어한다. 서비스 매니저는 액세스 포인트, SIP 프록시 서버, SIP 레지스트라, 집 안의 기존 장비들은 X.10 프로토콜로 연결되어 있어 이 장비들과의 통신을 위해 SIP-X.10 변환기로 구성되어있다. 서비스 매니저는 집 안의 제어 가능한 장비 리스트와 장비의 정보를 관리하고, 사용자는 통신을 통해 서비스 매니저에 접근하여 필요로 하는 장비에 대해서 SIP-X.10 변환기를 거쳐 서비스 에이전트에 연결되어 있는 장비들과 통신이 가능하게 된다.

본 논문에서 구현하는 홈오토테이션을 위한 로봇 제어 시스템은, 통신망을 통하여, 원격지의 사용자가 로봇에 장착된 카메라를 통해 홈오토테이션을 수행하는 서비스를 제공하게 되며, 그림 3과 같이 시스템을 구성하였다. 즉 가정의 로봇을 통한 홈오토테이션을 구현하기 위해, 외부의 휴대영상단말기를 사용해 홈오토테이션 제어가 가능하도록 한다.



(그림 3) 홈 오트메이션 시스템 설계 구조

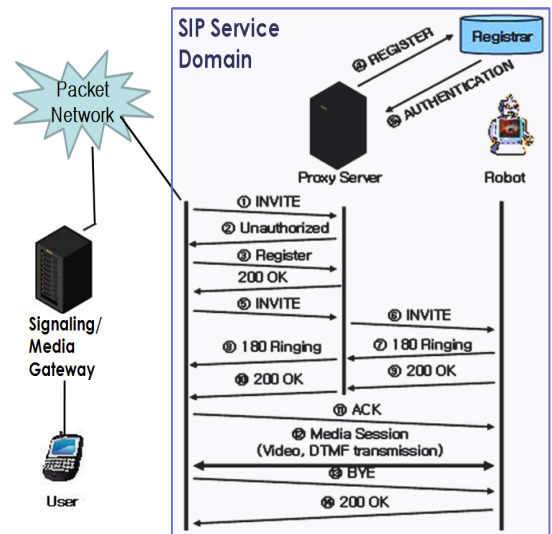
카메라가 장착된 로봇을 이용한 영상 전화 시 나리오는 다음과 같다. 사용자가 외부에서 영상 통화 단말기를 사용하여 집에 있는 로봇과 영상 통화를 시도한다. 그림 3과 같이 사용자는 영상 통화를 시도함으로써 로봇에 장착된 카메라를 통해서 보내지는 집 안의 영상을 볼 수 있고, 사용자가 전화기 버튼을 이용하여 로봇의 동작을 제어하게끔 로봇 영상 전화기의 어플리케이션을 구성한다. 예를 들어 로봇을 전진 혹은 후진 시켜 집 안 구석구석 이동할 수 있게 한다. 이때에 카메라를 통해 찍힌 영상을 모니터링 해 가스의 중간 밸브가 제대로 잠겨 있는지, 에어컨 및 전등이 꺼져 있는지 등을 체크하여 밸브를 잠그고, 에어컨과 전등의 전원을 끄며, 이는 홈 네트워킹과 연결되어 해당 작업을 수행할 수 있다. 이러한 알고리즘 구현을 위해, 로봇과의 영상 통화 기능 및 로봇에 장착된 카메라를 이용하여, 집 안의 원하는 곳을 보도록 로봇 움직임 제어를 위한 DTMF 신호 연결 기능을 수행하도록 영상 통화를 설계한다.

### 3.1 로봇 위치 확인 및 접속을 위한 시그널링 프로토콜 설계

영상 통화를 지원하는 인터넷 폰, 유선 영상 전화기, 휴대폰 영상 전화기를 이용하여, 무선 환경에서 집 안을 자유롭게 움직이는 로봇을 제어하기 위해서 로봇에 장착되는 영상 전화기는 SIP 프로토콜을 기반으로 하며, 전체적인 호처리 순서는

그림 4와 같이 구성한다.

본 논문의 영상전화기는 기존 통신망인 KTF망을 이용하여 실험한다. 사용자는 집의 외부에서 고유의 IP를 가진 가정 로봇에게 전화를 시도한다. 이때 사용자는 먼저 기존의 프록시 서버에 호를 요청한다. 프록시 서버는 콜 설정, 취소, 종료 등 사용자의 콜 관리 요청 기능을 수신 시 해당 SIP 주소로 등록된 현재 사용자의 위치가 복수로 존재할 수 있으므로 등록된 모든 위치에 동시에 요청 메시지를 전달하는 기능을 담당 한다. 또한 사용자의 응답이 오는 경우, 사용자에게는 최선의 응답을 전달해 주고, 동시에 발생시킨 다른 요청 메시지의 요구에 대한 취소 등의 처리를 담당하는 기능을 한다.



(그림 4) 가정용 로봇 제어를 위한 SIP 호 처리

KT망 프록시 서버의 특성상, 인증 단계를 포함하게 되며, 인증 단계가 완료된 후, REGISTER, INVITE 메시지를 통해 호 설정이 이루어진다. 처음 사용자가 로봇과 통신하기 위하여 프록시 서버에게 INVITE 메시지를 보내면 프록시 서버는 레지스터 서버에 인증 여부를 확인하고, 인증이 되어있지 않으면 사용자에게 인증되지 않았음을 알린다. Registrar는 SIP 주소에 대하여 사용자의

위치 정보 즉 IP 주소를 우선순위 정보와 함께 지정된 시간만큼 저장, 관리하며, 인증과정을 통해 인증된 사용자의 요구에 대해 요구된 로봇의 정보를 제공해 준다. 따라서 시스템 구현 상, 로봇의 통신 준비 상태가 되도록 하기 위해, 로봇은 사전에 KT망 프록시 서버에 INVITE 메시지를 전송하여 REGISTER 메시지를 레지스타 서버에 전송하고 정보를 서버에 등록하여 인증절차를 추가적으로 수행해야 한다.

REGISTER 메시지에는 로봇의 SIP URL 주소를 포함한다. 레지스타 서버는 데이터베이스에 SIP URL과 로봇의 IP 주소를 저장한다. 데이터베이스에 접근하는 프록시 서버가 사용자로부터 로봇으로 지정된 INVITE 요청 메시지를 받았을 때, 요청은 로봇의 저장된 IP 주소로 대체한다. 사용자는 프록시 서버에게 INVITE 메시지를 보낸다. 이후 호 처리는 SIP 프로토콜의 호 처리에 따르고 사용자와 로봇과의 전화를 연결한다. 이후 RTP를 통하여, 비디오 및 오디오용 미디어 세션을 설정하며, 미디어 세션을 통해, 로봇에 장착된 카메라를 통하여 집 안의 화면을 사용자에게 전송한다. 사용자가 로봇 이동을 자유롭게 제어하기 위해서 DTMF 신호를 이용한다.

### 3.2 로봇 액추에이터 제어를 위한 DTMF 프로토콜 및 영상코덱 설계

영상 전화로, 로봇의 카메라를 통해 들어오는 영상을 확인하면서, 혹은 로봇에 장착된 여러 센서의 반응에 대한 확인 작업을 수행하기 위해 사용자는 전화 단말기로 로봇의 이동을 제어하여 부엌의 가스 밸브가 잠겨 있는지 출입문이 잠겨 있는지 등 집 안의 여러 상황을 파악한다. 사용자는 전화 단말기의 번호 버튼을 눌러, DTMF 신호를 발생시키고, 해당 DTMF 신호는 SIP INFO 메시지 [9]를 사용하여 로봇에 전달한다. 로봇의 이동을 위한 버튼은 표 1과 같이 제공한다.

(표 1) 로봇 제어를 위한 DTMF 전달 프로토콜 및 SIP INFO 메시지 변환

번호	행동	SIP INFO 메시지 변환
1	로봇 이동 모드 설정/해제	INFO sip:robot@192.168.93.13 SIP/2.0 Via: SIP/2.0/UDP 192.168.100.150:5060;branch=...
4	좌로 이동	Max-Forwards: 70
6	우로 이동	To: robot@192.168.93.13;tag=...
2	앞으로 이동	From: user@192.168.100.150;tag=...
8	뒤로 이동	...
5	화면 확대/축소	CSeq: 3 INFO Content-Type: audio/telephone-event Content-Length: 4 xx xx xx xx

DTMF 신호는 SIP 프로토콜의 INFO 메시지의 Content-Type 에 telephone-event를 명기하여 전달한다. 그리고 전달 할 DTMF 신호는 RFC2833의 payload 형식에 따라 “xx xx xx xx” 형태로 표 1의 SIP INFO 메시지 변환의 예와 같은 방식으로 전송한다.

본 논문에서 구현된 시스템은, 로봇에 장착된 카메라를 통해, 원격지의 사용자는 집 안의 상황을 살필 수 있으며, 필요 시, 로봇 주위에 있는 타인과 영상 통화를 시작할 수 있는 기능을 제공한다. 본 논문에서는 멀티미디어 전송을 위해 보편적인 RTP를 사용하며, RTCP를 사용하며, QoS를 일정 수준 유지할 수 있도록 고려한다. 음성 코덱으로는 CELP 방식을 사용하며, 영상 코덱으로는 H.263 코덱을 사용한다. 최신 코덱인 H.264를 고려하였으나, 양방향 전송을 위한 코덱을 사용하므로, H.264의 1/4 화소 움직임 예측 등을 사용하기에는 실시간 동작을 지원하기 어렵다는 점과 현재의 영상 전화 단말이 H.263 코덱을 지원하고 있어, 트랜스코딩을 제공하면서 실시간성을 더욱 훼손시킨다는 점을 고려하였다.

### 4. 영상 및 로봇제어 신호 전달 시스템 구현

본 논문에서 제안된 시스템 구현을 위해서 실험 환경에는 가정에서 이용할 로봇, 로봇과 사용자간의 통신에 이용될 UA, UAC 와 UAS 간의 호



시그널링에 이용되는 프록시 서버, registrar 서버가 기본적으로 요구된다. 전체시스템은 영상단말과 로봇으로 구성되며, 로봇의 환경은 Mobile Intel Celeron 733MHz의 CPU, RAM 256M에 OS는 Windows XP Service Pack 2 가 탑재하였다.

그림 5는 로봇이 장착된 스마트 폰과 영상 통화 기능이 가능한 휴대폰을 이용한 영상 전화 이용의 예로 좌측은 사용자 휴대폰에서 수신한 영상이고 우측의 영상은 각각 왼쪽은 휴대폰에서 송신한 영상이고 오른쪽은 로봇에 장착된 카메라를 통해 입력된 영상이다. 최초 통화 설정 시, 코덱 지연시간은 1초 이내로 설정하였으며, 초당 15프레임을 송·수신할 수 있도록 구성하였다.



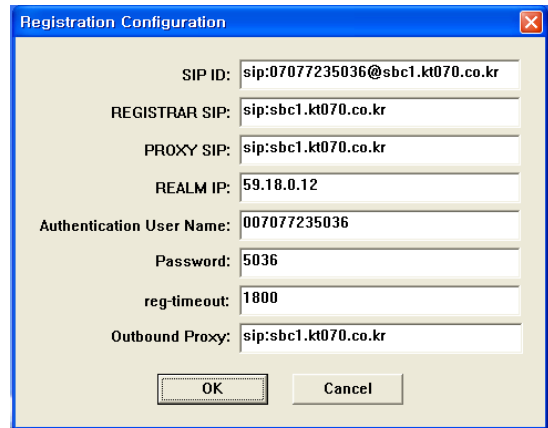
(그림 5) (a) 원격사용자의 영상단말과 (b)로봇에 장착된 홈오메이션 영상서버 구현

사용자와 로봇 영상 전화 프로토콜 구현을 위한 SIP 기능은, 오픈 소스인 sipX[10] 프로그램을 이용하였으며, 그림 6은 통화 설정을 위한 환경 설정 화면이다.



(그림 6) 로봇에 장착된 소프트 폰 및 환경설정

네트워크 서버는 한국통신(KT)의 프록시 서버 및 레지스타 서버를 사용하였다. 제안한 시스템에서 추가한 인증과정에 요구되는 프로토콜은 그림 7과 같이 구현하였다. 인증 요청에 대한 응답으로 레지스타 서버는 “200 OK” 메시지를 프록시 서버에 보내고, 프록시 서버는 “200 OK” 메시지를 다시 사용자에게 보낸다. 사용자는 수신된 “200 OK” 메시지를 통하여 등록된 사실을 인지하고 다시 “INVITE” 메시지를 보내게 되며, 이후에는 앞서 제시한 프록시 서버를 통한 호 절차를 따르게 되며 호 설정 성공 시 영상 통화를 통하여 로봇을 제어하게 된다.



(그림 7) 가입자 망에 특화된 인증 요청

본 논문에서 제안한 시스템을 실제 KTF 망을 통하여 실험한 결과, 핸드폰의 수신 감도에 따른 화질 저하 요소를 제외하고는 영상 화질 저하는 거의 발생하지 않았다. 화질 측정을 위해 영상단말의 이미지를 추출할 수 없는 관계로, 본 논문에서는 객관적 화질 측정 방법인 PSNR (peak signal to noise) 대신에, 가장 많이 사용되는 주관적 화질 측정 방법인 MOS (mean opinion score)를 사용하였으며, 20명의 화질측정 관찰자를 이용하였다. 표 2는 원격지의 사용자가 가정 내 로봇을 제어하기 위해, 로봇에 장착된 카메라를 사용하여 가정 내 환경을 모니터링하는 상황에서의 화질을 측정

한 실험결과이며, 표 2에서 보는 바와 같이, 평균 MOS가 9.15/10.0 을 취득하여, 대부분 수신 감도 변화에도 불구하고 높은 화질을 보이고 있다. 또한, 네트워크 트래픽의 변화에 따른 화질변화 요소를 측정하기 위해, 트래픽이 집중되는 통화시간대에 대하여 수신 화질을 비교하였다.

표 2에서의 실험영상화면 중에서 로봇수신영상은 로봇의 화면을 통해 보이는 원격지의 사용자를 보이고 있으며, 사용자수신영상은 원격지의 사용자가 로봇에 장착된 카메라를 통해 전송되는 가정환경을 영상단말의 디스플레이를 통해 본 화면이다. 표 2의 실험결과 (사용자수신영상)는 영상단말에 보이는 화면을 디지털카메라로 다시 캡처하는 과정에서 화질 저하가 발생하였다.

## 5. 결 론

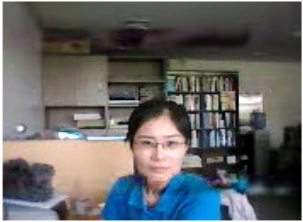
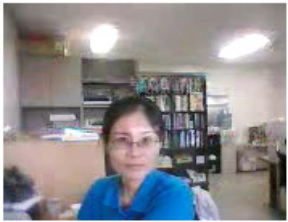
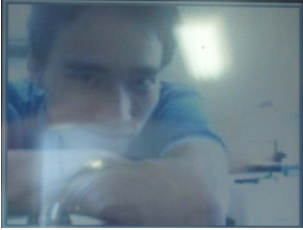







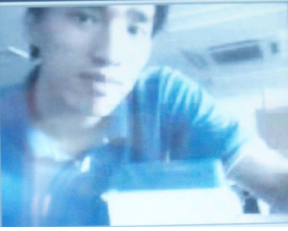
본 논문에서는 홈 네트워킹 구축의 일부인 가정용 로봇 제어 및 로봇에 장착된 카메라를 통한 영상 전화기를 설계하고 구현하였다. 영상 전화기 설계를 위해, VoIP 서비스 및 메신저, 차세대 이동통신 표준에 사용되고 있는 SIP를 사용하였으며, 로봇 움직임 제어를 위한 제어정보 송·수신을 위해 DTMF 메시지 전달을 위한 방식을 설계하였다. 구현은 카메라가 부착된 Mobile Intel Celeron 733MHz의 CPU, RAM 256M에 OS는 Windows XP Service Pack 2 기반의 저사양 로봇에서 실험하였으며, 추후 로봇의 영상 전화 기능 구현 시 고가의 하드웨어를 추가 할 필요 없는 장점을 가진다. 원격지의 사용자는 영상 전화 기능이 제공되는 영상 단말기를 사용하여, 로봇의 카메라 영상을 모니터링하게 되며, 전송 프로토콜은 RTP/RTCP, 코덱은 CELP/H.323 영상코덱을 사용하였다. 초당 15 프레임 이상의 영상을 화질 저하 없이 전송할 수 있으며, DTMF 신호를 전송하여, 로봇 액추에이터 제어를 통하여, 로봇을 원격 사용자 임의대로 조정할 수 있도록 하였다.

## 참 고 문 헌

- [1] "H.323 versus SIP: A Comparison," Packetizer, Aug. 2001, [http://www.packetizer.com/voip/h323\\_vs\\_sip/](http://www.packetizer.com/voip/h323_vs_sip/)
- [2] 민재홍, 조평동, "VoIP 기술 동향," 주기기술동향, 1021호, 한국전자통신연구원, Nov. 2001, pp. 14-24, .
- [3] "H.323: Packet-based multimedia communications systems," ITU-T Recommendation H.323. June 2006
- [4] A. B. Johnston, "SIP: Understanding the Session Initiation Protocol," Artech House, 2001.
- [5] J. Rosenberg et al., "SIP: Session Initiation Protocol," IETF RFC 3261, June 2002
- [6] H. Schulzrinne et al., "RTP: A Transport Protocol Real Time Applications," IETF RFC 1889, Jan. 1996.
- [7] 고영선 외, "Transcoding 기술동향 분석," SITI review, 창간호, 2002
- [9] S. Donovan, "The SIP INFO Method," IETF RFC 2976, Oct. 2000.
- [10] "sipX", open source of SIP, <http://sipxecs.sipfoundry.org/pub/sipXecs/>
- [11] L. Kagal, V. Korolev, S. Avancha, A. Joshi, T. Finin, "A Smartspace Approach to Service Management in Pervasive Systems," Grace Hopper Celebration of Women in Computing, 2002.
- [12] J.Y. Kwak, "Ubiquitous Services System Based on SIP," IEEE Trans. Consumer Electronics, vol. 53, no. 3, Aug. 2007



(표 2) 시간대 별 영상통화 캡처 화면 및 MOS 실험치

통화 시각	14시	15시	16시
로봇 수신 영상			
사용자 수신 영상			
MOS	9.6	8.7	8.9
통화 시각	17시	18시	19시
로봇 수신 영상			
사용자 수신 영상			
MOS	9.3	8.8	9.4

## ● 저 자 소 개 ●

### 조 명 지

1999년 덕성여자대학교 전산학과(공학사)

2007년 ~ 현재 서울시립대학교 대학원 컴퓨터통계학과 석사과정 재학중

관심분야 : 멀티미디어

E-mail : leisure7@uos.ac.kr

### 김 성 환

1999년 한국과학기술원 전자전산학과 박사 졸업

1996년 - 2000년 LG 전자 이동통신연구소 선임연구원

2000년 - 2002년 Cisco Systems EWBU 책임연구원

2002년 - 현재 서울시립대학교 컴퓨터과학부 부교수

관심분야 : 컴퓨터보안, 영상통신, 게임기술

E-mail : swkim7@uos.ac.kr