

# 이동통신환경에서 XHTML을 이용한 무선인터넷 문서변환기 분석 및 구현

백진영\* · 이종옥\* · 조성언\* · 조경룡\*

Analysis and Implementation of a Web Document Converter for Wireless Internet  
Use XHTML On Mobile Communication Environment

Jin-young Baek\* · Jong-ok Lee\* · Sung-eon Cho\* · Kyung-ryong Cho\*

## 요 약

본 논문은 사용자가 휴대용 단말기를 이용하여 기존의 웹에 접속시 웹서버내의 XHTML 문서를 WML 문서로 변환하는 도구의 설계·구현을 목적으로 한다. 사용자가 XHTML(또는 기존의 HTML문서)의 웹페이지에 접속해 정보를 요청하게 되면 본 문서 변환기는 XHTML 문서 구조를 인식하고, 기준이 되는 기준 텍스트 정보를 추출하여 이를 <select>와 <option> 태그를 이용하여 간결한 WML 문서로 재구성하게 된다. 또한 이에 맞게 기타 태그 및 이미지 등을 변환·치환 또는 제거함으로써 구조화시킨다. 그 결과 well-formed한 XHTML의 특성으로 인해 기존의 어떤 마크업언어보다 WML 문서와의 1:1 태그 매핑의 효율이 높음을 알 수 있었다.

## ABSTRACT

This paper is purposed in design and implement of a device which can convert XHTML documents in web-Server into WML documents when users access the web by using portable devices. Users access XHTML(so-called HTML) web page and ask for informations, this document converter recognizes of XHTML documents structures, reconstructs into simple WML documents by using <select> and <option> tags. Also document converters construct XHTML documents structure by conversion, transposition and deletion of other tags and images. As a result, because of the features of well-formed XHTML, it showed higher validity in 1:1 tags mapping with WML than any other Markup Language.

## 키워드

- HTML(eXtensible Hypertext Markup Language) : 웹 브라우저 상에 정보를 표시하기 위한 마크업 심볼들의 집합
- WML(Wireless Markup Language) : 무선 접속을 통하여 셀룰러폰 등에 웹페이지가 표시될 수 있도록 해주는 언어

## 1. 서 론

무선인터넷의 관점은 실질적인 무선인터넷 구현에 있어 기존의 인터넷 정보를 얼마나 자유자재로 이용할 수 있는가이다. 그러기 위해서는 먼저 유·무선 마크업 언어들의 호환 내지 통합성

에 그 열쇠가 있다고 할 수 있다. HTML은 작은 휴대 장치들의 웹 페이지들을 표시하는데 필요한 메모리와 네트워크 대역폭을 갖추기가 힘들다. 그리고 그런 장치들은 특별한 코딩을 필요로 한다. 그러나 어떠한 언어를 사용하는 사람이 많으면 많을수록 그 언어의 통신 수단으로서의 유용성은 더욱 커지게 된다. HTML은 수많은 사람들이 사

용하고 있다. XML이 아무리 우월한 언어라고 해도, 그 사람들 모두를 단번에 XML로 돌려세우는 것은 불가능하다. 이에 대한 W3C의 해결책이 바로 XHTML이다[4]. 따라서 이 마크업 언어에 대한 많은 연구가 필요하다.

본 연구는 이러한 점을 인지하여 XHTML의 특성에 대한 연구와 함께 기존 무선인터넷을 위한 WML문서의 상호 자유로운 변환을 하기 위함이며, 그 결과 well-formed한 XHTML의 특성으로 인해 기존의 어떤 마크업언어보다 WML 문서와의 1:1 태그 매핑의 효율이 높음을 알 수 있었다.

## II. 본 론

### 1. 문서변환 기술의 필요성

국내의 이동통신사는 WML, HDML, mHTML 중 하나를 자사의 표준으로 채택하고 있다(표1참조). 따라서 국내 콘텐츠 제공자는 무선인터넷 접속이 가능한 모든 이용자를 대상으로 콘텐츠를 제공하기 위해서 세 가지 형식에 맞춘 문서를 모두 갖추어야 한다. 이것은 비효율적이다. 또한 아무리 우수한 언어가 개발되었다라도 기존의 언어 체계를 무시하는 것 역시 비현실적이다. 그러므로 미래지향적으로 well-formed한 언어를 준비하면서 기존의 문서를 재사용하여 단말기 환경에 맞는 WML[2] 문서로 변환하는 기술이 필요하다.

표 1. 국내 이동통신사별 무선인터넷 마크업언어 채택현황

서비스	이동통신회사	서비스명	컨텐츠 양식	웹서버	
PCS	한국통신 프리텔	016	magic	HTML	ME
		018	magic	HTML	ME
	019 LG 텔레콤	ez Web	HDML/WML	WAP	
Cellular	011 SK Telecom	n.top	WML	WAP	
	017 SK신세기통신	itouch 017	HDML/WML	WAP	

### 2. XHTML의 필요성

XHTML은 웹 콘텐츠를 개발하기 위한 현재 표준인 HTML이 무선 디바이스들에게 전혀 쓸모없다는 사실에서 출발한다. 즉 HTML을 다른 태그 집합들과 결합될 수 있는 작은 조각들의 집합으로 재정의한 것이다. XHTML 문서는 일반적인 데스크탑 컴퓨터뿐만 아니라 휴대폰 등의 핸드헬드 장치에도 이용이 되어야 하므로, 문서를 작성할 때에는 그러한 장치들에서도 문서의 내용이 제대로 사용될 수 있는지에 대해 고려해야 한다.

HTML은 짧은 글씨를 위한 <b>나 가운데 정렬을 위한 <center> 같은 표현적인 요소들을 추가하면서 발전해 왔으나, 이동단말이 그러한 요소들을 제대로 처리하기에는 용량이나 능력면에서 한계가 있다. 이를 위해 W3C에서는 문서의 표현(문서가 표시되는 방식)과 문서의 내용(문서가 전달하고자 하는 정보)을 분리시키기 위해서 CSS (Cascading Style Sheets)라는 것을 만들어냈다[7]. 즉 문서의 내용은 건드리지 않고도 서로 다른 장치들에 대해 서로 다른 표현 스타일을 적용할 수 있게 된 것이다. CSS는 표시규칙들의 모음으로 사용자 에이전트가 문서에 스타일을 적용하는 데 필요한 정보를 제공한다. 이 기능으로 마크업언어는 보다 엄격화 될 수 있다.

### 3. 문서변환기 작동원리 및 문서변환

문서변환기의 작동을 위해 콘텐츠(URL) 요청에서 응답까지 크게 세 부분으로 나눈다(그림1).

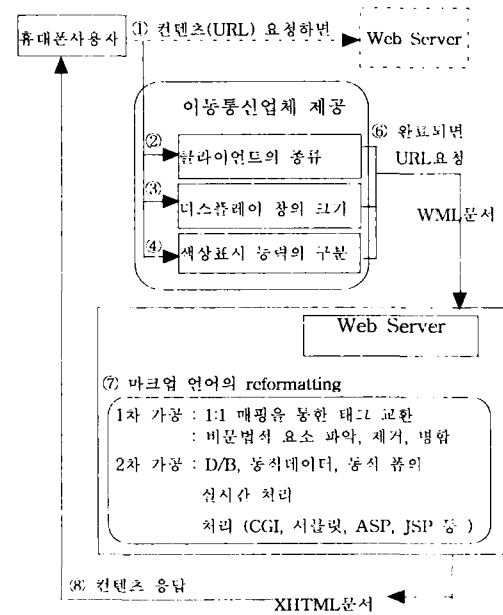


그림 1. XHTML의 WML 문서로의 변환 과정

어떤 단말기 사용자가 콘텐츠를 요청하게 되면, 해당 웹서버는 요구에 응답하기 위해 먼저 이동통신 업체의 장비를 경유하게 된다. 이동통신 업체는 접속해 들어오는 휴대폰의 고유 ID(휴대폰번호 등)를 기반으로 기기 정보(클라이언트의 종류, 화면의 크기, 색상표시능력 등)를 제공해 준다. 이를 기반으로 웹 서버는 클라이언트 분류 및 문서를 재구성(reformatting)하게 된다. 리포매팅 시 문서변환기는 2단계를 거치게 된다. 먼저 태그의 1:1 매핑을 통한 태그 교환이 이루어지게 되는

데 교환시 비문법적 요소를 파악하여 제거·통합하는 작업을 거쳐야 하며, 2차 가공시 CGI, 서블릿, ASP, JSP 등에 의한 동적 데이터, 동적 폼을 실시간 처리해야 하는데 이를 상용화하기 위해서는 개별화에 대한 각각의 지원이 필요하므로 현실상 어렵다. 본 논문에서도 태그 요소의 매핑을 통한 1차적인 가공적 문서변환 틀 설계를 목표로 한다.

본 논문은 문서 변환 첫 단계로서 XHTML 문서를 읽어들이며 객체화하였다. 마크업 언어는 태그들 사이의 포함 관계가 성립되어 있기 때문에 (그림2)와 같이 구조상 트리 형식으로 나눌 수 있다.

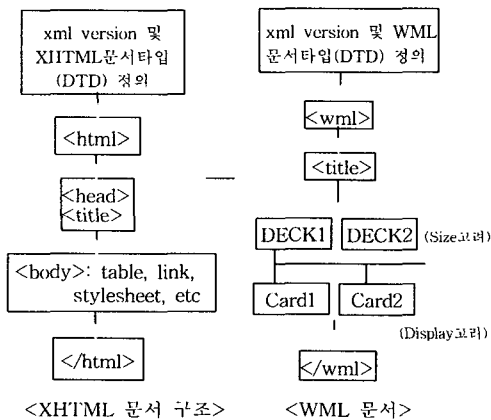


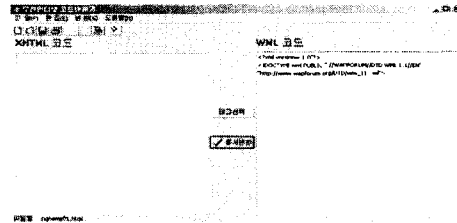
그림 2. 트리구조에 따른 문서변환

이와같이 XHTML 문서와 WML 문서를 상호 변환하는데 있어 가장 핵심이 되는 부분이 바로 body 영역이다. XHTML의 body 영역은 마크업 언어의 순차적 특성에 따라 WML의 Deck과 card로 XHTML의 문서를 분리한다. 이때 웹문서를 WML 문서의 글자 허용기준[6]인 1200바이트를 넘지 않는 한도 내에서 데이터를 잘라내어 각각의 WML문서를 만들어 낸다. XHTML의 <Title>은 WML 문서의 <Title> 페이지로 첫 페이지를 구성하게 된다.

두 번째로 1200바이트를 기준으로 재구성된 WML 페이지는 <select>와 <option> 태그를 이용하여 각각의 페이지들을 링크한다. 그후 이들을 태그에 따라 1:1 매핑한다. 이때 100 여개의 XHTML 태그들은 30여개의 WML 태그 매핑시 문서의 요소가 개념이 같거나 유사한 개념으로 N:1 매핑하게 된다. 기존 HTML문서의 경우 태그의 관대한 처리로 인해 태그 개폐성의 불확실성으로 문서변환시 미세한 작업이 필요했으나, XHTML의 경우 확실한 태그의 개폐성으로 인해 그러한 작업은 불필요하기 때문이다. 이는 기존의 어떤 마크업언어보다 직접적인 태그 매핑의 효율이 높아지게 함을 알 수 있다. 하지만 이는 태그 매핑의 관계만 논의한 것으로 그 외 script 및 image 처리의 보다 깊은 연구가 필요하다.

#### 4. 무선인터넷 변환기 구현

본 변환기는 여러 메뉴와 함께 두 개의 창으로 구성되어있다.

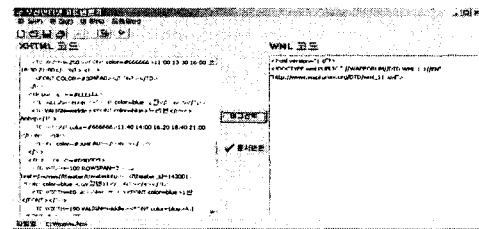


메뉴에는 파일을 읽거나 저장하는 부분, 편집 영역의 복사·붙여넣기 부분, 문서변환 영역의 태그 추출과 그에 따른 문서변환 부분, 검증 기능을 가진 변환 부분으로 구성된다. 그리고 마지막으로 이 프로그램의 목적과 정보를 가진 도움말 부분으로 구성되어 있다.

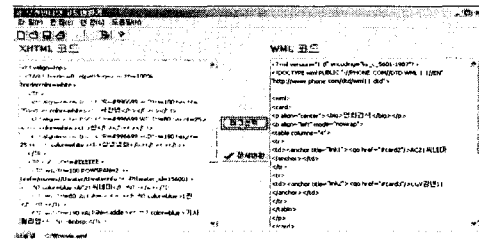
다음은 XHTML에서 WML로의 문서변환기를 이용한 변환과정이다.

가. 파일열기를 선택하여 XHTML(확장자: html 또는 htm) 파일을 연다.

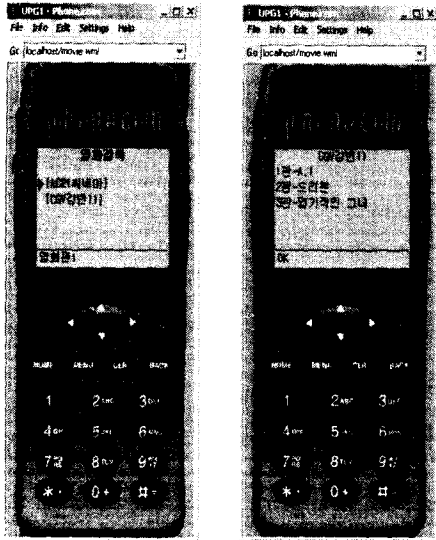
나. 가운데 있는 태그 선택 탭을 누른다. 이 탭을 누르면 변환기는 다음과 같이 DataBase화된 태그들을 선택하게 된다.



다. 다음으로 문서변환 탭을 누르게 되면 기존 XHTML 태그는 WML 태그로 대치되게 된다.



라. 마지막으로 위의 과정을 겪어 변환이 완성 되면 변환메뉴의 검증을 이용하여 WML 프로그램을 UP.SDK 4.0을 이용하여 검증한다.

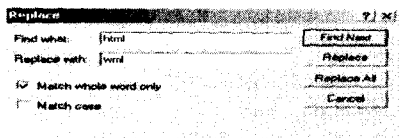


서변환기를 개발하였다. 그 결과 well-formed한 XHTML의 특성으로 인해 기존의 어떤 마크업언어보다 WML 문서와의 1:1 및 N:1 태그 매핑 효율이 높음을 알 수 있었다.

그러나 이는 아직 초기단계이며 텍스트 위주의 변환에 머무르고 있다. 또한 무선인터넷 문서 변환에 있어 개별화에 대한 지원이 이루어져야 하므로 상용화할 수 있는 제품을 제작하기란 한계가 있다. 또한 무선인터넷 변환도구는 기존 유선용 콘텐츠의 빠른 무선 서비스가 가능하며 필요한 정보에 즉시 접근이 가능한 반면 결과물로 나오는 콘텐츠의 질에 대한 보장이 불가하다.

이를 위해 콘텐츠 변환기의 핵심기능을 보완하여, 다양한 애플리케이션 프로그램들과의 연동은 물론 script 및 image 처리, CGI·ASP등 데이터베이스와 관련된 안정적인 작업에 관한 연구가 수행되어야 한다. 그러기 위해선 이를 위한 무선인터넷 문서 변환 구축에 맞는 개별적인 지원이 뒤따라야 한다. 무엇보다 향상적이고 안정된 무선인터넷을 위해서 XHTML에 대한 연구가 계속되어야 하며 XML로의 통합 연구가 계속 되어야 할 것이다.

마. 위의 검증을 통해 에러 발생시 해당 태그는 편집메뉴의 바꾸기를 선택하여 변환시킨다.



<시스템의 설계 환경>

- O/S 환경 : Windows 2000 Advanced Server
- IIS Server
- DataBase : ODBC
- 구현 프로그램 : DELPHI 5.0 Enterprise
- WML 문서 검증 툴 : UPSDK 4.0

III. 결 론

휴대폰의 작은 표시 창으로 웹을 사용한다는 것은 또 다른 도전이라고 할 수 있다. 문제는 기존의 HTML 문서가 이를 처리할 만한 능력을 가지고 있지 않다는 것이다. 그러나 XML의 well-formed한 특성을 따르고 있는 XHTML은 이를 훌륭히 수행해 줄 수 있다. 본 연구는 이러한 필요성에 맞춰 XHTML의 특성에 초점을 맞춰 문

참고문헌

- [1] 박영충, "모바일 환경에서 무선 인터넷을 위한 WAP 필터링 시스템의 설계 및 구현", 세종대학교 대학원 전산학과, 2000, 12.
- [2] [WML]"Wireless Markup Language", WAP Forum, April 30, 1998, URL : <http://www.wapforum.org/>.
- [3] [HTML]"HyperText Markup Language", W3C, December 24, 1999, URL : <http://www.w3.org/TR/html401/>.
- [4] [XHTML] "eXtensible HyperText Markup Language", W3C, January 26, 2000, URL : <http://www.w3.org/TR/xhtml1/>.
- [5] [XHTML-Basic] "eXtensible HyperText Markup Language-Basic", W3C, December 19, 2000, URL : <http://www.w3.org/TR/xhtml1-basic/>.
- [6] 김대송외, "무선인터넷 어플리케이션 프로그래밍", 삼양출판사, 2000.
- [7] Frank Boumphrey 외, Beginning XHTML, 정보문화사, p.55-71, 2000.