

Tele-Experience 실감 스마트워크 서비스를 위한 텔레프레즌스 기술

이미숙, 강진아, 황인기, 구기중, 김도영
한국전자통신연구원

요약

본고에서는 Tele-Experience 실감 스마트워크 서비스를 제공하기 위한 기반 기술인 텔레프레즌스 기술동향에 대해 살펴본다. 최근의 텔레프레즌스 시스템은 회의 참여자들의 몰입감을 향상시키기 위해 다양한 신기술들을 채택하고 있는데, 본고에서는 실감 스마트워크 서비스를 위해 특별히 차별화할 수 있는 3가지 핵심 기술 - 시선 맞춤, 실감 오디오 기술, 그리고 끊임없는 미디어 재생을 위한 패킷손실 복구기술을 중심으로 그 동향을 분석한다.

I. 서론

스마트워크는 시간과 장소에 얽매이지 않고 언제 어디서나 편리하고 똑똑하게 근무함으로써 업무 효율성을 향상시킬 수 있는 개념을 의미한다[1]. 이러한 스마트워크의 도입은 저탄소 녹색성장 시대에 대비하고, 저출산, 고령화, 낮은 노동생산성 등의 사회문제 해소에 도움을 줄 수 있어 정부에서도 많은 관심을 가지고 있다.

최근에는 네트워크에서의 미디어 처리 기술 및 단말기술 등의 발전과 더불어 원격에서 마치 사무실 공간에 있는 것처럼 업무를 처리할 수 있도록 하는 실감 스마트워크에 대한 요구도 증가하고 있으며 “사무실, 학교, 병원, 경기장, 회의장에 직접 가지 않아도 현재의 공간상에서 실시간으로 느끼고 상호작용이 가능한 Tele-Experience 실감 스마트워크 서비스” 기술개발을 진행하고 있다.

Tele-Experience 실감 스마트워크 서비스를 위한 핵심 기반 기술 중 하나는 텔레프레즌스 기술이다. 초창기의 텔레프레즌스는 대형 스크린과 고품질의 비디오와 오디오 지원을 주요 특징으로 하였으나 최근의 텔레프레즌스는 참여자들이 마치 앞에서 서로 마주하고 회의를 하는 듯한 느낌을 주어 회의에 대한 몰입도를 높이기 위해 다양한 기술들을 도입하고 있다.

몰입도를 높일 수 있는 요소 기술로는 끊임없는 미디어 (오디오 및 비디오) 재생을 기반으로 하여 참여자의 실물 크기 영상 (life-size), 원격 참여자간의 시선 맞춤 (Eye-contact), 참여자의 위치 정보 및 동일 공간감을 줄 수 있는 실감 오디오, 그리고 동일 공간감을 극대화 하기 위한 공간적 일치성 (Background-consistency) 등이 언급되고 있다. 회의 참여자의 실물 크기 영상을 재현하기 위해서 55" 이상의 스크린을 사용하고, 공간적 일치감을 주기 위하여 회의용 가구를 텔레프레즌스 솔루션에 포함을 시키기도 한다.

본 고에서는 외국 선도기관들을 중심으로 텔레프레즌스 기술 동향에 대해 살펴본다. 2장에서는 대표 기업들의 텔레프레즌스 시스템에 대한 특징에 대해 살펴보고, 3장에서는 텔레프레즌스의 몰입감을 높이기 위한 핵심 요소 기술 중에서 시선 맞춤, 실감 오디오 및 끊임없는 미디어 재생을 위한 패킷손실 복구에 대한 기술동향을 살펴본다.

II. 텔레프레즌스 시스템 동향

IDC등 여러 리서치 기관들의 시장 분석 결과를 보면 텔레

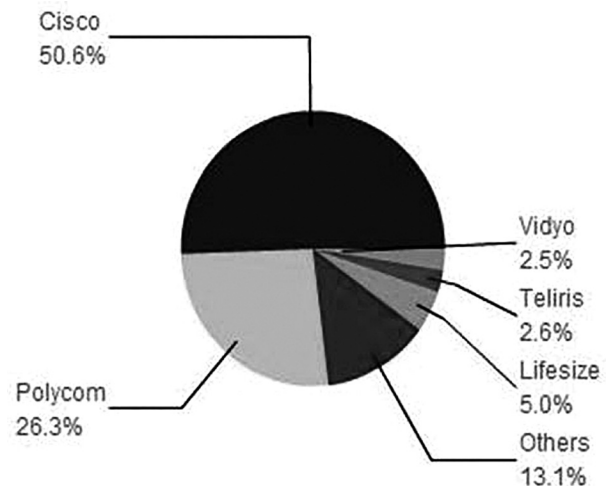


그림 1. 텔레프레즌스 시장 점유율

프레즌스 시장을 선도하는 기업은 Cisco와 Polycom이며, Huawei, Avaya, Vidyo, Logitech 등의 회사가 일부 시장을 점유하고 있다[2]. <그림 1>은 IDC에서 조사한 2012년 1분기 시장 조사 결과이다. 본 절에서는 Cisco와 Polycom 그리고 Vidyo사의 제품 특징에 대해 살펴본다.

1. Cisco

Cisco는 2006년 12월 시장에 진입한 이래로 최근까지 지속적으로 자사의 제품군을 확장해 왔고, 2010년 Tandberg 인수를 통해 계속해서 영상회의 시장에서 리더십을 유지하고 있다[3].

Cisco는 몰입감을 높이기 위한 전용 회의실에서부터 용도에 따라 변경 가능한 다목적 회의실, 개인용 데스크 탑이나 모바일 제품까지 폭넓은 제품군을 보유하고 있다. 또한, 텔레프레즌스 단말뿐만 아니라 네트워크 인프라, 클라우드 솔루션을 표준 규격에 따라 제공함으로써, 다양한 사용자 요구에 적합한 솔루션을 구축할 수 있게 해준다.

<표 1>과 <그림 2>는 Cisco사의 주요 제품과 특징을 정리한 것이다.

표 1. Cisco 주요 제품[4]

제품명	주요 특징
Cisco TelePresence TX9000	<ul style="list-style-type: none"> 전용 회의실 용 최대 6인 참석 동시 3개 1080p60 비디오 지원 3개 65"스크린
Cisco TelePresence MX300	<ul style="list-style-type: none"> 다목적 회의실 용 1080p30 비디오 지원 H.264/SVC 지원 1개 55"스크린
TelePresence EX Series	<ul style="list-style-type: none"> 데스크 탑 용 1080p30 비디오 지원



그림 2. Cisco의 TelePresence TX9000과 MX300

2. Polycom

Polycom은 Destiny Conferencing (Telesuite)의 인수와 함께 텔레프레즌스 시장에 뛰어 들었다. 현재 광범위한 텔레프레즌스 제품과 영상회의의 제품 모두를 보유하고 있으며, 표준 규격

에 맞는 제품을 제공함으로써 타사의 영상회의 또는 텔레프레즌스 장비와의 호환성을 제공한다[3].

Polycom의 텔레프레즌스 플랫폼인 RealPresence의 주요 특징은 사용자가 Apple iPad 태블릿을 이용하여 손쉽게 회의를 시작하거나 제어할 수 있는 SmartPairing 기술, 네트워크 사용 대역폭 감축 기술, true-to-life 협업을 제공할 수 있는 영상 및 음성 처리 기술, 사용자의 시스템 선택과 운용의 편리성을 제공하는 상호 연동 기능 등이 있다.

<표 2>와 <그림 3>은 Polycom사의 주요 제품과 특징에 대해 정리한 것이다.

표 2. Polycom 주요 제품[5]

제품명	주요 특징
Polycom RPX Solutions	<ul style="list-style-type: none"> 전용 회의실 용 4 ~ 28인 참석 3개 84"스크린 1080p60 비디오
Polycom HDX9000	<ul style="list-style-type: none"> 다목적 회의용 1080p 비디오 지원 H.264 High profile 지원 Polycom Lost Packet Recovery
Polycom RealPresence Mobile	<ul style="list-style-type: none"> 모바일 소프트웨어 iOS/안드로이드 지원 VGA 30fps 비디오 디코딩

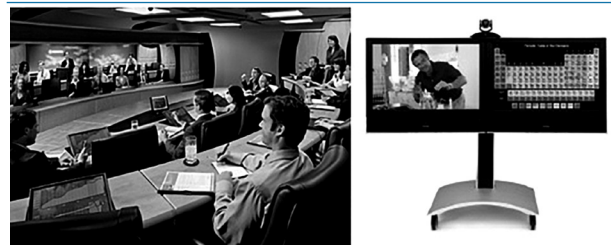


그림 3. Polycom의 RPX solution과 HDX9000

3. Vidyo

Vidyo 의 시장진출 전략은 보다 저렴한 텔레프레즌스 및 영상회의 시스템을 제공하고 대형 텔레프레즌스 기업들과의 직접적인 경쟁을 피할 수 있는 세분화된 시장을 목표로 하고있다[3].

Vidyo의 순음형 비디오 레이어링 아키텍처 기술은 H.264, SVC (Scalable Video Coding) 기반 압축 기술과 Vidyo의 IP를 활용하여 각 엔드포인트의 비디오를 능동적으로 최적화하는 기술로, 이 기술을 이용하면 값비싼 MCU를 사용하지 않고도 오류를 복원하고 네트워크 지연을 줄일 수 있다.

<표 3>와 <그림 4>는 Vidyo사의 주요 제품과 특징에 대해 정리한 것이다.

표 3. Vidyo 주요 제품[6]

제품명	주요 특징
VidyoPanorama	<ul style="list-style-type: none"> • 10% 비용으로 텔레프레즌스 서비스 제공 • 1080p60 화면 최대 6개 • 순응형 비디오 레이어링 아키텍처 지원
VidyoRoom	<ul style="list-style-type: none"> • 1080p30 비디오 지원 • 순응형 비디오 레이어링 아키텍처 지원 • 오류 복원 기능 제공
VidyoDesktop	<ul style="list-style-type: none"> • 소프트웨어 클라이언트 • Windows/Linux 지원 • 비디오 품질 및 네트워크 성능 실시간 모니터링
VidyoMobile	<ul style="list-style-type: none"> • 모바일 소프트웨어 • iOS/안드로이드 지원 • 720p 비디오 디코딩, VGA 비디오 인코딩



그림 4. Vidyo의 Panorama 와 VidyoMobile

III. 텔레프레즌스 요소 기술

1. 시선 맞춤 (Eye-contact) 기술

시선 맞춤 (Eye-contact)은 상대방이 어떤 자세로 대화에 임하고 있으며, 어느 정도 자신의 이야기에 관심을 가지고 있는지를 알게 해 주는 매우 중요한 무언의 언어 (non-verbal language)이다. 따라서 시선 맞춤은 대화에 대한 몰입도를 높일 뿐만 아니라 상대에 대한 신뢰성과 친밀감을 증가시킨다[7].

기존의 텔레프레즌스 환경에서는 상대방이 자신이 아닌 다른 곳을 보고 있는 것처럼 느끼는 경우가 많이 있다. 이는 <그림 5>와 같이 카메라가 아닌 스크린을 보고 이야기 하기 때문에 발생하는 문제이다. 나 (로컬참여자)는 스크린에 나타난 상대방 (원격 참여자)의 얼굴을 보고 있지만 카메라에 촬영된 영상 속의 나는 아래쪽을 쳐다보는 얼굴상을 하고 있다. 그렇다고 시선을 카메라에 두면 상대방의 얼굴을 자세히 볼 수 없어 시선 맞춤이 무의미해진다.

이러한 문제점을 해결하기 위해 많은 기관들에서 시선 맞춤 기술에 대한 개발을 진행 중이고 일부는 제품으로 출시되기도 하였다.

시선 맞춤 기술은 크게 다음과 같이 두 가지 유형으로 분류할

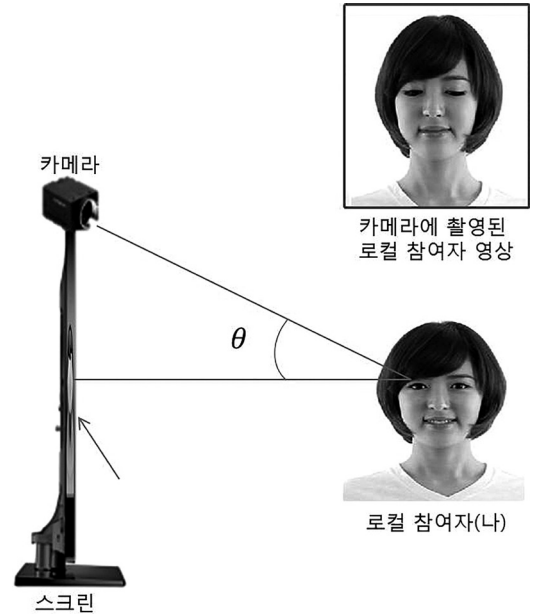


그림 5. 카메라의 위치와 시선 맞춤

수 있다.

- 1) 하드웨어 방식
- 2) 소프트웨어 방식

하드웨어 방식은 카메라의 위치를 조정하여 시선 맞춤 영상을 제공하는 기술을 의미하고, 소프트웨어 방식은 영상처리 기술을 이용하여 한 대 이상의 카메라로 얻은 영상을 기반으로 시선 맞춤이 된 영상을 합성하는 기술을 의미한다.

1.1 하드웨어 방식의 시선 맞춤 기술

하드웨어 방식의 시선 맞춤 기술은 카메라를 회의 참여자 시선의 연장선상에 위치시키는 방식으로, 현재 출시된 시선 맞춤 기능을 갖는 텔레프레즌스 시스템은 하드웨어 방식을 채택하고 있다. 하드웨어 방식은 자연스러운 시선 맞춤 영상을 제공하지만, 시스템의 구성이 복잡하고 많은 비용이 소요된다는 문제점이 있다.

하드웨어 방식을 도입한 대표적인 기관으로는 반투과 스크린 방식의 DVEtelepresence와 Sony가 있으며, 스크린과 카메라의 각도를 최소화하는 방식을 채택한 Cisco와 HP, 그리고 스크린 중간에 작은 구멍을 만들어 카메라를 설치하는 방식의 Polycom이 있다. 본 절에서는 DVEtelepresence와 Cisco의 기술에 좀 더 자세히 대해 살펴보기로 한다.

(1) DVEtelepresence

DVEtelepresence에서 제안한 반투과 방식을 이용한 시선 맞춤 방법에 대한 동작 원리는 <그림 6>과 같다. 먼저, Projec-

tor (①) 의 영상은 반투과 스크린 (③)을 통과하여 projector screen (②)에 도달하고, 다시 반사되어 반투과 스크린에 상을 맺게 되는데 회의 참여자는 단말의 뒷면 (④)에 상이 있는 것으로 느끼게 된다. 그리고 카메라 (⑤)는 반투과 스크린의 뒤 쪽에 사용자의 시선 상에 위치시켜 시선 맞춤이 가능하게 하는 방식이다[8].

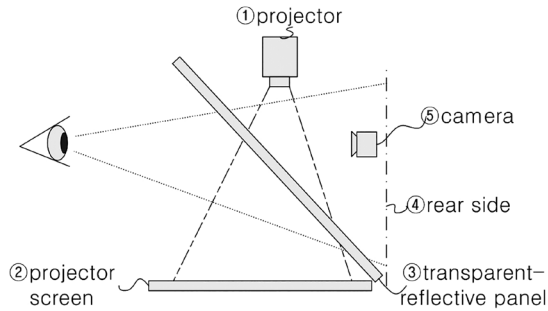


그림 6. DVTelepresence의 시선 맞춤 기술

이 방식은 반투과 스크린이 45도 기울어져 있어야 하므로 전체 시스템의 크기가 커지고, 반투과 스크린이 빛의 양을 반만 통과시키고 나머지는 반사시키기 때문에 사용자가 보는 화면이 흐려진다는 문제점이 있다.

(2) Cisco

Cisco에서는 <그림 7>과 같이 카메라의 위치를 최대한 낮춰 회의 참여자의 시선을 기준으로 스크린과 카메라간의 각도 θ 를 최소화하는 방식을 통해 시선 맞춤을 제공하고 있다[9]. 일반적으로 이 각도가 인지범위($< 5^\circ$) 내로 맞춰지면 참석자는 시선 맞춤이 되었다고 인식하는 것으로 알려져 있다[10]. 이 방식이 적용된 제품으로는 TX9000 이 있는데, 카메라가 스크린 위쪽 일부를 가리기 때문에 참여자의 시선을 방해한다는 문제점을 가지고 있다.

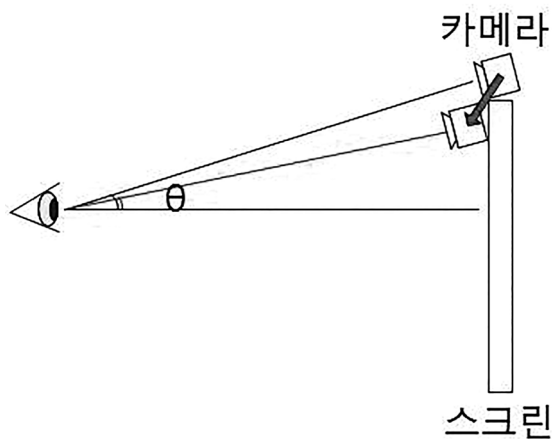


그림 7. Cisco의 시선 맞춤 기술

1.2 소프트웨어 방식의 시선 맞춤 기술

소프트웨어 방식의 시선 맞춤에서는 영상처리를 통해 스크린 밖에 있는 카메라로 촬영한 영상을 스크린 중앙에 있는 가상 카메라의 시점으로 변환하는 방식이다. 이 방식은 하드웨어 방식에 비해 기존 시스템의 변형을 최소화하면서 시선 맞춤을 제공할 수 있기 때문에 Fraunhofer HHI, Microsoft research, ETH Zurich등 많은 기관에서 개발을 진행하고 있다. 그러나 실시간 처리 및 시선 맞춤이 된 합성영상에서 나타나는 왜곡 등의 문제로 인해 아직까지 상용화가 이루어지지 않았다.

(1) Fraunhofer HHI

Fraunhofer HHI는 스테레오 정합 및 visual hull 알고리즘을 사용하여 시선 맞춤 영상을 합성하는 기술을 개발하였다. 특히, 스테레오 정합을 통해 정밀한 깊이 영상을 구하기 위해 다양한 카메라 구성을 제안하였는데, 대표적으로는 <그림 8>과 같이 작은 베이스라인 구조를 통해 깊이 영상의 오차를 줄이고, Trifocal 구조를 통해 깊이 영상의 정확도와 신뢰도를 높이고, 하단의 큰 베이스라인 구조를 통해 깊이 영상의 해상도를 높이고자 하였다[11].

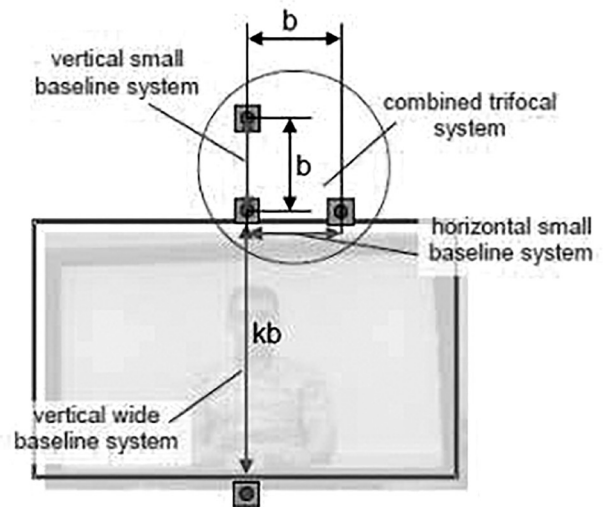


그림 8. Fraunhofer HHI의 카메라 구성도

최근에는 <그림 9>와 같이 “L”자 형태로 설치한 4대의 카메라로부터 얻은 영상을 이용해서 실시간으로 깊이 영상을 탐색하고, 화자의 시선을 추적해서 사용자가 바라보는 위치의 임의 시점의 영상을 합성하는 시연을 하였다[12].

(2) ETH Zurich

ETH Zurich에서 제안한 시선 맞춤 시스템은 <그림 10>과 같이 개인용 PC와 스크린 하단에 위치한 Kinect로 구성된다. Kinect로부터 얻은 깊이 영상에서 전경과 배경을 분리한 후 전경에 있는 홀을 채운다. 그리고 이 깊이 영상을 이용하여 색상



그림 9. Fraunhofer HH의 시선 맞춤 시연



그림 10. ETH zurich의 시선 맞춤 시스템

영상을 3차원 변환을 통해 시선 맞춤 영상을 구한 후 얼굴부분만 따로 추출하여 원래의 색상영상에 덮어 씌운 후 경계부근에서 발생한 왜곡을 제거하는 방식으로, 실시간으로 VGA급의 시선 맞춤 영상을 합성할 수 있다[13].

2. 실감 오디오 기술

오디오 기술은 텔레프레즌스 시스템이 사용자에게 궁극적으로 제공하고자 하는 실제감과 몰입감, 즉 한 공간에서 함께 일하는 것과 같은 느낌을 제공하기 위한 요소기술로서 적용되고 있다. 본 절에서는 먼저 텔레프레즌스를 위한 오디오 기술 개발 현황에 대해서 살펴본 후, 보다 향상된 몰입감을 제공하기 위해 적용될 수 있는 오디오 기술 대해서 살펴본다.

2.1 오디오 기술 개발 현황

텔레프레즌스를 위한 실감 오디오 기술은 마이크를 통해 입력되는 오디오 신호의 부호화, 전송, 출력 전반에 관여된다. 텔레

프레즌스 기술 선도기업인 Cisco, Polycom의 실감 오디오 기술을 살펴보면, HD급 고화질 영상에 맞추어 고품질의 오디오를 제공하기 위한 광대역 음성코덱 기술과 다채널의 마이크/스피커가 개방된 공간에 함께 존재함으로써 인해 발생하는 음향 에코 제거 (AEC: Acoustic Echo Cancellation)를 통한 공간 오디오 품질을 향상시키는 방향으로 진행되고 있다[4][5].

ETRI에서 개발한 텔레프레즌스는 이에 더하여 주화자 검출 (MSD: Main Speaker Decision)이라는 기능을 제공한다. 이는 여러 원격지에서 텔레프레즌스 회의에 참석할 경우 오디오 신호를 분석하여 어느 원격지가 회의의 주도권을 잡고 있는지를 자동으로 판별하는 기능이다. 그 결과에 따라 원격지 상대방으로부터 수신된 오디오 데이터를 효과적으로 믹싱하고 회의 참여자의 별도 작동 없이 주화면 (HD급 비디오) 과 부화면 (QCIF급 비디오)을 구성함으로써 사용자에게 보다 향상된 실감성을 제공할 수 있다[14].

오디오 신호 압축을 위한 코덱은 제품간 상호연동을 위해 표준화가 필요한 기술로써, 원격지에 있는 상대방 음성이 같은 회의실 안에서 들리는 것처럼 느껴지도록 하기 위해, 오디오 신호 대역폭을 기존의 Narrowband (300-3,400Hz)와 Wideband (50-7,000Hz)에서 Superwideband (50-14,000Hz) 또는 Fullband (20-20,000Hz)로 확장하고 있다. 이와 관련된 표준은 ITU-T와 IETF를 중심으로 이루어지고 있다. 또한 Cisco의 텔레프레즌스 시스템은 오디오 코덱으로 AAC-LD (Advanced Audio Coding with Low Delay)를 지원한다[14][15][4].

2.2 몰입감 향상을 위한 오디오 기술

앞에서 살펴본대로 현재까지의 텔레프레즌스를 위한 오디오 기술은 대역폭이 Fullband에 이르는 고품질 오디오 코덱 개발과 회의 룸에서의 음질향상을 위한 다채널 음향 에코를 제거하기 위한 기술개발로 수행되었다. 그러나 텔레프레즌스 시스템이 궁극적으로 추구하는 실제감과 몰입감을 실현하기 위해서는 공간 오디오 및 객체 기반 오디오 처리기술과의 접목이 필요하다.

즉, 텔레프레즌스 시스템 화면에 보여지는 화자와 동일한 공간상의 위치에 오디오를 표현하기 위해서는 3D 오디오 기술이 적용될 수 있고, MSD와 같이 주화자와 부화자를 결정한 후 그에 따른 오디오 신호 레벨 조정 등의 가공을 위해서는 오디오 객체 분리 기술을 적용되는 것이 적합하다. 또한 텔레프레즌스 서비스를 제공받는 다양한 단말 시스템의 재생 환경에 맞춰서 최적화된 오디오 품질을 제공하기 위한 오디오 채널 업/다운 믹싱 및 스피커 어레이를 이용한 공간 음향 재현 기술이 적용될 수 있다[16].

공간 음향 재현 기술은 헤드폰 환경과 스피커 환경에 따라 구

분되는데, 헤드폰 환경에서는 HRTF (Head-Related Transfer Function)라는 사람의 머리 모양과 귓바퀴 등의 신체모양을 수학적으로 분석한 모델을 이용해서 기존 음원을 3차원적으로 가공하여 재현하는 방법이 대표적이다. 스피커 환경에서는 10채널 이상의 다채널을 스피커 어레이를 이용하여 물리적으로 음장을 재현하기 위한 기술이 연구되고 있고 있다[17].

오디오 객체 분리 기술은 일반적으로 3차원 공간상에 존재하는 실제 소리를 마이크로폰 어레이를 가지고 마이크로폰간의 시간차 (channel-time difference) 혹은 강도차 (channel-level difference)를 이용한다. <그림 11>은 단일지향성 마이크로폰 2개와 및 쌍지향성 마이크로폰 1개를 이용한 마이크로폰 어레이로써, 정면과 후면을 향하는 단일지향성 마이크로폰 사이에 쌍지향성 마이크로폰이 배치된다. 이러한 마이크로폰 어레이를 통해 획득된 빔 패턴을 수학적으로 조합하여 원하는 방향의 빔 패턴을 형성함으로써 front left, front right, center, rear left, rear right 채널들에 대한 방향으로 원하는 신호를 획득할 수 있다. 또한 획득된 소리는 오디오 채널간 신호 강도차 분석을 통해 특정 오디오 객체로써 획득이 가능하다 [18]. 이렇게 오디오 신호를 객체단위로 분리하는 기술은 화면상에 표시되는 화자의 위치나 역할에 맞추어 오디오 신호를 가공하는데 적용될 수 있다.

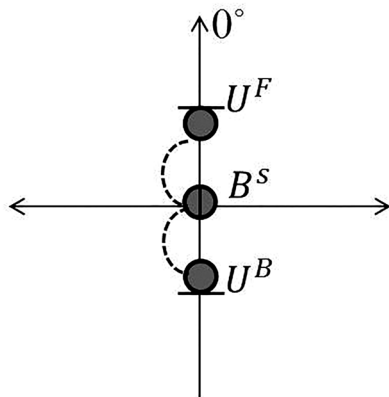


그림 11. 오디오 객체 분리를 위한 마이크로폰 어레이 구성 예

3. 응용계층 미디어 패킷 손실 복구 기술

응용계층 미디어 (오디오 및 영상) 패킷손실복구는 유선 및 모바일 (3G, LTE, Wi-Fi) IP망에서 끊김 없는 영상회의가 가능하도록 손실된 미디어 패킷을 응용계층에서 실시간으로 복구하는 기술이다.

영상회의에서는 아주 적은 양의 미디어 패킷 손실도 사용자의 체감 품질에 심각한 영향을 끼칠 수 있으므로, 일정 수준 이하의

패킷 손실률에서는 100% 미디어 패킷을 복구할 수 있는 패킷손실복구 기능을 제공해야 한다. 모바일 단말을 통한 영상회의 참여를 위해 유선에 비해 상대적으로 패킷 손실률이 더 높고 변화가 큰 무선 환경을 고려하여 네트워크의 품질에 적응적인 미디어 패킷손실복구 기능을 제공해야 한다. 또한 양방향 통화에서는 통화 품질이 미디어 전송 지연에 민감하므로, 통화 품질을 저하시키지 않는 실시간 패킷손실복구 기능을 제공해야 한다.

본고에서는 글로벌 영상회의의 시스템 시장의 70% 이상을 차지하고 있는 Polycom과 Cisco의 패킷손실복구 기술 및 최신의 응용계층 패킷손실복구 기술인 퀄컴 (Qualcomm)의 RaptorQ 기술을 소개한다.

3.1 Polycom

Polycom은 2007년 4분기에, 웨인하우스 리서치 (Wainhouse Research) 기관을 통해 자사의 HDX 비디오 시스템에 탑재된 패킷손실복구 기술인 LPR (Lost Packet Recovery) 에 대한 성능 평가를 의뢰하였다[19].

LPR은 IP 영상통화 중 네트워크 패킷 손실의 영향을 줄일 수 있는 패킷손실복구 기술로서 응용계층 순방향 오류정정 (AL-FEC: Application level - Forward Error Correction) 방식을 사용한다. 송수신시스템에서 통화 대역폭의 일부를 FEC 데이터를 보낼 데이터 채널로 할당하고, 손실된 패킷을 모두 복구할 수 있도록 FEC 데이터 채널의 대역폭을 동적으로 조정한다.

LPR에 추가로, 현재 비디오 프레임의 콘텐츠를 추정하기 위해 이웃 매크로블록, 이전 프레임과 향후 프레임의 패킷 정보를 이용해서 패킷 손실 효과를 보상하는 패킷손실은닉 기술 (PVEC: Polycom Video Error Concealment)을 제공한다.

<그림 12>는 LPR과 PVEC 기술을 적용한 결과이다. 패킷손실률이 5% 미만일 때 XGA (1024 x 768)급 영상을 끊김없이 선명하게 복원하고, 음성은 10~20초 마다 아주 미세한 끊김 (hiccup)이 발생한다. 비디오와 오디오 외에 프리젠테이션 데이터에 대한 패킷손실을 복구한다.

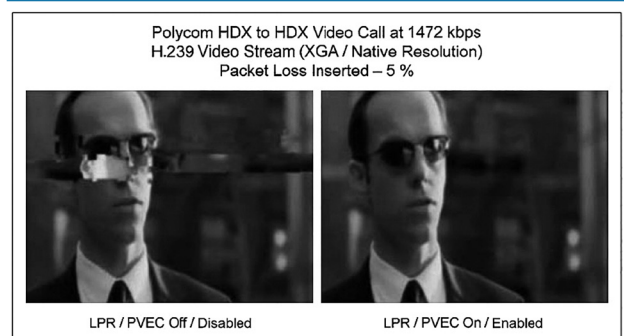


그림 12. 패킷손실 5% 조건에서 통화품질 - LPR+PVEC 기능 OFF/ON

LPR 기술은 Polycom의 리얼프레즌스 제품군 (RealPresence Group 300/500/700)에 탑재되어 있다.

3.2 Qualcomm

RaptorQ는 네트워크 패킷 손실에 대해 응용계층 데이터를 보호하는 소프트웨어로 구현된 AL-FEC 기술이다[20][21][22]. RaptorQ는 채널부호화 기술인 Digital Fountain 기술에서 시작된 Raptor 기술 중에서 가장 유연하고 강력한 기술이며, 비디오 스트리밍, 영상회의 또는 파일 전송 서비스에서 별도의 제어채널 없이 전송 중에 손실된 데이터를 복원할 수 있다. <그림 13>은 RaptorQ의 동작 개념으로 송신 시스템에서 콘텐츠를 RaptorQ로 인코딩하고, 인코딩된 콘텐츠는 네트워크를 통해 목적지 장치 (device)로 전달되며, 원래 콘텐츠를 복원하기 위해 RaptorQ 디코딩을 한다.

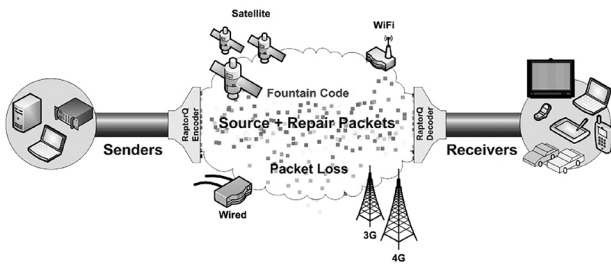


그림 13. RaptorQ의 손실복구 개념도

RaptorQ는 선형시간 인코딩 및 디코딩이 가능하고, 충분한 수의 부호화된 데이터가 수신 장치에 도달한다면, 패킷 손실 패턴에 관계없이 손실복구가 가능하다는 특징이 있다.

<그림 14>는 Airbone 네트워크[24]에서 비디오 스트리밍에 RaptorQ 기술을 적용한 결과이다. 25%의 오버헤드(overhead)를 추가하여 원영상을 복원한다.

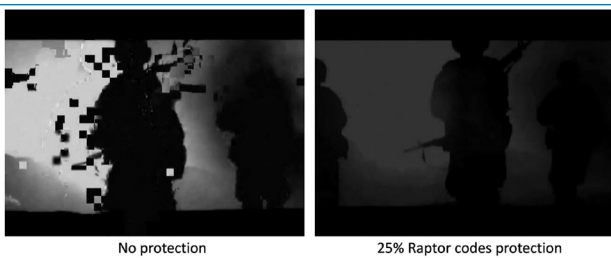


그림 14. Airbone 네트워크에서 비디오 스트리밍 품질 비교

3.3 Cisco Systems

Cisco의 IPLR(Intelligent Packet Loss Recovery) 기술은 IP망에 이미 설치되어 있는 대역폭을 예약하거나 비디오 트래픽에 우선 순위를 할당하는 기존의 QoS 기술을 대체하는 것이

아니라 기존 QoS 기술이 패킷 손실 제어에 실패하는 아주 드문 경우에 순간적인 패킷 손실 상황을 극복하고, 패킷 손실이 현재 비디오 품질에 미치는 영향을 줄이기 위한 기술이다[23].

송신단말에서 빈번한 패킷 손실 (약 1~2%)이 감지 되는 경우, 인코더는 강인한 모드 (robustness mode)로 전환되어, 손실시 영상 품질에 심각한 영향을 끼치는 인트라블록 (intrablocks)을 여러 프레임에 나눠 전송한다. 수신단말에서 영상 패킷 손실이 감지됐을 때, 디코더는 패킷 손실에 의해 어떤 정보가 사라졌는지 추정하여, 손실 정보의 영향이 최소화되도록 한다.

또한, IPLR 외에도 2초 동안 10% 이상의 과도한 패킷 손실이 검출되면, 허용 가능한 패킷 손실이 감지될 때까지, 자동으로 플로우 컨트롤을 이용해 상대방의 데이터 전송률을 64kbps 단위로 줄이는 다운스피딩 (Downspeeding) 기법이 같이 사용된다.

IPLR 기술은 Cisco TelePresence System Codec 6000 MXP, TX-9200 제품군 등에 탑재되어 있다.

3.4 ETRI

ETRI의 패킷손실복구 (PLR: Packet Loss Recovery) 기술은 영상 정보의 중요도를 고려하여 영상 프레임 단위로 손실복구를 수행한다. H.264/AVC 기반의 Full-HD 영상회의를 지원하고, 실시간 손실복구가 가능하도록 짧은 지연 시간 (< 33.3ms (1-frame) @30frames/sec)을 제공한다. <그림 15>와 같이 채널상태를 모니터링하고, 동적으로 복구 알고리즘을 적용하는 구조이다.

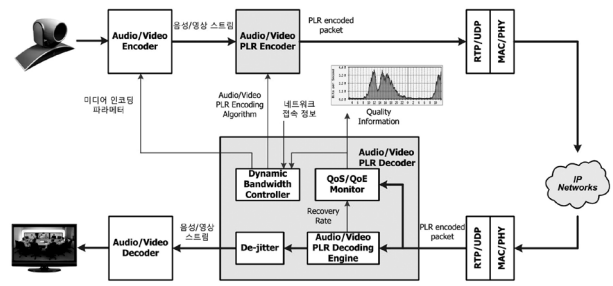


그림 15. 패킷손실복구 기술 구조도

<그림 16>은 20~30%의 추가 정보를 이용하여 5% 수준의 랜덤 패킷손실을 복구한 결과이다. <그림 17>은 패킷손실에 따른

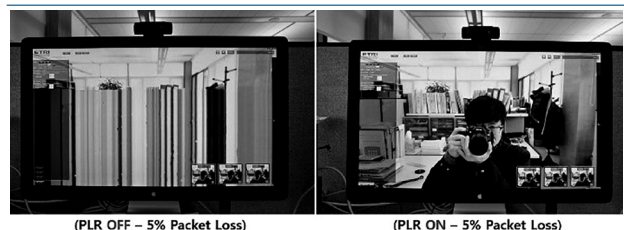


그림 16. 패킷손실 5% 조건에서 영상 품질 비교 - PLR OFF/ON

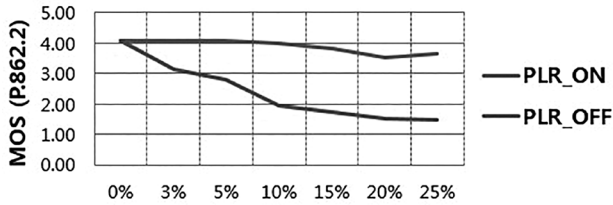


그림 17. 패킷손실환경에서 음성품질 비교

음성품질 측정 결과로, 130% 수준의 부가정보를 이용하여 10% 패킷손실 환경에서도 MOS 4.0 수준의 품질을 유지한다.

IV. 결론

스마트워크는 종래의 사무실 개념을 탈피하여, 언제 어디서나 (Anytime, Anywhere) 편리하게 효율적으로 업무에 종사할 수 있도록 하는 미래지향적인 업무환경을 의미한다. 정부에서는 이러한 스마트워크의 도입이 저출산, 고령화 등의 사회문제 해소와 GHG 가스의 배출감소에 도움을 줄 수 있을 것으로 기대하고 2015년까지 전체 근로자의 30%가 스마트워크 환경에서 근무할 수 있도록 하는 것을 목표로 스마트워크 활성화를 추진하고 있다.

본고에서는 기가미디어급 환경에서 Tele-Experience 실감 스마트워크 서비스를 구현하기 위한 기반 기술로서 현재 텔레프레즌스 기술동향에 대해 살펴보았다. 특히 원격 참여자들이 같은 공간에 있는 것과 같은 체험을 제공하기 위한 핵심기술인 시선 맞춤 기술, 실감 오디오 기술, 그리고 끊임없는 미디어 재생을 통해 몰입감을 높일 수 있는 패킷손실 복구 기술을 중심으로 기술동향을 살펴보았다.

Acknowledgement

본 연구는 미래창조과학부 '범부처 Giga KOREA 사업'의 일환으로 수행하였음. [GK13P0100, Giga Media 기반 Tele-experience 서비스 SW플랫폼 기술 개발]

참고 문헌

[1] 방송통신위원회, “스마트워크 활성화 추진계획,” 2011년 1월.
 [2] “Worldwide Enterprise Videoconferencing and Telepresence Market Growth Decelerates on Weakness

in the Immersive Telepresence Segment,” IDC, May 2012.

[3] 전자통신동향분석 제26권 제3호, “텔레프레즌스 시장분석과 기업동향,” 2011년 06월.
 [4] <http://www.cisco.com>
 [5] <http://www.polycom.com>
 [6] <http://www.vidyo.com>
 [7] <http://www.imcca.org/news/the-importance-of-eye-contact-in-visual-communications>.
 [8] US 5639151, “Pass-through reflective projection display”.
 [9] <http://www.cisco.com/c/en/us/products/collaboration-endpoints/telepresence-tx9000-series/index.html>.
 [10] Milton Chen, “Leveraging the Asymmetric Sensitivity of Eye Contact for Videoconferencing,” CHI 2002, Apr. pp. 20–25, 2002.
 [11] O. Schreer, N. Atzapadin, I. Feldmann, “Multi-baseline Disparity Fusion for Immersive Videoconferencing,” 2nd Int. Conference on Immersive Telecommunications (IMMERSCOM 2009), University of California, Berkeley, CA, USA, pp. 27–29, May 2009.
 [12] W. Waizenegger, N. Atzapadin, O. Schreer, I. Feldmann, “Patch-sweeping with robust prior for high precision depth estimation in real-time systems,” 18th IEEE International conference on Image Processing, 2011.
 [13] Claudia Kuster, Tiberiu Popa, Jean-Charles Bazin, Craig Gotsman, Markus Gross, “Gaze Correction for Home Video Conferencing,” Proceedings of ACM SIGGRAPH Asia, 2012.
 [14] 김도영, 최승한, 이미숙, 강헌식, 전우직, 장두원, “실감 텔레프레즌스 기술,” 한국통신학회논문지, 제 29권, 제 12호, pp. 10–17, 2012년 12월.
 [15] 이종화, 강신각, “텔레프레즌스 기술 표준화 동향,” 한국통신학회논문지, 제 29권, 제 12호, pp. 25–30, 2012년 12월.
 [16] 서정일, 유재현, 강경옥, 장세진, “실감방송을 위한 오디오 표준화 현황,” 방송공학회지, 제 19권, 제 1호, pp. 37–47, 2014년 1월.
 [17] 김정훈, 권기수, “사용자 맞춤형 실감 음향 기술의 현황과

전망,” 방송공학회지, 제 19권, 제 1호, pp. 10-21, 2014년 1월.

- [18] C. J. Chun and H. K. Kim, “Sound source separation using interaural intensity difference in real environments,” Proceedings of Audio Engineering Society Convention, New York, NY, preprint 8976, Oct. 2013.
- [19] Ira M. Weinstein, “Polycom’s Lost Packet Recovery (LPR) Capability,” Wainhouse Research, 2008.
- [20] <http://www.qualcomm.com/Raptor>, “RaptorQ-Datasheet,” 2012.
- [21] <http://www.qualcomm.com/Raptor>, “RaptorQ Forward Error Correction Technology Overview and Use Cases,” 2012
- [22] Michael Luby, “Raptor codes Application Layer FEC,” ICNC, Jan. 2012.
- [23] White paper, “Tandberg and Packet Loss,” D50165, Rev2,0, Jun 2009.
- [24] http://en.wikipedia.org/wiki/Airborne_Networking

약 력



이 미 숙

1991년 호서대학교 전자공학과 졸업
 1993년 호서대학교 전자공학과 석사
 2001년 한국과학기술원 전기 및 전자공학과 박사
 2004년~현재 한국전자통신연구원 책임연구원
 관심분야: 음성·오디오 신호처리



강 진 아

2001년 제주대학교 통신공학과 졸업
 2005년 제주대학교 통신공학과 석사
 2012년 광주과학기술원 정보통신공학부 박사
 2014년~현재 한국전자통신연구원 선임연구원
 관심분야: 음성·오디오 전송 및 신호처리



황 인 기

1999년 성균관대학교 전자공학과 졸업
 2001년 성균관대학교 전자공학과 석사
 2001년~현재 한국전자통신연구원 선임연구원
 관심분야: 병렬처리 시스템, 영상회의



구 기 종

1999년 충남대학교 전자공학과 졸업
 2001년 충남대학교 전자공학과 석사
 2012년~현재 충남대학교 정보통신공학과 박사과정
 2000년~현재 한국전자통신연구원 선임연구원
 관심분야: 멀티미디어 통신



김 도 영

1985년 성균관대학교 전자공학과 졸업
 1987년 성균관대학교 대학원 전자공학과 석사
 2007년 충남대학교 대학원 컴퓨터공학과 박사
 1987년~현재 한국전자통신연구원 실장/
 책임연구원
 2008년~2009년 미국 SJSU 전기공학과 객원교수
 관심분야: 스마트워크, 실감 멀티미디어통신,
 실시간 대용량 미디어 처리