
하이브리드 웹 기반의 실시간 화상통신 시스템

반태학* · 조태범* · 정희경**

Hybrid Web-Based Real-Time Video Communication System

Tae-Hak Ban* · Tae-Bum Cho* · Hoe-Kyung Jung**

요 약

오늘날 HTML5와 JavaScript를 활용한 하이브리드 웹 기반의 실시간 화상통신 기술이 대두되고 있다. 이에 본 논문에서는 HTML5를 지원하는 웹 브라우저를 이용한 실시간 화상통신 기술에 대해 연구하였다. HTML5와 JavaScript를 기반으로 시스템을 설계 및 구현하였으며, 서로 상이한 웹 브라우저들 간의 웹서버를 통해 RTC API와 HTTP/WebSockets 등을 사용하여 웹 브라우저를 활용한 실시간 화상통신을 가능하게 하였다. 이는 PC 및 노트북, 스마트패드, 스마트폰 등 모든 웹 브라우저의 사용 가능한 단말 기기로 사용 가능으로 별도의 회원가입이나 플러그인 등의 설치 없이 이용 가능하다.

이는 네트워크가 연결된 모든 환경에서 웹 브라우저의 이용이 가능한 모든 단말에서 사용가능하며, 제한적인 환경에서의 화상통신분야에 활용될 것이다.

ABSTRACT

Real-time video communication technology of hybrid Web-based using JavaScript and HTML5 today is on the rise. Therefore, in this paper, we have studied the real-time video communication technology using a Web browser that supports HTML5. It has been designed and implemented a system based on HTML5 and JavaScript, allowing for video communication of real-time, such as you would using the HTTP / WebSockets and RTC API through the Web servers in the Web browser between different each other, utilizing a Web browser was. It is available in a usable terminal equipment of a Web browser PC or notebook, smart pad, such as a smart phone, which is available without installation, such as a plug-in register or another.

In all environments the network is connected, it is available in all terminals use the Web browser is enabled, which will be utilized in the field of video communication in a restricted environment.

키워드

WebRTC, 실시간 통신, 화상 시스템, 실시간 화상통신, HTML5

Key word

WebRTC, real-time communications, imaging system, real-time video communication, HTML5

* 준회원 : 배재대학교 컴퓨터공학과

** 종신회원 : 배재대학교 컴퓨터공학과(교신저자, hkjung@pcu.ac.kr)

접수일자 : 2013. 03. 02

심사완료일자 : 2013. 03. 13

Open Access <http://dx.doi.org/10.6109/jkiice.2013.17.5.1113>

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서 론

웹 브라우저에서의 음성통신, 화상회의 등의 실시간 통신을 위한 표준기술이 IETF(Internet Engineering Task Force)의 rtcweb(real-time communication in Web-browsers) 워킹그룹을 통해 표준화되고 있다. 이에 따라, IETF에서는 프로토콜과 API 요구사항을, W3C에서는 표준 API 개발을 진행하고 있다.

기존의 화상통신 방식은 각각 서비스 제공사의 특정 서버와 프로그램을 통해 화상통신을 제공하는 방식이었다. 이용자들이 서버에 접속하거나 별도의 플러그인(Plug-in)[1]이나 액티브엑스(Active-X)[2] 등을 설치하여야 화상통신이 가능했다. 그러나, 최근 Google 등에서 별도의 화상통신 서버나 어플리케이션, 소프트웨어를 사용하지 않고, 화상통신 서비스에 대해 웹 브라우저만을 사용하여 사용자간 화상통신 서비스를 제공해 주는 WebRTC[3,4] 기술을 진행하고 있다.

기존의 웹을 활용한 화상통신 시스템들은 HTML 구조와 PHP를 활용한 구조였고, 동영상이나 플래시, 액티브엑스나 플러그인 등의 SW를 사용하기 때문에 많은 저장용량을 필요로 하고 효율적인 면에서 부작용들이 많이 존재하였다. 하지만 HTML5에서는 별도의 태그들을 사용하여 간단 명료하게 스트림(stream)에 관한 요소들을 사용할 수 있는 웹상에서 HTML5[5]와 자바스크립트(JAVAScript)를 사용하였다.

본 논문에서는 웹 브라우저에서 화상통신 시스템을 설계 및 구현하였다. 이는 위의 설명과 같이 효율적인 시스템 측면과 기술적 측면에서 우수함이 검증되었다.

II. 관련연구

2.1. SRTP(Secure Real-Time Transport Protocol)[6,7]

SRTP 표준은 인터넷전화에서 음성과 영상 트래픽을 전송하기 위하여 사용되는 RTP 패킷에 대한 암호화 기술 표준이다. 양단간 통화에서 SDP 메시지에 미디어 스트림의 암호화/복호화를 위한 세션 키를 전달할 수 있으며, 이를 이용해 RTP 패킷에 대한 암호화 및 복호화가 가능하다. 실시간 데이터를 암호화 또는 복호화하기 위해 SRTP에서는 기본적으로 AES 알고리즘을

사용한다.

2.2. STUN(Session Traversal Utilities for NAT)[8]

STUN은 실시간 음성/화상/메시지 채팅 및 각종 대화형 IP 통신을 위한 프로그램에서의 NAT 순회(Traversal)를 위하여 사용되는, 네트워크 프로토콜 정의를 포함한 각종 표준화된 방법들의 집합을 의미한다.

2.3. ICE(Interactive Communication Environment)[9]

ICE는 UDP기반 멀티미디어 스트림 세션수립을 위한 NAT 순회 프로토콜이다. ICE 프로토콜을 사용하는 양단간 호스트는 대화형으로 세션을 수립하게 된다. ICE는 비표준인 NAT 환경을 극복하고, 다양한 네트워크 환경에 비 의존적인 범용적인 NAT 순회 솔루션을 제공한다.

III. 웹기반 실시간 화상통신 시스템 설계

웹 기반 실시간 화상통신 시스템 설계에서는 크게 3가지 부분으로 나뉘며, 첫 번째로는 비디오 API 부분, 두 번째로 오디오 API 부분, 세 번째는 P2P API 부분이 필요하고, 별도의 코덱과 전송프로토콜을 사용하였다.

3.1. 웹 기반 실시간 화상통신 시스템 구성도

그림 1의 웹 기반 화상통신 시스템 동작 구성은 크게 두개의 웹 브라우저와 두개의 웹 서버로 구성된다.

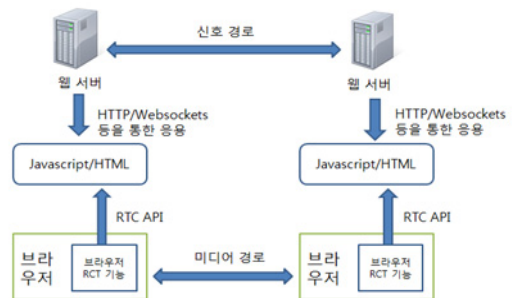


그림 1. 웹 기반 화상통신 시스템 동작 구성도
Fig. 1 Operation Configuration of the Web-Based Video Communication Systems

연결요청을 받은 사용자는 다시 연결을 요청한 사용자에게 응답 메시지를 보내주게 되면, 중간 연결매체인 어플리케이션은 다시 사용자에게 응답에 대한 메시지를 전송하고, 사용자 간 통신을 위한 연결을 실행하여 실시간 화상통신을 진행할 수 있게 한다.

웹 기반 실시간 화상통신 시스템의 특성상 HTML과 자바스크립트를 사용하기 때문에 웹 브라우저 상에서의 영상을 위한 영상 부분과 각각의 API 연결을 위한 소스파일, 컴파일을 위한 `scons` 파일들이 존재한다.

서버 동작을 위한 클래스들은 WebRTC를 중심으로 메인과 라이브러리로 나뉜다. 메인에서는 크게 일곱 개 부분으로 나뉘는데, 첫 번째로는 웹, 두 번째로는 빌드, 세 번째 부분에서는 `xmlite` 부분으로, 이 부분에서는 `xml` 변환과 파싱을 위한 빌더와 파서, 내용들에 대한 콘텐츠들, 내용에 대한 요소에 관련된 엘리먼트, 출력을 위한 프린트로 구성되고, 네 번째 부분에서는 `xmpp` 부분으로, 콘텐츠들과 모듈, 인증 프로토콜을 핸들링하는 부분과, `xmpp` 클라이언트와 엔진, `xmpp`를 처리하기 위한 `task` 부분으로 나뉜다. 다섯번째 부분에서는 세션 부분으로 화상통신을 담당하는 `talk` 부분과 연동되어지고, 이 부분에서는 화상통신을 시작하는 콜부분과, 통신을 하기 위한 채널부분, 영상과 음성의 코덱, 미디어스트림, `rtp`를 사용하는 부분, 영상과 음성에 대한 클립부분으로 구성된다. 여섯번째는 `Base` 부분으로, 이 부분에서는 웹을 사용하기 위한 요소들로 구성되는데, `http`를 이용하기 위한 서버와 클라이언트 부분, 메시지 전송에 대해 핸들링하기 위한 부분, 소켓과 스레드로 구성된다. 마지막은 전송을 위한 `p2p` 부분이고, 이 부분에서는 `p2p` 종류의 하나인 `icc`와 `stun`, `p2p` 전송포트, `http` 포탈 연결부분으로 구성된다.

3.4. 웹 기반 화상통신 응답 시스템 설계

웹 기반 화상통신 응답 시스템 설계에서는 세가지 단계로 구성되는데 연결을 요청하는 사용자(PeerConnection)와 요청을 받는 사용자(Remote Peer), 두 사용자를 이어주는 어플리케이션인 사용자 연결 부분(PeerConnection Observer)으로 구성된다. 연결을 요청하는 사용자로부터 화상통신 요청결과에 대하여 로컬 사용자로부터 요청을 수락하게 되면, 어플리케이션이 응답에 대한 수행 결과를 분석하여, 요청한 사용자에게 응답결과를 통보하여, 상호간 화상통신을 가능하게 한다.

먼저 연결을 요청받은 사용자는 연결을 요청한 사용자로부터 요청에 대하여 응답을 수락하게 되면, 상대방 피어에서는 어플리케이션으로 응답에 대한 결과를 넘겨주게 된다.

어플리케이션에서는 요청받은 결과에 대해 사용자 연결부분을 생성하고 전달하면서 연결을 요청한 사용자에 대한 정보를 생성하게 되고, 이 과정에서 생성된 정보들은 연결요청 사용자에게 메시지로 전달되게 된다. 이 단계에서 연결을 요청하는 사용자에 대한 연결생성과, 생성된 연결에 대한 정보와, 비디오, 오디오를 사용하게 하는 트랙에 대한 정보들을 생성하게 된다. 이 단계를 거쳐 사용자에게 생성된 스트림에 대한 정보를 다시 전송한다. 이후 사용자는 어플리케이션으로 통신정보들에 대한 메시지를 전송하고 마지막으로 어플리케이션은 다시 연결을 요청받는 상대방 연결자에게 연결요청을 보내게 된다. 이러한 과정의 진행이 종료되면 두 사용자간의 실시간 화상통신을 가능하게 한다. 연결 응답 시스템은 그림 4와 같다.

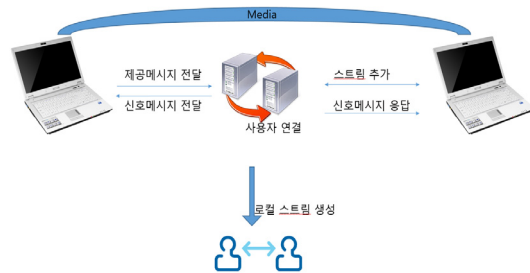


그림 4. 연결 응답 시스템
Fig. 4 Connection Response System

IV. 웹 기반 실시간 화상통신 시스템 구현

4.1. 웹 기반 실시간 화상통신 시스템 구성도

구현을 위한 시스템 구성도에서는 첫 번째로 연결을 요청하는 사용자와 연결의 요청을 응답받는 사용자, 두 번째로는 실시간 화상통신을 위해 웹 브라우저를 사용하는 WebAPI, 세 번째로는 웹 기반 실시간 화상통신 시스템에서 사용자간 상호 연결해 주는 WebRTC API, 네 번째로는 실시간 화상통신 시스템에서의 음성과 영상

을 분석하여, 스트림을 전송해 주는 API들로 구성된다. 구성도는 그림 5와 같다.

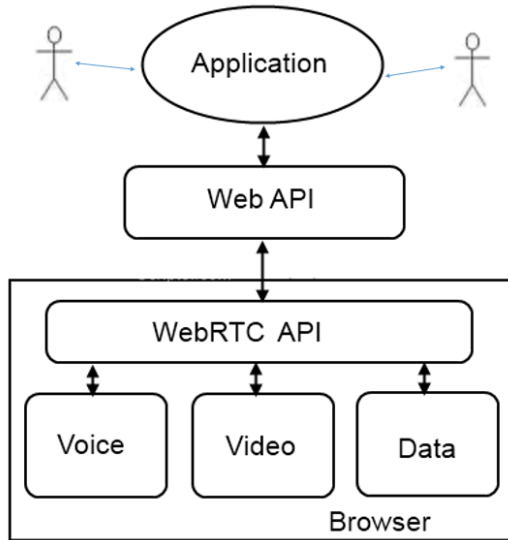


그림 5. 웹 기반 화상통신 시스템 구성도
Fig. 5 Web-based Video Communication Systems

4.2. 웹 기반 화상통신 시스템 구현

사용자가 웹 브라우저 상에서 실시간 화상통신 시스템을 사용하는 단계이다.

크롬 웹 브라우저 상에서 사용자의 시스템에 연결된 영상장치(웹캠, 카메라)의 사용여부를 알려주게 되는데 허용하게 되면 연결을 요청하는 사용자 영상이 웹 브라우저 상에서 나타나게 된다. HTML5 이전의 버전에서는 HTML 태그에 대하여 영상과 음성에 대하여 비디오, 오디오 태그의 링크정보, 화면의 사이즈와 색상표에 대해서 자세히 명시해야 했지만 HTML5에서는 특정한 태그와 값의 입력만으로 화면의 구성이 가능하다.

웹 기반 화상통신 요청 시스템의 구현은 웹 브라우저 상에서 화상통신 시스템의 연결을 요청받은 상대방 사용자를 위한 시스템 구성이다.

연결을 요청받은 상대방 사용자는 웹 브라우저 상에서 URL 주소 검색 창에 상대방의 화상통신 시스템의 URL 주소를 입력하면 상대방의 IP 주소에 대해 HTTP를 사용하여 자동으로 네트워크 상에서 분석하여 연결을 요청한 실시간 화상통신 시스템 사용자의 화상통신 시

스템과 연결 된다. 그림 6은 상대방 사용자와 연결 요청 사용자간의 상호 연결 화면이다.



그림 6. 웹 기반 실시간 화상통신 시스템 상호 연결 화면
Fig. 6 Web-Based Real-Time Video Communication System Interconnects Screen

그림 6과 같이 연결 요청 사용자는 시스템 상에서 서버 구축을 통해 시간과 장소의 제약 없이 인터넷이 연결되고 웹 브라우저의 사용이 가능한 환경이면 상대방 사용자에게 언제든지 실시간 화상통신 요청을 하고 연결함으로써 실시간 화상통신이 가능하다.

V. 결 론

본 논문에서는 HTML5를 활용한 하이브리드 웹 기반 실시간 화상통신 기술을 활용하여 웹 브라우저를 통해 실시간으로 인터넷이 가능한 환경에서 화상통신이 가능한 시스템을 구현하였다. 사용자가 다른 사용자에게 화상통신을 요청하고, 요청받은 사용자는 응답을 함으로써, 상호간 화상통신이 가능하게 하였다. 본 시스템은 웹 상에서 액티브엑스나 별도의 플러그인이 필요하지 않고, 기존의 화상통신 시스템과 같은 서버나 특정 프로그램은 필요로 하지 않는다. 본 논문의 구현 시스템은 웹 브라우저를 지원하는 모든 단말 및 스마트 기기에서 사용 가능하다.

향후 연구 과제로는 현재의 PC 환경이 아닌, 클라우드 컴퓨팅이나 셋톱박스 등에 적용하여 웹을 사용할 수 있는 모든 플랫폼 환경에서 실시간 화상통신이 가능하게 하는 연구와, 1:1 환경이 아닌 1:N의 다수의 사용자를 대상으로 하는 화상통신 연구가 필요하다.

참고문헌

[1] TaeHyuk Sung, MinJae Park, JungHun Lee, Kwanghoon Kim, "Plugin based Messenger Implementation Using Message Protocol", Korean Society for Internet Information Fall Conference. Vol.6 No.2. 1993

[2] <http://www.active-x.com>

[3] <http://www.webrtc.org>

[4] HaSung Koo, JongIk Shim, "The Study of Dynamic Flow Control Method using RSST in Video Conference System", The Journal of the Korean Institute of Information and Communication Engineering. Vol.9 No.8. Dec.2005

[5] HTML5 Open Community, <http://htmlfive.co.kr>

[6] <http://whitelka.blog.me/60119534928>

[7] MinYeong Kim, Khongorzul D., ByungCheol Shin, InSung Lee, "Design of RTP/UDP/IP Header Compression Protocol in Wired Networks", The Journal of the Korean Institute of Information and Communication Engineering. Vol.9 No.8. Dec. 2005

[8] <http://frontjang.info/279>

[9] <http://www.smoothweb.com/ICE?sid=wlxwpxtf>

저자소개



반태 학(Tae-Hak Ban)

2011년 배재대학교 컴퓨터공학과 (공학사)
 2013년 배재대학교 컴퓨터공학과 (공학석사)

2013년~현재 배재대학교 컴퓨터공학과 박사과정
 ※ 관심분야: 멀티미디어문서정보처리, 차세대 인터넷, USN, HTML5, 영상처리 시스템



조태범(Tae-Beom Cho)

1999년 배재대학교 정보통신대학원 정보통신학과(이석사)
 2001년~2006년 배재대학교 IT교육센터 책임강사

2007년~2011년 배재대학교 교양지원센터, 컴퓨터공학과 외래강사
 2010년 배재대학교 대학원 컴퓨터공학과(공학박사)
 2008년~현재 을지대학교 외래강사
 ※ 관심분야: XML, 모바일 시스템, 웹 보안, 멀티미디어, 데이터베이스



정희경(Hoe-kyung Jung)

1985년 광운대학교 컴퓨터공학과(공학사)
 1987년 광운대학교 컴퓨터공학과(공학석사)

1993년 광운대학교 컴퓨터공학과(공학박사)
 1994년~현재 배재대학교 컴퓨터공학과 교수
 ※ 관심분야: 멀티미디어 문서정보처리, XML, Web Services, Semantic Web, MPEG-21, Ubiquitous Computing, USN, Cloud Computing