

논문 2005-42CI-6-1

# 무선 인터넷을 위한 프레임 지원 HTML 변환기의 설계 및 구현

(Design and Implementation of an HTML Converter Supporting Frame  
for the Wireless Internet)

한진섭\*, 박병준\*\*

(Jin Seop Han and Byung Joon Park)

## 요약

본 논문은 무선 응용 프로토콜 환경에서 무선 인터넷 액세스를 위한 HTML 변환기의 설계 및 구현에 대해 기술한다. 구현된 HTML 변환기는 콘텐츠 변환, 변환 규칙 집합, WML 파일 생성, 프레임 콘텐츠 재구성 모듈로 구성된다. 일반 텍스트 콘텐츠의 경우, 변환 규칙 집합을 참조하여 콘텐츠 변환 모듈에서 일대일 태그 매핑을 통해 WML 콘텐츠로 변환이 되며, 프레임으로 구성된 콘텐츠는 프레임 콘텐츠 재구성 모듈에서 최초로 받은 프레임 셋 소스를 파싱하여, 포함되어 있는 모든 파일 이름들을 가지고 요구 메시지를 구성하여, 파일 개수만큼 다시 서버에 연결해서 각각의 문서들을 받아 처음에 받은 HTML 문서에 연결시키고, 일련의 재구성 과정을 거친 후, 테이블 형태의 WML 콘텐츠로 변환이 된다. 이미지 맵이 포함된 콘텐츠는 이미지 맵 관련 태그를 파싱한 후, 링크되어 있는 HTML 문서의 이름을 추출해서 WML 콘텐츠 데이터로 대체하고, 해당 콘텐츠로 링크시킴으로써 변환이 가능하게 한다. 제안한 프레임 콘텐츠 변환 방법은 기존 상용 변환기와는 달리 사용자의 편의와 상호 운용성이 고려된 변환 방법이며, 이미지 맵의 경우 상용 변환기에서는 지원되지 않는 부분이다.

## Abstract

This paper describes the implementation of HTML converter for wireless internet access in wireless application protocol environment. The implemented HTML converter consists of the contents conversion module, the conversion rule set, the WML file generation module, and the frame contents reformatting module. Plain text contents are converted to WML contents through one by one mapping, referring to the converting rule set in the contents converting module. For frame contents, the first frameset sources are parsed and the request messages are reconstructed with all the file names, reconnecting to web server as much as the number of files to receive each documents and append to the first document. Finally, after the process of reformatting in the frame contents reformatting module, frame contents are converted to WML's table contents. For image map contents, the image map related tags are parsed and the names of html documents which are linked to any sites are extracted to be replaced with WML contents data and linked to those contents.

The proposed conversion method for frame contents provides a better interface for the users convenience and interactions compared to the existing converters. Conversion of image maps in our converter is one of the features not currently supported by other converters.

**Keywords:** HTML Converter, Wireless Application Protocol, WML, Frame Contents, WML's Table Contents

## I. 서론

최근 무선 단말기 사용자의 급속한 증가와 네트워크 인

프라의 구축 및 발전으로 인하여 기존 유선망에서와 함께 무선망에서 데이터를 서비스 받고자하는 요구가 급증하고 있다. 그러나 일반 데스크탑 컴퓨터나 워크스테이션을 기준으로 제작된 기존의 웹 콘텐츠를 그대로 무선 단말기 사용자에게 제공하는 것은 낮은 대역폭, 적은 메모리, 작은 디스플레이 화면 등의 제약 사항 등을

\* 학생회원, \*\* 정회원, 광운대학교 컴퓨터과학과  
(Department of Computer Science, Kwangwoon University)

접수일자: 2005년5월4일, 수정완료일: 2005년11월3일

고려할 때, 무선 단말기에 적합하지 않으며, 무선 인터넷 콘텐츠를 새롭게 제작하는 것은 추가로 엄청난 시간과 비용이 들게 된다.<sup>[7]</sup> 따라서 기존의 웹 콘텐츠를 무선 인터넷 환경에 적합하도록 변환해서 무선 단말기 사용자에게 제공하는 변환기가 필요하다. 즉 변환기는 기존 유선망에서 HTML로 되어 있는 웹 콘텐츠를 WML(Wireless Markup Language : 무선 마크업 언어)<sup>[7]</sup>과 같이 무선 단말기에서 지원하는 마크업 언어로 변환하는 것이다.

WML이란 무선 네트워크를 통해 무선 단말기와 같은 소형 장치로 보내질 애플리케이션을 만들기 위한 마크업 언어이다. 이러한 무선 마크업 언어는 WAP Forum(Wireless Application Protocol Forum : 무선 애플리케이션 프로토콜 포럼)에 의해 개발된 공개 표준이다. 그밖에 무선 단말기에서 지원 가능한 마크업 언어로는 mHTML, HDML, cHTML 등이 있으며 최근에 WAP 2.0에서 제안한 XHTML이 있다.

위와 같은 WAP이라고 하는 무선 인터넷 환경에서 기존의 HTML로 되어 있는 웹 콘텐츠를 WML 등과 같은 무선 마크업 언어로 변환해서 무선 단말기 사용자에게 제공하는 것이 HTML 변환기이다.

특히, 본 논문에서 제안한 HTML 변환기는 HTML 프레임 콘텐츠의 경우에 효과적으로 변환하도록 설계하였다. 대부분의 홈페이지에서 좌측 프레임에 메뉴 콘텐츠가 위치하게 되고, 우측에는 여러 메뉴에 대해서 링크되어 있는 HTML 문서들이 보여주는 등의 사용자 인터페이스를 제공한다. 그러나, 기존의 HTML 변환기는 최초 화면이 해당 페이지의 콘텐츠가 나오는 것이 아니라, HTML 프레임 태그의 속성 값이 출력되어, 콘텐츠의 내용과 상관없이 임의의 콘텐츠를 선택했을 때, 해당 콘텐츠에 접근이 가능하다.<sup>[8]</sup> 이는 유선의 HTML 프레임에서 제공하는 사용자 인터페이스와 상호 운용성을 제공하지 못하는 것이다. 따라서 본 논문에서는 HTML 프레임 변환에 있어서 무선 단말기의 최초 화면에서부터 콘텐츠를 보여주도록 하고, 기존 상용 서비스와 차별화하는 사용자 인터페이스와 상호 운용성을 제공하는 프레임 지원 HTML 변환기를 설계하고 구현하였다.

또한, 현재 많은 홈페이지에서 이미지에 해당 콘텐츠를 링크시키는 이미지 맵 콘텐츠를 많이 사용하고 있으나, 기존 상용 서비스에서는 변환 자체가 불가능하다. 이

를 지원하기 위해 이미지에 링크되어 있는 HTML 문서를 보여주도록 변환하여 가능하도록 했다. 일반적인 텍스트 콘텐츠는 태그 변환 방법을 이용하였다.

본 논문의 구성은 II장에서 HTML 변환기 관련 연구에 대해 기술하고, III장에서 본 논문에서 제안하고, 상용 서비스와 차별화된 HTML 프레임 변환을 지원하는 HTML 변환기의 구조와 각 구성 요소들에 대해 기술한다. IV장에서는 실행 결과 및 평가를 하고, V장에서 결론을 맺는다.

## II. 관련 연구

본 장에서는 HTML 변환기 관련 연구에 대하여 기술한다.

오늘날 빠르게 성장하는 모바일 시장에서의 운전자와 제조업체들의 차별화 되고, 빠르고, 유연한 서비스 창출의 도전을 충족시켜주기 위하여, WAP은 전송 세션, 어플리케이션 레이어 분야의 프로토콜을 규정하고 있으며, 인터넷 상의 정보를 빠르게 검색, 표시하기 위하여 이동전화망과 인터넷 사이에 그림 1과 같이 게이트웨이를 두는 형태로 실현되며 독특한 프로토콜을 사용하고 있으며, 때문에 기존 인터넷 표준 프로토콜(HTTP, TCP등)과 호환성은 없다.

HTML과 호환성이 없으므로 기존의 HTML 베이스의 인터넷 사이트를 접속하기 위하여 WML로 변환하기 위한 게이트웨이가 필요하다. 전용 게이트웨이에서 수행하는 작용은 프로토콜을 변환해주는 것과 전용 콘텐츠를 전송하는 것 이외에 기존의 HTML 베이스의 콘텐츠들을 WAP에서 규정하고 있는 WML이라는 언어로 변환하여 주는 것이다. 이러한 게이트웨이에 HTML 변환기 모듈이 탑재되어 동작할 수 있고, 웹 콘텐츠를 WML로 변환하는 HTML 변환기의 기본적인 구현 방법은 HTML 문서의 레이아웃과 콘텐츠를 동시

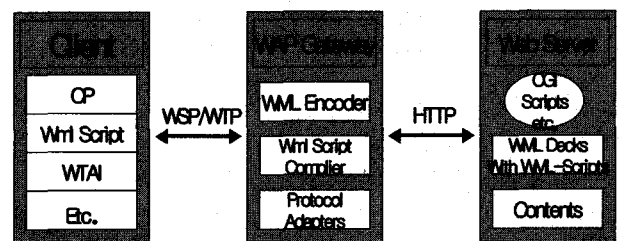


그림 1. WAP 구조  
Fig. 1. WAP Architecture.

표 1. HTML 변환기의 비교  
Table 1. Comparison of HTML Converters.

Automated Converter	Phone.com WAP Gateway	텍스트 데이터의 WML 변환, 1.4kbyte의 양적 분할 변환
	Argo Actigate	WML, CHTML, HDML, XML 변환, 이미지 압축 변환, 최상 기능 제공
	인사이드코리아의 IWAP	Hdml, Wml, mHtml로 변환
	HTML Filter	WML로의 양적 분할을 통한 변환
Configurable Converter	IBM Webspherer	HTML, HDML, WML, cHTML 변환가능. 이미지 압축 변환
	SpyGlass	텍스트와 이미지 압축 변환기능. WML과 HTML 계열로 변환
	Oracle Wireless Edition	HTML 파일을 파싱한 후, XML 문서를 생성하고, 클라이언트 단말기에 적합한 xsl를 이용해서 WML, tinyHTML, HTML, VoXML 등의 마크업 언어들로 변환
	엠펙소프트의 M-Converter	WML, mHTML, HDML, cHTML, XML로의 변환. 이미지 변환

에 변환해주는 HTML 재구성<sup>[3][9]</sup>, 레이아웃은 손대지 않고 HTML 문서 태그만 WML 태그로 변환해주는 태그 변환<sup>[1][10]</sup>, HTML과 WML 태그가 혼재되어 있는 인터넷 문서에서 WML 문서만을 찾는 웹 클리핑<sup>[11]</sup> 등 3가지 방법이 있다. 이 3가지 방법 중 태그 변환 방법과 웹 클리핑 방법은 개발이 상대적으로 쉽기 때문에 이미 상용화되어 있으며 HTML 재구성 방법은 최근 연구가 한창 진행 중에 있다.

이러한 HTML 변환기는 여러 기업이나 학교에서 연구 및 개발을 진행하고 있는 상황이다. 다음은 HTML 변환기의 사례연구로 제품들의 특징을 살펴보고 지원되는 기능들을 표 1에 나타내었다. 표 1과 같이 HTML 변환기는 크게 고정적으로 설정된 변환 규칙에 따라 자동 변환을 수행하는 Automated Converter와 관리자가 설정된 변환 규칙의 수정을 할 수 있는 Configurable Converter로 구분 할 수 있다. Argo Actigate, IBM Webspherer, 그리고, SpyGlass는 이미지 압축 변환이 가능하며, Oracle Wireless Edition, 엠펙소프트의 M-Converter 등은 xsl를 이용해서 다수의 무선 마크업 언어로의 변환이 가능한 특징이 있다.

기타 국내 상용 서비스(SKT, KTF, LGT 등)는 일반 콘텐츠에 경우 정상적인 변환을 수행하지만 프레임 콘텐츠 변환에 있어서는 그림 2와 같이 최초 화면이 해당 페이지의 콘텐츠가 출력되지 않고, HTML 프레임 태그

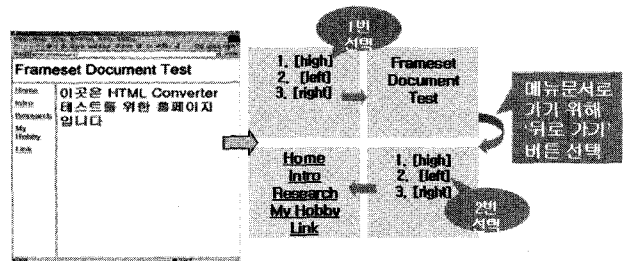


그림 2. 상용 서비스의 프레임 변환  
Fig. 2. Frame Conversions in the Commercial services.

의 속성 값이 출력된다. 대부분의 HTML 프레임 홈페이지는 좌측 프레임에는 메뉴 콘텐츠가 위치하게 되고, 우측에 여러 메뉴에 대해서 링크되어 있는 HTML 문서들이 타겟 프레임에 하나로 고정되어 출력되는 방식으로 되어 있는 반면에 상용 서비스의 프레임 콘텐츠 변환의 경우, 프레임 태그 속성 값이 출력됨으로써 해서 콘텐츠의 내용과 상관없이 임의의 콘텐츠를 선택했을 때, 해당 콘텐츠에 접근하도록 되어 있는 것이다. 만약에 그림 2와 같이 무선 단말기 사용자가 메뉴 화면(Home, Intro, Research, My Hobby, Link로 구성된 콘텐츠)에 접근하고자 한다면 3개의 HTML 프레임 속성 값이 출력된 화면에서 내용 파악 없이 임의의 속성 값(1번 선택)을 선택하게 되고 원하는 콘텐츠가 아니면 단말기의 "뒤로 가기" 버튼을 클릭한 후에 다시 3개의 속성 값 중에 하나를 내용 파악 없이 임의로 선택(2번 선택)해야 비로소 유선의 HTML 프레임 페이지 좌측에 위치하고 있는 메뉴 문서에 접근할 수 있게 되는 것이다. 이는 유선의 HTML 프레임에서 제공하는 사용자 인터페이스와 상호 운용성이 전혀 고려되지 않은 변환 방법이다.

따라서, 본 논문은 프레임 콘텐츠를 단순 변환 할 것이 아니라, 기존 설계와 차별화하여 그림 3과 같이 프레임 콘텐츠 Reformatting 모듈을 두어서 WML의 테이블 콘텐츠로 변환하여, 유선의 HTML 프레임과 유사한 사용자 인터페이스와 상호 운용성을 무선의 단말기 사용자에게 제공하고자 설계하였다.

추가로, 현존 하는 많은 HTML 웹 페이지에서는 이미지에 해당 콘텐츠를 링크시키는 이미지 맵을 많이 사용하고 있으나, 기존 상용 서비스에서는 지원되고 있지 않다. 이를 지원하기 위해 이미지 맵 관련 태그를 파싱한 후, 링크되어 있는 HTML 문서의 이름을 추출해서 WML 콘텐츠 데이터로 대체하고, 해당 콘텐츠로 링크 시킴으로써 변환이 가능하게 했다.

### III. HTML 변환기 설계

본 장에서는 기존 HTML 콘텐츠를 무선 단말기 사용자에게 적합한 WML 콘텐츠로 변환하도록 설계한 HTML 변환기의 구조와 이를 구성하는 구성 요소들의 동작 원리 및 기능에 대해 기술한다. 1절에서 시스템 개요를 기술하고, 2절에서 4절은 일반 콘텐츠의 변환에 대해서 기술하며 5절에서 HTML 프레임 변환에 있어서 무선 단말기의 최초 화면에서부터 콘텐츠를 보여주도록 하고, 기존 상용 서비스와 차별화하는 사용자 인터페이스와 상호 운용성을 제공하도록 한 프레임 콘텐츠 재구성 모듈에 대해서 기술한다. 마지막 6절에서는 역시 기존 상용 서비스에서 제공하지 않는 이미지 맵 변환에 대해서 기술한다.

#### 1. 시스템 개요

본 논문에서 설계하고 구현한 HTML 변환기의 구조는 그림 3에 나타낸 바와 같이 콘텐츠 변환, 변환 규칙 집합, WML 파일 생성, 프레임 콘텐츠 재구성 모듈로 구성되어 있다. 무선 단말기 사용자는 WAP 게이트웨이에 WML request를 보내고 WAP 게이트웨이는 이를 HTML request로 변환하여 웹 서버로부터 HTML 콘텐츠 파일을 읽어 온다. 읽어온 HTML 콘텐츠는 구현한 HTML 변환기에서 일반 콘텐츠와 프레임 셋 콘텐츠를 분류한다. 일반 콘텐츠의 경우에 HTML 콘텐츠를 콘텐츠 변환 모듈에서 SP(SGML Parser) 파서를 이용하여 태그와 콘텐츠 데이터를 파싱해서 토큰화된 태그와 데이터를 트리 구조로 노드에 저장<sup>[7]</sup>하고, 변환

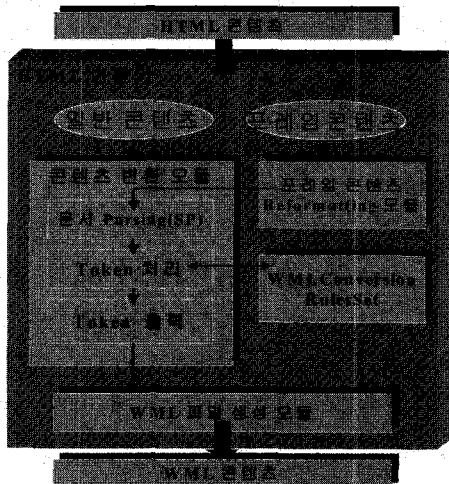


그림 3. HTML 변환기 구조  
Fig. 3. HTML Converter Architecture.

규칙 집합을 참조하여 각각의 HTML 태그들을 WML 태그로 일대일 매핑을 시켜서 변환하는 토큰 처리 과정을 거쳐서 WML 태그로의 토큰 출력이 된다<sup>[2][7]</sup>. 다음으로 WML 파일 생성 모듈에서 WML 헤더와 DTD 정보를 추가하여서 최종적으로 WML 콘텐츠를 생성하여 무선 단말기 사용자에게 제공한다. 콘텐츠가 프레임 셋 콘텐츠인 경우는 기존 상용 서비스와 달리, 프레임 콘텐츠 재구성 모듈을 통하여 태그 변환 규칙과 콘텐츠 데이터를 HTML 프레임과 유사한 사용자 인터페이스가 지원되도록 변환한다. HTML의 이미지 맵 콘텐츠 역시 기존 태그 변환 규칙과 데이터의 수정을 통해서 변환이 가능하다. 다음은 HTML 변환기의 각 구성 요소들에 대해 상세히 기술한다.

#### 2. 콘텐츠 변환 모듈

콘텐츠 변환 모듈은 SP(SGML Parser)를 이용하여 태그와 콘텐츠 데이터를 파싱해서 토큰화된 태그와 데이터를 트리 구조로 노드에 저장하고, 변환 규칙 집합을 참조하여 각각의 HTML 태그들을 WML 태그로 변환시킨다. 여기서는 SGML Parser의 개요와 객체지향 SGML 파싱 킷인 SP에 대해 기술한다.

##### 가. SGML Parser

실제 작성된 문서는 검증이 완료된 문서형 정의부(DTD)와 적합한지를 점검하여야 하는데, 이런 문서 검증은 SGML 파서에 의해서 처리된다. SGML 문서를 검증하는 SGML 파서는 우선 SGML 문서형 정의부를 입력으로 이것이 표준 SGML 규칙에 맞게 선언되어 있는지를 분석한다. 이 분석에 따라 문서형 정의부의 적합성이 확인되면 이 문서형 정의 정보는 이미 작성된 문서와의 적합성을 점검하는 단계에 적용되어야 하고 또 SGML 응용모듈에서 이용할 수 있어야 한다. 일반적으로 SGML 파서는 2가지형으로 구분된다. 첫 번째형은 간단한 SGML 파서로 SGML 문서내의 마크업만을 인식한다. 둘째, 유효한 SGML 파서는 기록할 만한 마크업 오류가 존재한다면 이를 발견하고 알려주는 파서로, 마크업 인식뿐 아니라 검색 및 오류 출력기능을 가지고 있다.

##### 나. SP (SGML Parser Toolkit)

실제 SP는 SGML Parser를 가리키기도 하지만,

SGML 문서 파싱과 엔터티 조작을 위한 공개 객체 지향 툴킷이기도 하다. SP는 Unix, MS-DOS, Windows, OS/2 등의 운영 체제 환경에서 사용 가능하며, 문서를 파싱하고 엔터티를 조작하는 API를 제공하며 다음과 같은 클래스를 제공한다.<sup>[2]</sup>

- SGMLApplication : 어떤 SGML 문서의 태그 정보에 일치하는 이벤트를 핸들하는 객체.
- EventGenerator : SGMLApplication에 의해서 핸들된 이벤트의 시퀀스를 발생시키는 객체.
- ParserEventGeneratorKit : SP를 사용해서 이벤트를 발생시키는 EventGenerator를 만들.

다수의 이벤트 중 중요한 이벤트 타입은 다음과 같다.

- StartElementEvent : 시작 엘리먼트 즉, 문서의 시작 태그에 대해서 발생하는 이벤트로 Position pos, CharString gi, Element::ContentType contentType 등의 멤버를 가짐.
- EndElementEvent : 종료 엘리먼트 즉, 문서의 종료 태그에 대해서 발생하는 이벤트로 Position pos, CharString gi 등의 멤버를 가짐.
- DataEvent : 문자 엘리먼트 즉, 문서의 문자 데이터에 대해서 발생하는 이벤트로 Position pos, CharString gi 등의 멤버를 가짐.
- SdataEvent : 콘텐츠에서 내부 sdata 엔터티로의 레퍼런스에 대해서 발생하는 이벤트로 Position pos, CharString text, CharString entityName 등의 멤버를 가짐.
- ExternalDataEntityRefEvent : 외부 data 엔터티로의 레퍼런스에 대해서 발생하는 이벤트로 Position pos, Entity entity 등의 멤버를 가짐.

즉, HTML 변환기의 콘텐츠 변환 모듈은 HTML 콘텐츠의 태그를 파싱해서 토큰화 한 후, 위에서 언급한 EventGenerator 클래스에서 해당 이벤트를 발생시켜서 변환 규칙 집합을 참조, 반영하여 WML 콘텐츠로 변환이 가능하도록 한다.

### 3. 변환 규칙 집합

HTML 콘텐츠를 WML로 변환 시 가장 간단한 방법은 일대일 태그 매핑을 하는 방법이다. 그러나 WML의 태그가 모든 HTML의 기능을 제공하지 못하므로 HTML 태그에 따라서 효과적인 변환을 수행할 수 있는 변환 규칙을 규정할 필요가 있다. 또 고려해야 하는 사항은 일반적인 무선 단말기는 매우 작고 제한적인 화면과 제한적인 폰트를 가진다. 이것은 무선 단말에서 스타일 시트를 지원하는 것이 어렵다는 것을 의미한다<sup>[3][4][5][6][9]</sup>. 따라서 기존의 웹 문서 내의 스타일 관련 정보와 다양한 폰트를 표현하는데 사용되는 텍스트 포매팅과 관련된 요소 및 속성은 제거되어야 하며, 대부분의 무선 단말은 스크립트 언어를 지원하지 못하기 때문에 이벤트 및 스크립트와 관련된 요소 및 속성은 제거되어야 한다<sup>[4][5]</sup>. 그리고 웹 문서에서 다양한 외부 객체를 지원하기 위해 사용되는 객체 및 자바 애플릿 역시 이것을 지원 할 수 있는 무선 단말기가 거의 존재하지 않기 때문에 제거되어야 할 것이다. 위와 같은 사항들을 고려한 변환 규칙 설정 방식에는 여러 가지 방식이 가능한데, HTML 태그에 따라 단순 지원 가능한 WML 태그로 변환할 수도 있고 HTML 태그와 함께 속성도 같이 변환해야 하는 경우도 있다. 이외에 태그는 변환되지 않고 태그의 속성이나 태그 내부의 데이터만을 변환하는 경우, 또는 태그와 콘텐츠 데이터 모두를 삭제하는 방식도 태그 변환 규칙의 예가 될 수 있다. 그러나

표 2. HTML 태그 분류  
Table 2. Classification of HTML Tags.

분 류	HTML 태그
valid(25개)	html, body, h1, h2, h3, h4, h5, br, b, p, I, u, center, strong, ...
validAttributes(2개)	<a href =“x” target=“y”>, <area >
validAttributesAsData(1개)	<img alt=“x”> => [IMG]-x
disable(44개)	map, table, span, font, overlay, noscript, noframes, spacer, ...
discard(29개)	applet, option, script, style, input, media, form, select, ...
Invalid	frameset, frame, ...

HTML 태그의 경우 100여 개에 이르는 각 태그에 대한 변환 규칙을 모두 정의하기에는 역부족이다. 따라서 본문에서는 표 2에서 보는 바와 같이 태그에 대한 변환을 크게 6개의 클래스로 구분하여 각 클래스에 대한 다른 변환 규칙을 적용함으로써 보다 간단하고 빠른 변환을 수행하도록 했다.

각각의 클래스를 설명하면 다음과 같다.

- valid : WML로의 변환 가능한 태그가 지원되며, 단순히 태그만을 이용한 변환을 한다.

예) <html> -> <wml>, <body> => <card><p>

- validAttributes : WML로의 변환 가능한 태그가 지원되며, valid와는 다르게 HTML 태그내부의 특정 Attribute를 함께 변환한다.

예) <a href = "x" target = "y"> => <a href = "x">

- validAttributesAsData : validAttributes와 마찬가지로 HTML 태그내부의 Attribute를 함께 이용 변환하지만 변환된 결과가 WML 태그가 아니고 콘텐츠 데이터로 변환해서 사용된다.

예) <img alt="x"> => [IMG]-x

- disable : 해당 HTML 태그는 WML에서 지원되지 않는 것으로 간주하여 삭제하나 시작 태그와 종료 태그 사이에 내포해 있는 태그와 콘텐츠 데이터들은 삭제하지 않고, 보존한다.

- discard : 해당 HTML 태그 뿐 아니라 시작 태그와 종료 태그 사이에 내포해 있는 모든 태그와 콘텐츠 데이터들을 삭제 한다.

- invalid : 위의 5가지 클래스 어디에도 속하지 않는 경우로 이런 태그 발생 시, 예외 처리

#### 4. WML 파일 생성 모듈

WML 파일 생성 모듈에서는 태그 변환 규칙을 가지고 생성된 WML 콘텐츠를 실질적으로 무선 단말기 사용자에게 보낼 WML 파일을 생성하는 기능을 수행한다.

기존 파일 이름에 확장자 .wml을 붙이고 그림 4와 같은 wml 파일 헤더와 DTD 정보를 문서 첫 부분에 삽입하게 된다. 그리고 마스터 HTML 콘텐츠를 가지고 WML 콘텐츠로 변환했을 때 2Kbytes 초과 시에는 양

```
<?xml version="1.0" encoding="ks_c_5601-1987"?>
<!DOCTYPE wml PUBLIC "-//WAPFORUM//DTD WML 1.1//EN"
"http://www.wapforum.org/DTD/wml_1.1.xml">
```

그림 4. WML 헤더  
Fig. 4. WML Header.

적 분할을 해서 다음 텍스트의 링크를 생성해 준다.

#### 5. 프레임 콘텐츠 재구성 모듈

프레임 콘텐츠 재구성 모듈은 최초 변환기가 받은 HTML 콘텐츠가 프레임 셋 콘텐츠를 읽어 들였을 경우, 올바른 변환을 수행해 줄 수 있도록 콘텐츠를 교정 처리한다. 사실상 무선 단말기의 작은 디스플레이 화면과 지원되는 브라우저가 없는 제약 사항 등을 고려할 때, HTML에서와 같은 상호 운용성을 제공할 수 있는 WML 콘텐츠로 변환한다는 것은 어려움이 많이 있다. 기존 HTML 변환기에서 프레임 변환이 어려웠던 것은 무선 단말기 사용자가 프레임 콘텐츠를 서버로 요구할 경우, 변환기가 받는 콘텐츠는 전체 프레임 요소를 정의해 놓은 프레임 셋 소스만을 받아서 변환을 하기 때문에 올바른 변환을 수행하지 못했다. 따라서 최초에 받은 프레임 셋 소스를 파싱하여, 포함되어 있는 모든 파일 이름들을 추출해서 추출한 파일 이름을 가지고 요구 메시지를 다시 구성하여, 파일 개수만큼 매번 서버에 연결한다. 그리고 각각의 문서들을 받아서 처음에 받은 HTML 문서에 연결시킨 다음 html, body, /body, /html 등의 중복 태그들을 삭제해서 변환이 가능하도록 처리한다. 여기서 프레임 콘텐츠를 변환 시키는데 있어서 고려되어야 할 점이 있다. 대부분의 HTML 홈페이지가 여러 메뉴에 대해서 보여 지게 되는 타겟 프레임은 하나로 고정되어 있어 우측에는 메뉴에 링크 되어있는 HTML 콘텐츠를 보여주는 등의 사용자 인터페이스를 제공하나 상용 서비스에서는 그렇지 못한 실정이다. 따라서 그림 5와 같이 무선 단말기의 디스플레이 화면에 1행 2열의 테이블로 변환이 되어서 좌측 열에는 여러 메뉴 콘텐츠를 우측 열에는 타겟 프레임에 보여 지게 되는 콘텐츠를 위치시킴으로 해서 기존 서비스와 차별화하여 프레임 콘텐츠의 변환을 가능하도록 했다.

변환 방법을 좀 더 상세히 기술하면 다음과 같다. 단말기 화면의 좌측 열에는 HTML 콘텐츠 중에서 태그 <a>와 "target" 속성을 가지고 있는 데이터를 추출해서

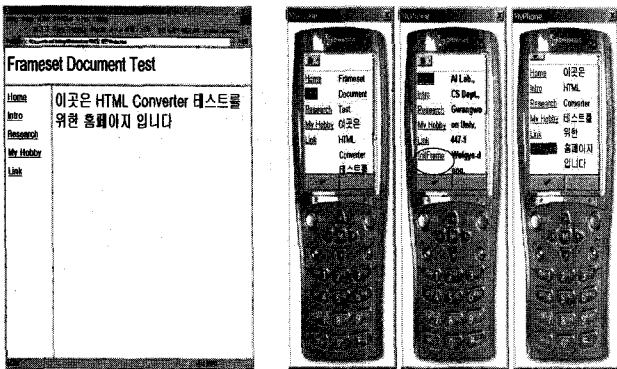


그림 5. 프레임 콘텐츠 변환  
Fig. 5. Conversion of Frame Contents.

위치시키고 우측 열에는 그 나머지 모든 데이터들을 위치시키다. intro 메뉴로 이동 할 때에는 intro.html이 변환이 되는데 역시 1행 2열의 테이블로 변환이 되어서 우측 열에 위치 하게 되고 좌측열에는 역시 태그 <a>와 "target" 속성을 가지고 있는 데이터를 위치시키고 InitFrame.wml 파일 링크를 추가 시켜 줌으로 해서 최초 변환된 WML 파일을 액세스 할 수 있도록 처리 해 준다. 이와 같이 해서 기존 HTML 프레임에서 클라이언트와의 상호 운용성이 반영된 변환을 수행 하게 된다.

6. 이미지 맵 변환

이미지 맵이란 일반적으로 <a>태그를 사용할 때 단순히 하나의 문서나 하나의 사이트로밖에 연결할 수 없는 단점을 극복할 수 있는 방법으로서, 한 개의 이미지를 정해 놓고 그 이미지를 영역별로 분할하여 저마다 다른 링크를 설정한 후, 마우스 포인터로 해당 영역을 선택하였을 때 정해진 링크를 따라 해당 문서나, 사이트를 열도록 하는 태그이다. 이미지 맵을 사용하기 위해서는 그림 6에서 보는 바와 같이 <img>, <map>, <area> 태그 등이 필요하다. <img> 태그는 클릭 가능한 맵에서 사용할 이미지 파일의 이름 또는 URL을 지정하고, <map> 태그는 사용할 맵 이름을, <area> 태그는 맵의 형태, 좌표 값, 그리고 클릭 했을 경우의 이동할 URL을 지정한다. 많은 웹사이트에서, 메뉴를 꾸미기 위해 위와 같은 이미지 맵을 이용하기 때문에, 이에 대한 변환이 필요하다. 변환 방법은 이미지 맵 관련 태그를 파싱한 후, 링크되어 있는 HTML 문서의 이름을 추출해서 WML 데이터로 대체한 후, 해당 문서로 링크 시킨다.

```

<map name="ImageMap1">
<area shape="rect" coords="37,45,124,166" href="링크 될 주소">
```

그림 6. 이미지 맵 태그  
Fig. 6. Image Map Tag.

IV. 실험 결과 및 평가

구현한 HTML 변환기는 리눅스 7.1 환경에서 KDevelop 1.4.1을 사용하여 개발하였고, 시뮬레이터는 오픈웨이브에서 제공하는 Openwave SDK 5.1의 Openwave Mobile Browser 5.0.2를 Windows 2000 Professional Edition에서 사용하였다.

테스트를 위해 사용한 콘텐츠는 리눅스 아파치 웹 서버에서 동작하는 HTML 콘텐츠를 사용하였다. 실험 결과는 그림 7, 그림 8과 같고 본 논문에서 구현한 HTML 변환기의 경우 프레임 콘텐츠와 이미지 맵 콘텐츠가 보다 효과적으로 변환 된 것을 알 수 있다. 프레임 콘텐츠의 경우, 그림 7에서 보듯이 1행 2열의 테이블로 변환이 되며 메뉴를 클릭 했을 때에도 역시 테이블로 변환이 되어 좌측 열에 메뉴를 우측 열에 링크되어 있는 문서를 위치 시켜서 무선 단말기 사용자의 화면에 출력한다. 이때 초기 화면을 액세스하기 위해 테이블 변환 시, 그림 7에서 보는 바와 같이 좌측 열 메뉴 마지막에 IntiFrame(IntiFarne.wml) 메뉴 링크를 추가 해 준다. 이와 같은 방법으로 변환함으로써 HTML 에서와 유사한 상호 운용성을 제공할 수 있게 된다.

다음으로 그림 9와 같은 3개, 4개, 5개의 프레임으로

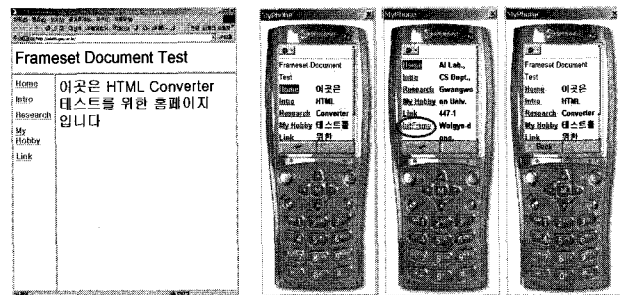


그림 7. 프레임 콘텐츠 변환  
Fig. 7. Conversion of Frame Contents.

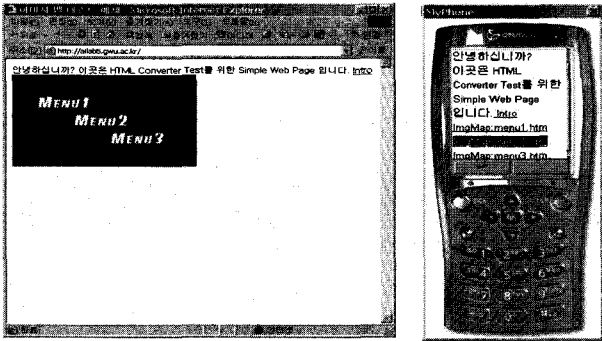


그림 8. 이미지 맵 콘텐츠 변환  
Fig. 8. Conversion the Image map Contents.

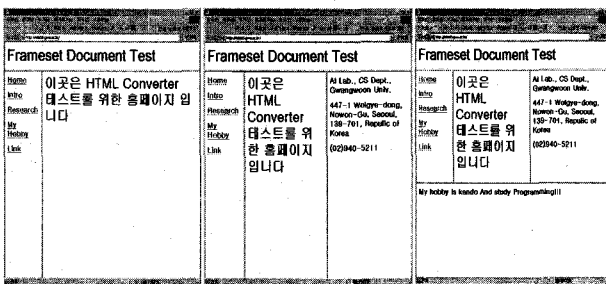


그림 9. 성능 분석 콘텐츠 I  
Fig. 9. Performance Analysis Contents I.

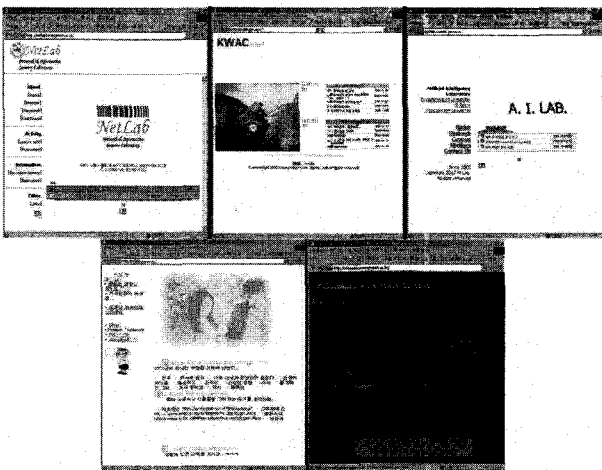


그림 10. 성능 분석 콘텐츠 II  
Fig. 10. Performance Analysis Contents II.

구성된 3개의 예제 콘텐츠와 그림 10과 같은 실제 웹상에 존재하는 프레임 콘텐츠 중에서 2개의 프레임으로 구성된 콘텐츠, 한 프레임에서 다른 프레임 셋이 중첩된 중첩 프레임 콘텐츠 등 5개의 임의의 콘텐츠를 변환시켰을 때 역시 위의 그림 7과 같은 1행 2열의 테이블 형태로 원활하게 변환이 수행된다. 성능 평가로써 표 3과 같이 기존의 단순 변환과, 본 논문에서 제안한 테이블 콘텐츠 변환시의 오버헤드를 측정비교 했다. 변환

표 3. 성능 분석  
Table 3. Performance Analysis.

	단순 변환		테이블 변환		오버 헤드	
	변환양	변환시간	변환양	변환시간	데이터 증가	시간 지연
3개프레임	480bytes	0.15초	534bytes	0.19초	11.2%	26.6%
4개프레임	606bytes	0.16초	660bytes	0.22초	8.9%	37.5%
5개프레임	654bytes	0.20초	708bytes	0.25초	8.2%	25.0%
실제 콘텐츠 (중첩프레임)	549bytes	0.16초	603bytes	0.21초	9.8%	31.2%
실제 콘텐츠 (2개프레임)	559bytes	0.18초	609bytes	0.23초	8.9%	27.7%
실제 콘텐츠 (2개프레임)	942bytes	0.23초	978bytes	0.28초	3.8%	21.7%
실제 콘텐츠 (2개프레임)	1656bytes	0.22초	1686bytes	0.27초	1.8%	22.7%
실제 콘텐츠 (2개프레임)	1082bytes	0.25초	1101bytes	0.29초	1.7%	16.0%

시간과 변환된 WML 파일의 데이터 양을 측정해서, 각각의 데이터 증가와 시간 지연의 오버헤드를 계산했을 때, 변환양의 경우는 HTML 콘텐츠 데이터의 양이 클수록 기존 단순 변환 때에 비해서 테이블 변환 시 오버헤드가 작게 나타나는 것을 알 수 있었으며, 변환 시간의 경우는 테이블 변환 시, 작게는 약 16% 많게는 약 37%의 시간 지연 오버헤드가 발생했다.

### V. 결 론

본 논문에서는 WAP이라고 하는 무선 인터넷 환경에서 무선 단말기 사용자가 WAP 게이트웨이를 통해 HTML 콘텐츠를 요구했을 때 WML 콘텐츠로 변환하는 HTML 변환기를 설계 및 구현하였다. 구현된 HTML 변환기의 구조는 콘텐츠 변환, 변환 규칙 집합, WML 파일 생성, 프레임 콘텐츠 재구성 모듈로 구성된다.

콘텐츠 변환 모듈은 HTML 콘텐츠를 SP(SGML Parser toolkit)를 이용, 파싱해서 토큰화하여 트리 구조로 저장하며, SP에서 제공하는 API를 통해 WMLConversionRulesSet을 참조하여 WML 콘텐츠로 변환시킨다. 변환 규칙 집합에서는 변환 규칙을 적용하기 위해서 HTML 태그를 valid, validAttributes, validAttributesAsData, disable, discard, invalid 등의 6개 클래스로 분류했다.

프레임 콘텐츠 재구성 모듈은 단말기 사용자가 요구



하는 문서가 프레임으로 구성된 문서인 경우에 최초로 받은 프레임 셋 소스를 파싱하여, 포함되어 있는 모든 파일 이름들을 가지고 서버에 연결해서 문서들을 받아 받은 HTML 문서에 Append 시키고, 일련의 재구성 과정을 거친 후, 1행 2열의 테이블 형태의 WML 콘텐츠로 변환이 된다. 좌측열에는 <a href="x" target="y">와 같이 target 속성을 가지고 있는 콘텐츠를 위치시키고 우측열에는 그 나머지 모든 데이터들을 위치시켜서 무선 단말기 사용자의 화면에 출력함으로 해서 상용 서비스와 차별화되고, 기존 HTML에서와 같은 사용자 인터페이스와 상호 운용성을 제공 할 수 있도록 변환했다.

상용 서비스에서 제공하지 않는 이미지 맵 콘텐츠 변환은 이미지 맵 관련 태그를 파싱한 후, 링크되어 있는 html 문서의 이름을 추출해서 WML 콘텐츠 데이터로 대체하고, 해당 콘텐츠로 링크시킴으로써 변환이 가능하게 한다.

위와 같이 본 논문에서는 HTML 변환기를 설계하고 구현하였으며, 오픈웨이브에서 제공하는 Openwave SDK 5.1의 Openwave Mobile Browser 5.0.2 시뮬레이터를 사용하여 프레임 콘텐츠와 이미지 맵 콘텐츠가 효과적으로 변환됨을 확인하고 평가하였다.

향후 연구 과제로는 HTML 콘텐츠의 스크립트의 처리, 웹에서의 CGI 기능처리 등의 연구를 필요로 한다.

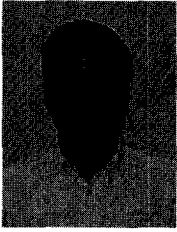
Mobile Devices”, IEEE Computer , June 2001, pp54-58.

- [7] 박영충, 신동일, 신동규, “모바일 인터넷을 이용한 HTML-WML Converting System의 설계 및 구현”, 한국정보과학회 가을 학술발표논문집Vol.27, No.2, pp.301-303, 2000.
- [8] 이승진, 김대건, 최린, 강철희, “확장성 있는 웹서비스를 위한 무선 응용 프로토콜 기반의 HTML Filter구현”, 한국정보과학회 봄 학술발표 논문집, Vol.28, No.1, pp.391-393, 2001.
- [9] 우찬일, 신인철, “실시간 정보 제공을 위한 웹 트랜스 코딩”, 대한전자공학회논문지CI, 1229-6376, 제39권2호, pp.54-63, 2002.
- [10] 이미경, 장민수, 김재홍, 손주찬, “웹 콘텐츠 변환을 위한 태그 변환기의 설계 및 구현”, 한국정보처리학회 2003년 춘계학술대회, Vol.10, No.01, pp.1297-1300, 2003.
- [11] 조승호, 차정훈, “Clipping 기반의 무선 인터넷 사이트 구축용 변환 서버 구현”, 한국정보처리학회 논문지A, Vol.11, No.02, pp.0165-0174, 2004.

## 참 고 문 헌

- [1] WAP Forum, <http://www.WAPforum.org>
- [2] James Clark, <http://jclark.com/sp/>
- [3] Marcin Metter, Dr Robert Colomb, “Wap enabling existing HTML applications,” User Interface Conference 2000, Page(s): 49 ?57, 2000.
- [4] Mario Canataro, Domenico Pascuzzi, “An Object-Based Architecture for WAP-Compliant Applications” Database and Expert systems Applications, 2000 Proceedings 11th International Workshop on 2000, p178~185.
- [5] Chi-Hsing Chu, Chien-Hsuan Huang, Michel Lee, “Building a XML-Based Unified User Interface System under J2EE Architecture”, Multimedia Software Engineering, 2000 Proceedings, International Symposium on, 2000.
- [6] Subhasis Saha Mark Jarntgaard, John Villasensor, “Bringing the Wireless Internet to

저 자 소개



한진섭(학생회원)  
 2001년 광운대학교 전자계산학과  
 학사 졸업(이학사).  
 2003년 광운대학교 컴퓨터과학과  
 석사 졸업(공학석사).  
 2003년~2005년 (주)한국컴퓨터  
 기술연구소 연구원.

2005년~현재 광운대학교 컴퓨터과학과  
 박사 과정.

<주관심분야: Artificial Intelligence, Data  
 Mining, Ubiquitous System>



박병준(정회원)  
 1984년 서울대학교 컴퓨터공학과  
 학사 졸업(공학사).  
 1988년 University of Minnesota,  
 Computer Science  
 석사 졸업(공학석사).

1989년~1996년 US Army  
 Construction Engineering  
 Research Lab. Researcher.

1997년 University of Illinois at  
 Urbana-Champaign, Computer Science  
 박사 졸업(공학박사).

1997년~1998년 Epic Systems Corp, Senior  
 Researcher.

1998년~2000년 SPSS Inc, Senior Researcher

2000년~현재 광운대학교 컴퓨터소프트웨어학과  
 교수.

<주관심분야 : Artificial Intelligence, Data/web  
 Mining, Knowledge-Based Systems>