

# 디지털 방송에서 문자 전송을 위한 2차 한글 허프만 부호 설계

강신일, 윤자천, 설상훈

고려대학교 전자컴퓨터공학과

[sikang@mpeg.korea.ac.kr](mailto:sikang@mpeg.korea.ac.kr), [jcyoon@mpeg.korea.ac.kr](mailto:jcyoon@mpeg.korea.ac.kr), [sull@mpeg.korea.ac.kr](mailto:sull@mpeg.korea.ac.kr)

## Second Order Hangul Huffman Code Design for Data Broadcasting in Digital Broadcast

Shinil Kang, Ja-Cheon Yoon, Sanghoon Sull

Dept. of Electronics and Computer Engineering, Korea University

### 요 약

DTV 전송방식이 확정되고 디지털방송이 본격적으로 시행되면 주식, 환율, 일기예보 등의 실시간 정보와 인터넷 서비스 등의 데이터 방송 서비스가 더욱더 증대될 것이며 이에 따른 데이터 전송의 고효율을 위한 방법 또한 필요할 것이다. 따라서 본 논문에서는 디지털 방송에서 문자 전송을 위한 한글 허프만 부호 설계 기법으로 한글의 문법상의 특징과 한글코드와 ASCII 코드 길이의 차이를 고려한 2차 한글 허프만 부호 설계 기법을 제안한다.

### I. 서 론

최근 DTV(Digital Television) 전송방식을 둘러싼 논쟁이 미국식의 ATSC(Advanced Television Systems Committee, 이하 ATSC)방식으로 매듭지어졌다. ATSC 방식은 최대 19.4 Mbps까지 영상과 음향, 데이터 등의 정보를 전송할 수 있으며 시청채널정보 이외의 부가적인 정보를 전송하는 프로토콜로 PSIP(Program and System Information protocol, 이하 PSIP)을 정의 하여 사용한다. DTV는 특성상 고품질, 고음질의 동영상뿐만 아니라 추가적인 데이터를 전송할 수 있으므로 대화형 방송, 인터넷 서비스 등의 다양한 기능의 데이터 방송을 제공할 수 있다. 그러나 이와 같은 데이터 방송을 서비스하기 위한 전파자원은 한정되어 있기 때문에 주어진 자원을 최대한 효율적으로 이용해야만 한다.

ATSC 전송방식에서는 데이터 방송의 대부분인 텍스트 데이터 전송 규격을 PSIP 표준 내에 정의했다. 이 표준에 정의된 영문 텍스트 데이터의 압축을 위해 2차로 정의된 허프만 코드를 사용하며 타이틀과 내용에 따라 서로 다른 타이틀을 사용한다. ATSC 전송방식을 채택한 국내 DTV 방송 규격은 한글 코드 압축을 위한 타이틀이 마련되어

있지 않아 현재까지 압축을 하지 않은 한글 텍스트를 전송하고 있다. 따라서 한글 데이터가 차지하는 전송량의 증대는 심각한 문제를 발생시킬 수 있다.

본 논문에서는 2차 영문 허프만 코드의 기술의 장점을 살려 한글에도 적용 가능한 2차 한글 허프만 코드를 제안한다. 본 논문의 구성은 다음과 같다. II장은 한글을 표현하기 위한 표준조합형과 표준완성형, 유니코드를 대해 알아보고 III장에서는 PSIP에서 정의한 다중 문자열 구조에 대해 서술한다. IV장은 한글이 가지고 있는 특성과 이를 이용한 2차 한글 허프만 코드의 생성 기법에 대해 알아보고 V장은 실험 결과들, VI장은 최종 결론에 대해 기술한다.

### II. 한글코드의 종류

한글을 표현하기 위한 한글코드는 표준조합형(KSX 1001 : KSC 5601-1982), 표준완성형(KSX 1001 : KSC5601-1987), 유니코드(KSX 1005 1 : ISO/IEC 10646 1) 세 가지가 있다.