

CC/PP와 애노테이션에 기반한 웹 문서 트랜스코딩

김희모⁰, 이경호

연세대학교 컴퓨터과학과

hmkim@icl.yonsei.ac.kr, khlee@cs.yonsei.ac.kr

Web Document Transcoding based on CC/PP and Annotation

Hwe-Mo Kim⁰, Kyong-Ho Lee

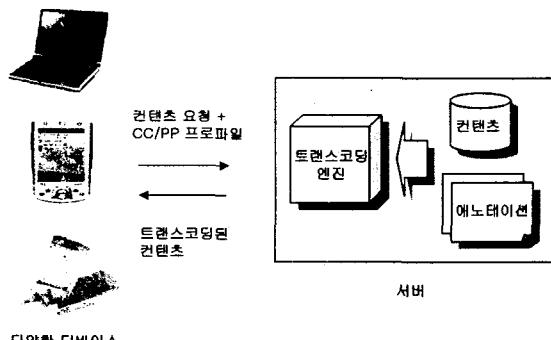
Dept. of Computer Science, Yonsei University

요약

모바일 디바이스가 널리 사용됨에 따라 이를 통한 웹 컨텐츠의 이용이 증가하고 있다. 그러나 모바일 디바이스를 통하여 기존의 웹 컨텐츠를 이용하는 데에는 한계가 있다. 본 논문에서는 CC/PP 프로파일에 따라 웹 문서를 적절히 가공하여 전송하는 트랜스코딩 방법을 제안한다. 제안된 방법은 보다 정교한 수준의 맞춤형 서비스를 지원하기 위하여 원본 문서에 애노테이션(annotation)을 기술할 수 있는 방법을 지원한다. 제안된 애노테이션은 모바일 디바이스에서 표시할 수 없는 컨텐츠를 임의의 리소스로 대체할 수 있다. 또한 제안된 방법은 디바이스의 스크린 사이즈를 고려하여 컨텐츠를 적절한 크기로 나누어 보여주며, 문서의 구조를 효과적으로 전달하기 위한 내비게이션 맵을 제공한다.

1. 서론

무선 네트워크 및 모바일 컴퓨팅 기술의 발전에 힘입어 언제 어디서든 원하는 정보를 얻을 수 있는 환경이 구축되고 있다. 이에 따라 핸드폰, PDA 등 다양한 종류의 디바이스를 통한 인터넷 접속이 가능해졌다.[1][2]



다양한 디바이스

그림 1. CC/PP와 애노테이션에 기반한 트랜스코딩

모바일 디바이스의 컴퓨팅 파워가 증가하고 종류도 다양해졌지만 모바일 디바이스는 기기의 특성상 표현 가능한 컨텐츠의 종류 및 스크린 사이즈 등 다양한 측면에서 제약을 가질 수밖에 없다. 기존의 웹 컨텐츠는 모바일 디바이스에서 제대로 표현되며 흥미로운 만 아니라 표현된다고 해도 사용자가 컨텐츠를 사용하기 위해서는 과다한 화면 스크롤이 필요한 경우가 많다.

이러한 문제를 해결하기 위해서는 모바일 디바이스의 특성을 고려한 트랜스코딩 기법이 필요하다. 그러나 기존의 웹 컨텐츠들은 주로 PC를 대상으로 하여 작성되었기 때문에 컨텐츠 자체에 트랜스코딩을 위한 정보를 포함하고 있지 않다. 따라서, 본 논문에서는 디바이스의 특성을 기술한 CC/PP(Composite Capability/Preference Profile) [3][4] 프로파일과 원본 컨텐츠에 대한 부가적인 정보를 기술할 수 있는 애노테이션(annotation)에 기반한 정교한 수준의 트랜스코딩 방법을 제안한다.

제안된 트랜스코딩 방법은 세 가지의 주요 기능을 지원한다. 먼저 스크린 사이즈가 작은 모바일 디바이스를 위해 원본 컨텐츠를 그룹으로 나누어 보여 주는 기능을 지원하는데 이때 각각의 그룹에 대한 중요도를 계산하여 중요한 그룹은 우선적으로 보여주고 중요도가 낮은 그룹은 생략한다. 또한, 사용자의 디바

이스에서 표현 불가능한 컨텐츠는 미리 작성한 애노테이션 파일을 바탕으로 대체 가능한 리소스로 대체한다. 또한, 페이지를 나눔으로 페이지 수 증가로 인해 사용자가 원하는 컨텐츠를 탐색하는데 오랜 시간이 걸리는 것을 막기 위해 내비게이션 맵을 제공하여 보다 효율적으로 웹 컨텐츠를 접근할 수 있도록 하였다. 한편 <표 1>은 CC/PP나 애노테이션에 기반한 기존 연구를 요약한 것이다.

표 1. CC/PP 또는 애노테이션에 기반한 트랜스코딩 기법

저자	연도	방법	특징	애노
Hori 등 [5]	2000	X	O	리소스 대체 및 페이지 분할, 저작 단계에서의 변환 확인
Lemlouma 와 Layaïda [6]	2001	O	O	구조적 트랜스 코딩 및 미디어 트랜스 코딩
Intel [7]	2001	O	X	컨텐츠 선택, 컨텐츠 변환, 동적 컨텐츠 생성
Yasuda [8]	2001	O	X	이미지 포맷 변환, 이미지 컬러 수 감소, 이미지 크기 감소, HTML의 XHTML-basic으로의 트랜스코딩, 일본어 문자 코드 트랜스코딩, 컨텐츠 요약/추출 등
Korolev 와 A. Joshi [9]	2001	O	X	CC/PP 프로파일에 따라 미리 저작 해둔 컨텐츠 중 하나를 선택해서 전송하거나 미리 만들어 둔 XSLT 를 적용하여 새로운 컨텐츠를 전송함
Ma 등 [10]	2002	O	O	레이아웃 변환, 미디어 변환
Kinno 등 [11]	2003	O	X	XACML로 표시된 접근 규칙에 따라 XSLT 생성

2. 애노테이션

어떤 애노테이션 속성을 제공하는가 하는 문제는 결국 트랜스코딩 시스템에서 제공하는 변환의 종류를 결정한다. 본 논문의 기반이 되는 애노테이션의 구조는 <그림 2>와 같다. 제안된 애노테이션 스키마는 크게 대체 리소스를 표기하는 엘리먼트 SUBSTITUTE와 논리적으로 단일의 그룹을 나타내는 엘리먼트 GROUP으로 이루어진다. SUBSTITUTE는 특정 리소스를 대체 할 수 있는 리소스가 존재할 경우, 해당 리소스가 요구하는 클라이언트측의 능력과 위치를 표시한다.

GROUP은 논리적으로 독립된 단위를 이루는 부분을 표현함으로서 페이지 분할을 지원한다. SUBSTITUTE 엘리먼트는 엘

리먼트 SRC와 ALTERNATIVE를 자식으로 갖는다. SRC는 원본 리소스의 위치를 가리키며 ALTERNATIVE는 대체 가능한 리소스를 가리킨다. ALTERNATIVE는 자식 엘리먼트로 REQUIRES와 PATH를 가지는데 REQUIRES는 특정 그룹에 해당 리소스가 적용된 이후 변하는 그 그룹의 제약 사항을 표시하고 PATH는 대체 리소스의 위치를 표시한다. 또한, ALTERNATIVE는 속성으로 GROUPNUM, NUM, FIDELITY를 갖는데 이들은 각각 대체 가능한 리소스가 속한 그룹의 번호와 대체리소스 번호 그리고 원본컨텐츠에 대한 충실도를 나타낸다.

엘리먼트 REQUIRES는 엘리먼트 ATTRIBUTES 그리고 엘리먼트 LESSTHAN, MORETHAN 또는 EQUAL을 자식 엘리먼트로 갖는다. ATTRIBUTE는 참조하고자 하는 CC/PP 프로파일의 속성의 이름을 표기하고 속성에 따라 LESSTHAN, GREATER-THAN, EQUAL 엘리먼트로 ATTRIBUTE 값이 가져야 할 조건을 표시한다. 대체 리소스의 위치가 외부 파일에 존재할 때는 엘리먼트 PATH의 SRC 속성을 이용해서 대체 리소스의 위치를 표시하고 PATH의 값에는 해당 리소스의 XPath 수식을 기술한다. 대체 리소스의 내용을 직접 표시하려면 엘리먼트 CONTENT를 사용한다.

엘리먼트 GROUP은 엘리먼트 PATH를 자식으로 갖는다. PATH는 해당 그룹에 포함되는 컨텐츠들의 XPath 수식을 값으로 갖는다. 또한 각각의 그룹에 대한 번호, 중요도, 요약, 그리고 역할을 나타내는 속성 NUM, IMPORTANCE, SUMMARY 그리고 ROLE을 포함한다. 트랜스코딩 과정에서 시스템은 사용자가 지정한 중요도와 역할을 바탕으로 해당 그룹의 중요도를 계산한다. 사용자 지정 중요도는 -1부터 1사이의 값을 나타내며 1에 가까울수록 중요함을 의미한다. 사용자 지정 중요도는 같은 역할을 가지는 그룹들이라도 사용자가 임의로 중요한 것과 중요하지 않은 것을 구분할 수 있는 방법을 제공한다. SUMMARY는 내비게이션 맵을 작성할 때 각각의 그룹이 해당 그룹의 요약으로 대체되도록 사용된다. ROLE은 각각의 역할에 따라 특별한 처리가 필요할 경우를 위해 그룹의 역할을 표시한다. 또한, PATH는 해당 리소스를 표현하는데 필요한 제약 사항을 표현하는 엘리먼트 REQUIRES를 자식으로 갖는다. REQUIRES는 생략 가능하며 생략된 경우, 텍스트만으로 이루어진 컨텐츠를 의미한다.

3. 제안된 방법론

본 절에서는 앞에서 소개한 애노테이션과 CC/PP 프로파일을 기반으로 사용자의 디바이스에 맞게 XML 문서를 트랜스코딩하는 방법을 자세히 기술한다. 제안된 트랜스코딩 방법은 <그림

3>과 같이 내비게이션 맵 생성, 변환 규칙 결정, 그리고 XSLT 스크립트 생성의 세 단계로 구성된다. 제안된 방법은 크게 3가지의 기능을 지원한다. 먼저 사용자의 위치에 따라 동적으로 내비게이션 맵을 생성하여 사용자의 디바이스에 맞게 보여주는 기능, 애노테이션 파일에 표시된 조건에 맞는 리소스를 중 사용자 디바이스에 가장 적절한 리소스로의 대체 기능, 그리고 사용자 디바이스에 맞게 페이지를 나누어 보여주는 기능을 포함한다.

3.1 내비게이션 맵 생성

스크린의 사이즈가 작은 모바일 디바이스의 경우, 한번에 사용자에게 보여줄 수 있는 컨텐츠의 크기가 제한되어 사용자가 원하는 정보를 찾기까지의 탐색 시간이 비교적 오래 걸릴 수밖에 없다. 따라서, 본 논문에서는 사용자가 원하는 컨텐츠를 보다 쉽게 찾을 수 있도록 내비게이션 맵을 제공하여 사용자가 보다 효율적으로 원하는 페이지를 찾아갈 수 있도록 한다. 내비게이션 맵은 일종의 목차 기능을 하는 것으로 해당페이지를 이루는 그룹의 엘리먼트 SUMMARY를 이름으로 갖고 해당 그룹의 링크를 포함하는 하이퍼링크로 이루어진다.

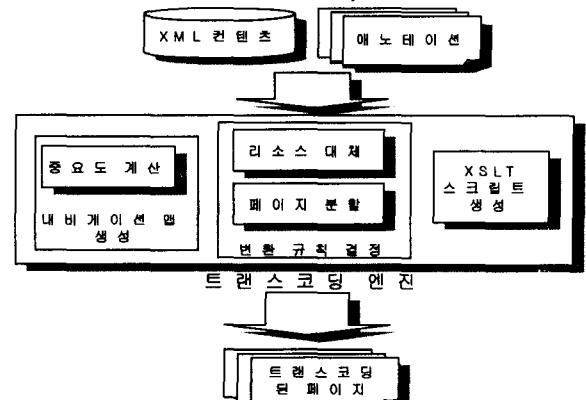


그림 3. 제안된 트랜스코딩 엔진의 구조

원본 컨텐츠의 내용을 단순히 스크린 사이즈를 고려하여 적당히 몇 개의 페이지로 나누어 준다면 사용자가 원하는 정보를 찾는데 어려움을 겪을 수밖에 없다. 따라서, 보다 중요한 내용을 우선적으로 표시하고 중요하지 않은 정보는 스크린 사이즈를 고려하여 생략 한다.

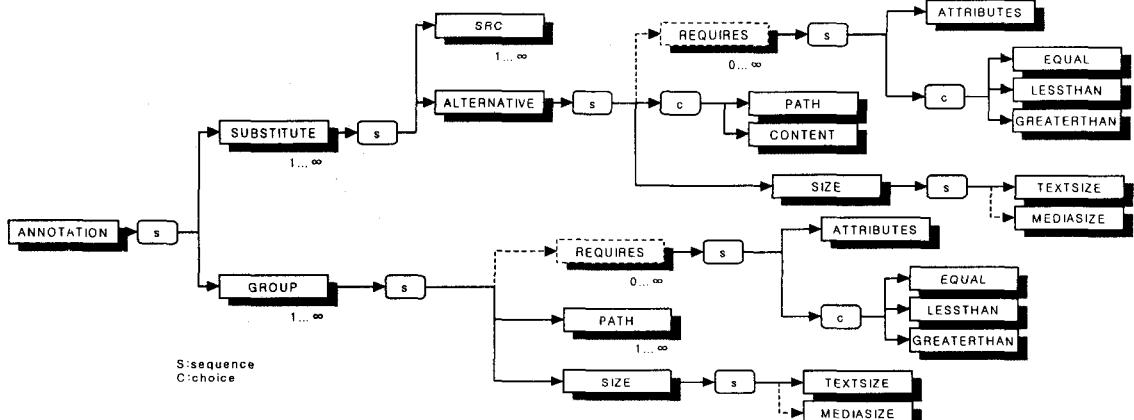


그림 2. 제안된 애노테이션 스키마의 엘리먼트 계층 구조

사용자가 특정 페이지를 요청하면 시스템은 해당 페이지의 애노테이션 파일로부터 개별 그룹들에 대한 정보를 받아와서 중요도를 계산한다. 중요도는 해당 그룹의 역할에 따른 중요도 값과 사용자 지정 중요도 값의 합으로 계산된다. 계산된 중요도를 바탕으로 중요도가 높은 그룹이 상위에 표시되도록 하며 중요도가 특정 수치 이하의 낮은 그룹의 경우는 생략된다. 중요도 값에 따라 표시되기도 결정된 그룹은 각각의 요약을 이용으로 가지는 링크페이지로 대체된다.

3.2 변환 규칙 결정

사용자가 내비게이션 맵을 통해 특정 그룹의 페이지를 요청하면 시스템은 사용자의 CC/PP 프로파일, 애노테이션 파일, 그리고 원본 컨텐츠를 살펴보고 변환 규칙을 결정한다. 사용자가 요청한 페이지에 해당하는 애노테이션 파일에서 대체 가능한 리소스를 찾고, 이중 사용자의 디바이스에서 제공할 수 있는 가장 높은 수준의 서비스를 제공할 수 있도록 대체 리소스를 선택한다. 애노테이션 파일의 내용에는 리소스가 포함되기 위한 조건이 표시되어 있으므로 조건을 만족시키는 가장 우선순위가 높은 리소스를 선택한다.

내비게이션 맵을 통하여 특정한 그룹에 대한 요청이 오면 먼저 해당 그룹을 표시하기 위한 제약사항을 만족시키는지를 검사한다. 제약 사항을 만족시키지 못하는 경우 해당 그룹의 대체 리소스들을 순서대로 살펴가며 해당 리소스의 제약사항을 만족시킬 수 있는 대체 리소스를 찾아 대체한다. 이때 원본 컨텐츠에 대한 충실성을 나타내는 속성 FIDELITY값이 높은 것을 우선적으로 선택한다. 대체 여부를 결정한 이후에는 해당 그룹을 표시하기 위한 스크린 사이즈를 계산하여 스크린 사이즈에 여유가 있을 경우 기존 내비게이션 맵의 나머지 엔트리들을 대상으로 특정 값 이상의 중요도를 가지는 것들을 내림차순으로 스크린 사이즈의 한도 내에서 표시한다. 해당하는 모든 엔트리를 텍스트로 표시하고도 스크린 사이즈의 여유가 있으면 내비게이션 맵의 나머지 엔트리를 중 보다 원본 컨텐츠에 가깝게 보여줄 수 있는 대체리소스를 찾아 대체 한다.

트랜스코딩에서는 보다 중요하고 원본에 가까운 리소스를 보여주는 것이 바람직하므로 이를 위해 내비게이션 엔트리들을 대상으로 표시 우선순위를 계산하였다. 각 그룹의 역할과 사용자가 애노테이션 파일에 지정한 중요도로 계산된 중요도 값에 해당 대체 리소스의 충실도를 곱한 값이 큰 것을 우선적으로 표시하기 위함이다. 재한된 스크린 사이즈에서 전체 컨텐츠의 (중요도x충실도)를 최대화하기 위해서 (중요도x충실도)/스크린 사이즈의 값이 큰 것부터 선택하여 (중요도x충실도)의 값을 최대로 하는 경우를 계산한다.

요청된 그룹을 표시하기 위한 제약사항을 만족시키는지를 검사한 후 제약 사항을 만족시키지 못할 경우 제약사항을 만족시키면서 원본 컨텐츠에 대한 충실도가 가장 높은 대체 리소스를 선택한다. 전체 스크린 사이즈에서 요청된 그룹을 표시하기 위한 스크린 사이즈를 재하고 남은 그룹들을 대상으로 계산된 표시 우선순위가 높은 순서대로 스크린 사이즈의 한도 내에서 보다 충실도가 높은 대체 리소스를 선택한다.

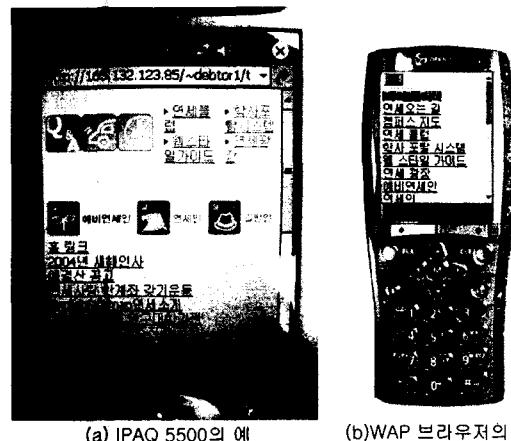
3.3 XSLT 스크립트 생성

요청된 페이지에 대한 변환 규칙이 결정되면 리소스의 대체와 페이지 분할을 위한 XSLT 스크립트를 생성하여 원본 컨텐츠에 적용하여 변환된 컨텐츠를 생성한다. 앞 단계에서 결정된 페이지 분할 정책과 리소스 대체 정책을 토대로 XSLT 스크립트를 생성한다. XSLT 스크립트를 통해 하고자 하는 내용이 이미 정해져 있으므로 XSLT 스크립트의 내용이 변할 뿐이지 전제적인 구조는 변하지 않는다.

4. 성능평가 및 결론

본 논문에서 제안한 방식은 기존의 연구들과 비교하여 다음과 같은 특징을 가진다. 첫째로, 원문에 대한 애노테이션 정보를 고려하지 않는 기존 연구에서는 다른 기획된 복잡한 구조의 문서를 지원한다. 애노테이션에 기반한 방법은 원본 컨텐츠에 대한 자동 분석을 통하여 추출할 수 없는 정보들까지 애노테이션

저작자가 표시해 줄 수 있으므로 보다 정교한 수준의 트랜스코딩이 가능하다. 둘째, 스크린 사이즈가 작은 장치의 경우 한번에 보여줄 수 있는 컨텐츠의 양이 적은 만큼 중요한 컨텐츠는 최대한 원본 컨텐츠에 가깝도록 트랜스코딩하고 중요도가 떨어지는 컨텐츠는 제외하도록 했다. 이를 통하여 사용자의 요구에 보다 적합한 컨텐츠를 가능한한 원본과 비슷하게 트랜스코딩 할 수 있다. 마지막으로, 사용자가 작성한 애노테이션에 사용자 지정 중요도와 대체 리소스의 원본에 대한 충실도를 추가함으로서 컨텐츠 저작자의 의도가 보다 잘 반영될 수 있도록 하였다. <그림 4>는 제안된 트랜스코딩 방법을 이용하여 실제 웹 컨텐츠를 다양한 모바일 디바이스를 통하여 브라우징한 결과이다.



(a) IPAQ 5500의 예 (b)WAP 브라우저의 예
그림 4. 제안된 트랜스코딩 방법의 적용 결과

5. 참고 문헌

- [1] L. Suryanarayana and J. Hjelm "Profiles for the Situated Web," Proc. WWW Conference May 2002.
- [2] M. Butler, F. Giannetti, R. Gimson and T. Wiley "Device Independence and the Web," IEEE Internet Computing, Vol.6, Issue 5, pp.81-86, September/October, 2002.
- [3] CompositeCapability/PreferenceProfiles(CC/PP), W3C Recommendation, <http://www.w3.org/Mobile/CCPP/>, Jan. 2004.
- [4] J. Indulska, R. Robinson, A. Rakotonirainy and K. Henricksen "Experiences in Using CC/PP in Context Aware Systems," Proc. MDM, Melbourne, 2003.
- [5] M. Hori, G. Kondoh, K. Ono, S. Hirose and S. Singhal Annotation-based Web Content Transcoding," Proc. 9th World Wide Web Conference (WWW-9), Amsterdam, Netherlands, May 2000.
- [6] T. Lemlouma and N.I Layaida, "NAC: A Basic Core for the Adaptation and Negotiation of Multimedia Services," OPERA Project, INRIA, Sep. 2001.
- [7] Intel CC/PP Toolkit, <http://www.intel.com/pca/developernewtwork/>
- [8] K. YASUDA "Implementation and Evaluation of Keio CC/PP Implementation" <http://yax.tom.sfc.keio.ac.jp/panda/>
- [9] V. Korolev and A. Joshi. "An End-End Approach to Wireless Web Access," Proc. Int'l Workshop on Wireless Networks and Mobile Computing, Apr. 2001.
- [10] W. Ma, I. Bedner, G. Chang, A. Kuchinsky, and H. J.Zhang. "A Framework for Adaptive Content Delivery in Heterogeneous Network Environments," Proc. SPIE MMN Vol.3969 pp.86-100, San Jose, USA, 2000.
- [11] A. Kinno, Y. Yonemoto, T. Nakayama and M. Etoh "Environmentally-Adaptive XML Transformation and Its Application to Content Delivery," Proc. Int'l Conf. On Communications(ICC), Vol. 2, pp.844 - 848, May. 2003.