

Embedded VoIP 시스템에서 음질개선을 위한 음성패킷 변환기법

*강진아, *양영배, **임재윤
*제주대학교 통신공학과, **제주대학교 통신컴퓨터공학부
전화 : 064-754-3635 / 핸드폰 : 016-696-1917

The method of Voice Packet Transformation for the Improvement of Voice Quality in Embedded VoIP System

Jin-Ah Kang, Young-Bae Yang, Jea-Yun Lim
Dept. of Telecommunication Engineering, Cheju National University
E-mail : the_dreamer@hanmail.net

Abstract

In this paper, we propose the method of voice packet transformation for the improvement of voice quality in embedded VoIP system terminals. For this purpose, it was analyzed about the RTP header in the voice packet receiving side and designed about the handling method of voice packets by using jitter buffer. Through this analyzed and designed results, by implementing the voice packet transformation method and testing on the self-product VoIP system, we reached that a conclusion is satisfied with performance of VoIP services.

I. 서론

통신 네트워크의 전반적인 발전 추세는 아날로그/디지털 네트워크에서 디지털 네트워크로 진화되고 있으며 네트워크 효율 및 경쟁력 향상을 위해 회선교환 방식에서 패킷교환 방식으로, 폐쇄형에서 개방형으로 전환·발전되고 있다. 이러한 통신 네트워크의 발전 추세에 맞추어 여러 인터넷 응용들 중에서 주목받고 있는 VoIP(Voice over IP) 기반의 인터넷 텔레포니 서비스는 일반 사용자들에게 새로운 기술을 인식시켰으며 기존의 음성통신 시장을 변화시키고 있다.

그리고 저렴한 통신비용, 부가 서비스 제공, 글로벌 로밍(global roaming) 서비스 제공, 통합 네트워크 구축 등의 장점에 기인하여 VoIP 기술은 인터넷의 최대 응용 서비스 중 하나로 급부상하고 있고, 동시에 급속한 시장 확대를 가져오고 있는 분야이다. 그러나 전화 체계, 과금 문제, 지능망을 통한 부가 서비스의 연동 기술, 표준화 문제 등 해결해야 할 과제들이 남아 있으며, 기존의 일반전화에 비해 떨어지는 음질은 개선해야 할 중요한 문제 중의 하나이다[1-4].

시스템 구현과 관련하여 VoIP 단말기는 여러 가지 종류가 개발되고 있는데, 그 중 Embedded VoIP 시스템이 각광받고 있으며, 그 이유는 사용이 편리하고 PC 기반의 단말기에 비하여 안정적인 성능을 보장하기 때문이다.

본 논문은 Embedded VoIP 시스템을 이용하여 음질 개선을 위한 연구를 수행한 것으로, 먼저 음질과 Embedded VoIP 시스템에 대해 살펴본 후 음성패킷 변환처리 기법을 구현하고, 실제 이 기법을 Embedded VoIP 시스템에 적용하여 그 개선 효과를 확인한다.

II: 음질과 Embedded VoIP 시스템

2.1 VoIP 음성패킷 구조와 음질 평가

VoIP 시스템에서 일반적으로 사용하는 음성패킷 구

조는 그림 1과 같고 RTP(Real Time Protocol)/UDP/IP 등으로 이루어진다. 패킷길이는 음성 데이터의 크기에 따라 20~160byte가 된다. RTP는 IP 네트워크에서 미디어를 전송하는 프로토콜이며 RTP 헤더 구조는 그림 1의 하단부와 같다.

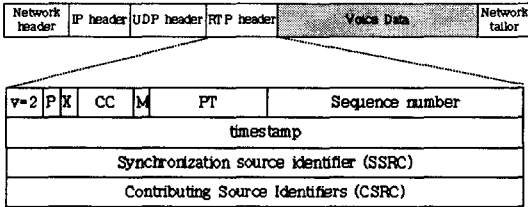


그림 1. VoIP 음성패킷 구조

VoIP 음성통신은 인터넷의 대역폭을 효율적으로 활용하기 위해 음성 데이터를 압축하여 전송하므로 음질이 떨어지게 된다. 한편 음질은 지연에 상당히 민감하나 일반적으로 그 시간이 150~200ms 사이이면 대부분의 이용자는 그 성능에 만족하게 된다[1,2]. 그림 2에 나타낸 것과 같이 통신망에서 일어나는 지연에는 전달지연과 처리지연이 있다. 전달지연은 광섬유나 구리를 매체로 사용하는 통신망에서 매체의 특성에 따라 나타나고, 처리지연은 정보를 처리하는 통신기의 정보 입력에서 출력까지 걸리는 시간에 의해 발생하며, 코덱지연과 대기열지연 등이 포함된다.

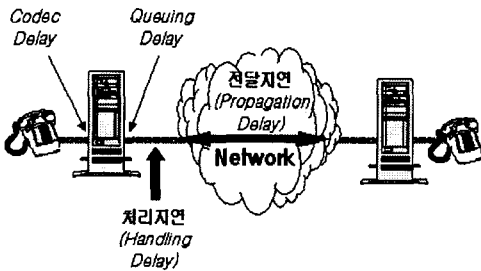


그림 2. 통신망에서의 지연 요소

전달지연과 처리지연 이외에 음질에 영향을 미치는 중요한 요소 중 하나가 지터(jitter)이다. 지터란 예정된 음성 데이터 패킷 수신과 실제 패킷 수신 사이의 편차이다. 그림 3에 나타나듯이 송신자가 패킷 A, B, C를 균일한 시간 간격(D_1)으로 송신하더라도 네트워크의 트래픽 상태에 따라 전달지연 및 처리지연이 균일하지 않으므로, 패킷 A, B, C가 수신자에 도착하는 시간 간격이 균일하지 못하게 된다($D_2 \neq D_3$). 이렇게 수신 데이터 사이의 간격이 균일하지 않을 때, 지터는 음질

을 현저히 떨어뜨린다. VoIP 시스템들은 이러한 지터를 해결하기 위해 지터 버퍼를 설정해서 수신자 측의 패킷 간격을 균일하게 만든다.

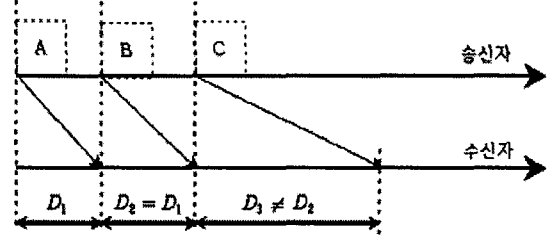


그림 3. 송신자-대-수신자간 지터 발생

2.2 Embedded VoIP 시스템 구조

그림 4는 Embedded VoIP 시스템 구조에 대한 대표적인 형태이다. 중앙처리장치인 MCU(Micro Controller Unit)를 중심으로 메모리, Ethernet 통신 모듈, 음성신호처리 모듈, 그리고 keypad와 LCD로 이루어진 사용자 입·출력 모듈로 구성되어 있다.

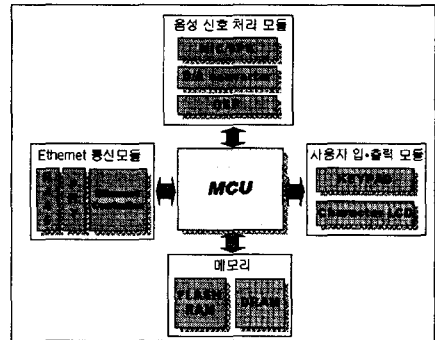


그림 4. Embedded VoIP 시스템 구조

이 시스템은 음성신호처리 DSP를 따로 두어 음성 데이터 압축을 하드웨어로 처리하도록 함으로써 소프트웨어의 부담을 줄이고 음질을 개선시킨 형태이다.

III. 음성패킷 변환처리 기법

3.1 RTP 헤더 분석

음성패킷을 수신한 VoIP 시스템은 패킷의 RTP 헤더의 sequence number를 분석하여 순서에 맞게 음성 데이터를 지터 버퍼에 쌓도록 한다. 음성패킷의 순서가 역행하였을 때에는 그림 5와 같이 재 정렬하도록 한다.

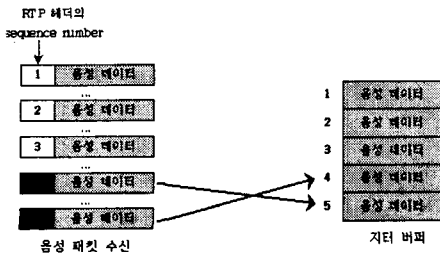


그림 5. 음성패킷의 순차적 정렬

3.2 지터 버퍼의 가변적 조절

네트워크 상태에 따라 지터가 변하는 것은 앞장에서 설명하였다. 따라서 지터 버퍼의 크기를 고정된 값으로 설정하지 않고 네트워크 상태, 즉 네트워크의 패킷 전달지연 크기에 따라 가변적으로 설정된다면 보다 안정적인 음질이 보장될 것이다. 따라서 음성통신 초기 단계에 지터 버퍼의 크기를 설정하는 절차를 거치도록 하며, 이는 지터 버퍼 크기의 효율적인 값이 송신단-대-수신단간의 고정적인 지연시간과 관계된다는 사실에 착안하여 구현한다.

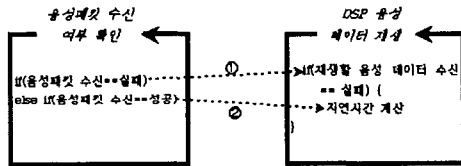


그림 6. 송신단-대-수신단 지연 측정

그림 6은 송신단-대-수신단의 지연시간을 측정하는 알고리즘을 나타낸다. ①에서 음성패킷이 수신되지 않는 동안은 DSP가 재생할 음성 데이터가 없는 상태가 되며 이 구간의 시간을 측정한다. ②에서 음성패킷이 수신되면 ①에서 측정된 시간을 패킷 전달지연으로 수용한다. 이렇게 측정된 단-대-단 지연시간을 권고되는 지터 크기 150ms에서 감산함으로써 지터 버퍼의 크기를 찾고, 그 값으로 지터 버퍼를 생성한다. 이는 그림 7과 같다.

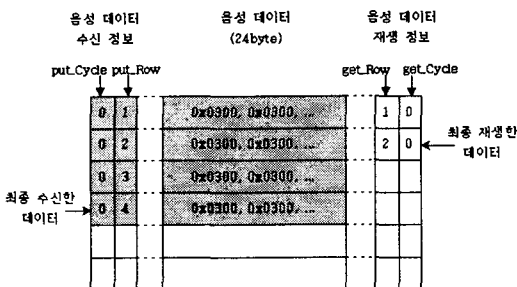


그림 7. 지터 버퍼 구성

지터 버퍼의 음성 데이터 액세스에 대한 정보는 네 가지 변수로 두었는데, 수신된 음성 데이터의 지터 버퍼상의 위치는 (put_Cycle, put_Row)로 나타내고, 재생할 음성 데이터의 지터 버퍼상의 위치는 (get_Cycle, get_Row)로써 나타낸다. 이러한 위치 정보를 통해 음성 데이터는 지터 버퍼 상에 쌓이거나 재생될 수 있으며, 또한 이 정보에 의해 수신량과 재생량을 비교하여 적절히 대처하도록 한다. 예를 들어 재생량이 수신량을 초과하는 경우에는, 이전의 음성 데이터를 다시 재생하도록 하여 음성두절 현상을 개선한다.

그림 8은 이러한 변환처리 기법을 적용한 전체 음성 통화 프로그램의 구조를 간략히 나타내었다. 좌측의 프로세스는 음성패킷을 수신하고 RTP 헤더를 분석 및 정렬하여서 음성 데이터를 지터 버퍼에 넣는다. 한편 우측의 프로세스는 지터 버퍼로부터 음성 데이터를 가져와서 재생(play)하고 동시에 녹음(record)된 음성 데이터를 RTP 패킷화하여 전송한다. 이러한 동작은 실시간으로 이루어지며 두 개의 프로세스는 스레드(thread)로 구현하여 동시에 수행되도록 한다.

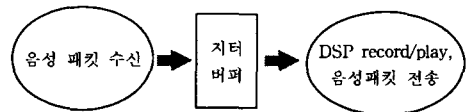


그림 8. 전체 음성통신 프로세스 구성

IV. 실험 및 분석

그림 9는 본 논문에서 수행한 실험의 구성을 나타낸 것이다. 구현된 Embedded VoIP 시스템은 ARM7TDMI Core의 32bit network MCU와 G.723.1 음성압축 방식의 DSP, 그리고 Ethernet 통신 블록을 통합한 형태의 DSP, 그리고 Ethernet 통신 블록을 통합한 형태이다. signaling protocol은 H.323 기반으로 이루어진다.

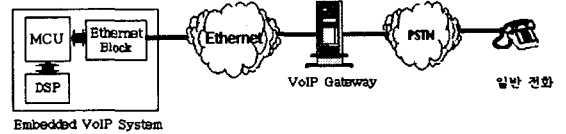


그림 9. 음성통화 실험 구성도

실험을 위하여, 앞서 설계한 음성패킷 변환처리 기법을 embedded linux 기반의 C 프로그램으로 구현하여 타겟 시스템에서 동작시킴으로써 음성통화 실험을 시행하였다. 그림 10과 그림 11은 음성통화의 실험결과를 호스트 PC 화면에 출력한 내용이다.

그림 10은 음성통화에서 수신된 RTP 헤더의 sequence number를 나타내며, 패킷 손실 또한 12[packet]/2000[packet] 만큼 발생하였다. 이 실험에

서 패킷순서의 역전 현상은 발생하지 않았다.

```

=====
TEST_1>>> RTP Header Parsing
=====
debug Parameter: [Sequence Number]
.
[1931] [1933] [1934] [1935] [1936] [1937]
[1938] [1939] [1940] [1941] [1942] [1943]
[1944] [1945] [1946] [1947] [1948] [1949]
[1950] [1951] [1952] [1953] [1954] [1955]
.
[1998] [1999] [2000]
- Total Rx RTP Packet = 2000 [pkt]
- RTP Packet Loss     = 12  [pkt]
- RTP Reverse         = 0   [pkt]
=====
    
```

그림 10. 수신 음성패킷의 RTP 헤더 분석결과

그림 11은 지터 버퍼를 이용한 음성패킷 변환처리 기법에 대한 실험결과로써, TEST_2는 기법을 적용하지 않았을 때를, TEST_3은 적용하였을 때의 내용이다.

```

=====
TEST_2>>> disable
=====
- Total Rx RTP Packet = 2000 [pkt]
- DSP access exceed   = 6   [time]
=====
TEST_3>>> enable
=====
debug Parameter:
[Put_Cycle][Put_Row][Get_Cycle][Get_Row]
.
[001] [000] [000] [001]
[001] [001] [000] [002]
[001] [002] [000] [003]
[001] [003] [001] [000]
[002] [000] [001] [001]
.
- Total Rx RTP Packet = 2000 [pkt]
- Jitter Buffer Size   = 4   [frame]
- DSP access exceed   = 0   [time]
=====
    
```

그림 11. 변환처리 기법의 비적용 · 적용 결과

TEST_2와 TEST_3 간의 비교되는 값은 “DSP access exceed”라고 표현하였고 이것은 DSP의 재생속도(또는 타이밍)와 음성패킷의 수신속도(또는 타이밍)의 차이로 인해 DSP가 음성 데이터의 재생을 놓치는 경우의 횟수이다. 이는 수신된 음성패킷의 RTP 헤더의 sequence number와 DSP가 재생한 음성패킷의 RTP 헤더의 sequence number를 비교함으로써 구한

값이다. TEST_2에서 6회에 이르렀던 값이 TEST_3에서는 0회로 발생하지 않았다. 이 때의 변환처리 기법에서 적용한 지터 버퍼의 효율적인 크기는 4 frame(4 word)으로 설정되었음을 나타내며, (Put_Cycle, Put_Row)와 (Get_Cycle, Get_Row)의 간격이 지터 버퍼 크기와 동일한 4 frame을 일정하게 유지하는 현상이 나타났다.

실험결과를 보면 그림 10의 실험에서는 예상했던 음성패킷 순서의 역전현상이 나타나지 않았는데, 이는 다양한 네트워크 환경에 대해 더 많은 실험이 필요하다. 그림 11의 실험에서는 구현된 음성패킷 변환처리 기법을 적용함으로써 적절한 지터 버퍼의 크기를 찾아 내어 음성패킷의 수신속도와 재생속도간의 효율적인 완충작용을 통해 안정적인 DSP의 음성재생을 보장받을 수 있었으므로 음질개선 효과를 얻을 수 있었다.

V. 결론

VoIP 기반의 인터넷 텔레포니 서비스가 급격히 주목받고 있는 현시점에서, VoIP의 음질개선 문제는 아직도 중요한 과제라 할 수 있다. 본 논문은 우선 VoIP의 음질개선에 있어 소프트웨어적으로 구현할 수 있는 방안을 연구하기 위해서 VoIP의 음질 평가 요소와 Embedded VoIP 시스템을 분석한 후, 수신된 음성패킷의 변환처리 기법을 설계하였다. 그리고나서 실제 구현함에 있어서, 사용의 편리성과 비교적 안정적인 성능제공의 장점을 가지고 있는 Embedded VoIP 시스템을 채택하여 실험하였으며, 그 결과 음성패킷 변환기법이 음질개선에 좋은 영향을 미침을 입증하였다.

본 논문은 Embedded VoIP 시스템에서 음질개선을 위한 하나의 방안으로 음성패킷 변환기법을 제시하였으며, 본 논문에서 미처 다루지 못한 음성패킷 손실에 대한 대처방안과 Embedded VoIP 시스템에서의 최적화 방안 등에 대한 연구가 지속되어야 하겠다.

참고문헌

- [1] 유승화, “인터넷 전화 IP Telephony”, 전자신문사, 2002.1.
- [2] 정윤찬 외, “VoIP 네트워크에서 패킷 전송지연시간 변이현상을 없애주는 적응식 변이 제어기 제안 및 성능분석”, 한국통신학회논문지 '01-12 Vol.26 No.12C, 2001.12.
- [3] Jonathan Davidson, James Peters, “Voice over IP Fundamentals”, cisco systems, 2000
- [4] <http://kldp.org/HOWTO/html/VoIP-HOWTO/>