

모바일 웹 서비스를 위한 PC 웹에서의 중요 이미지 결정 방법

박대혁*, 홍마리아**, 서정만***, 이근수****, 임영환*****

The method of Deciding PC Web Important Image for Mobile Web Service

Dae-Hyuck Park *, Maria Hong **, Jeong-Man Seo ***, Keun-Soo Lee ****, Young-Hwan Lim *****

요 약

인제 어디서나 네트워크에 연결되어 있는 유비쿼터스 환경에서의 무선 단말기의 콘텐츠에 대한 연구가 활성화되고 있다. 본 논문에서는 일반 컴퓨터의 인터넷 콘텐츠를 무선 단말기를 이용하여 얻을 수 있도록 하는 프레임워크를 제안하고자 한다. 프레임워크의 콘텐츠를 관리하는 서버의 성능 향상을 위해서 다수의 리소스 중에서 위한 중요 이미지를 선정하고, 선정된 이미지만을 단말기에서 재생 가능한 이미지로 변환하는 것이 본 논문의 목표이다. 이를 위해 웹 문서의 이미지 관련 정보 중 화면 표시율(DPR)과 이미지의 분포 영역을 고려하여 다수의 변환 요구 이미지 중에서 중요 이미지를 선별한 후 모바일 이미지로 변환 서비스하여 서버의 모바일 이미지로의 변환 시간 및 자원을 축소 시켜 서버의 성능을 향상시키고자 한다.

Abstract

There are many new studies on the contents for mobile devices that are all connected to networks in a seemingly ubiquitous environment. The purpose of this paper is to propose a framework for obtaining Internet content intended for computers on mobile devices. In other words, the primary concern of this paper is to select the best images for optimal performance and convert them into images that can be effectively reproduced on mobile devices. For this, the performance of the server is to be optimized through selecting high-priority images among the many available images and converting them while considering the display pixel rates (DPR) and the image distribution areas.

▶ Keyword : ubiquitous, Policy for Selecting Priority Images, mobile web service, Display Pixel Rate, Position Breadth First Search

• 제1저자 : 박대혁

• 접수일 : 2005.12.07, 심사완료일 : 2006.02.10

* 송실대학교 미디어학과 박사과정, ** 안양대학교 디지털 미디어공학과 교수,

*** 한국재활복지대학 컴퓨터게임과 교수, **** 한경대학교 컴퓨터공학과 교수, ***** 송실대학교 미디어학과 교수

※ 본 연구는 송실대학교 교내연구비 지원으로 이루어졌음.

1. 서론

다양한 기기간의 데이터의 전송과 모든 정보 통신 기술을 묶어서 융합하는 디지털 컨버전스 기술이 급성장하고 있다. 모바일의 고성능화로 디지털 컨버전스의 핵심에 모바일 단말기가 있다. 그래서 기존의 컴퓨터를 이용하던 서비스 기술들이 모바일을 이용하는 서비스 분야로 기술개발이 이루어지고 있다.

최근에는 유무선 통합된 서비스를 하기 위해서 콘텐츠 제작할 때, XML 형식을 제작하고 있다. 이와는 다르게, 기존에 제작된 많은 인터넷 콘텐츠를 재활용 하여 모바일 단말기에 직접 서비스하고자 하는 연구가 있다. 즉 모바일을 이용하여 웹 문서에 접근할 때, 웹 문서의 내용을 구조적으로 분석하고, 테두리로 분리하여 선택된 내용을 줄이거나 변형하여 모바일 페이지를 만들고, 사용하는 이미지를 모바일에서 재생할 수 있는 이미지로 변환하여 서비스 하는 시스템(6) 등의 연구가 있다. 또한 하나의 웹 페이지를 모바일 화면을 고려하여 점진적으로 접근하는 연구(7) 등이 있으며, 모바일에서 재생 가능한 이미지를 서비스하기 위해서 기존의 인터넷의 이미지의 크기 및 포맷을 변경하고 디바이스에 따라 다르게 서비스를 제공하는 연구(3)(4)가 있다.

다양한 플랫폼간의 데이터를 변환하여 서비스하는 것은 하나의 콘텐츠에 대해서 수많은 페이지를 생산하여 보관하여야 하는 문제점을 갖고 있다. 모든 웹 문서를 모든 플랫폼에 대해서 보관하면 서버의 부하가 커지고, 일부를 삭제하면 서비스의 양적 질적으로 손실을 갖는 등의 trade off 요소를 갖는다. 특히 디스크 I/O가 웹 프락시 서버의 성능을 심각하게 저해하는 중요한 요소 중의 하나임을 보이고 있다.[1][2] Mogul은 디스크 I/O 오버헤드가 캐시 히트를 높여 개선된 지연시간보다도 훨씬 크다는 연구 결과는 본 논문의 착안점이 되었다. 특히 디스크의 응답시간에 대한 문제에 대해서 Rousskov와 Soloview는 디스크의 지연시간이 객체가 히트(hit)되었을 때의 응답시간에 30% 정도의 영향을 미친다는 연구가 있다.[5] 따라서 본 논문에서는 웹 콘텐츠를 실시간으로 다양한 플랫폼에 컨버전스 한 후, 사용자에게 맞는 콘텐츠를 제공하고자 다수의 이미지 중에서 중요 이미지를 선정하고, 선정된 이미지만을 변환하여 서비스

하는 것을 제안하고자 한다. 이로 인해 웹 페이지의 모든 이미지를 변환하여 저장하는 것 보다 중요 이미지만을 변환 저장하면 서버의 디스크 효율성이 높아지고, 또한 다수의 사용자에게도 빠른 시간 내에 서비스를 할 수 있다고 예상된다.

논문의 구성을 보면 다음과 같다. 웹 데이터를 이용한 무선 인터넷 서비스를 위한 Trans-Gate System에 대한 동작 모델을 설명한다. 그리고 DPR(Display Pixel Rate)과 PBFS(Position Breadth First Search)를 이용하는 이미지 선정 정책을 제안한다. Trans-Gate System에 정책을 적용하여 실험 및 성능 평가한다.

II. 중요이미지 결정 기법

2.1 Trans-Gate System Architecture

웹 데이터를 기반으로 무선 인터넷 서비스하는 Trans-Gate 시스템은 유선 인터넷 콘텐츠를 유무선 통합 서비스 시스템으로 유선 콘텐츠를 사용자의 정보를 기반으로 컨버전스하여 서비스하는 서버 시스템 (그림 1)이다.

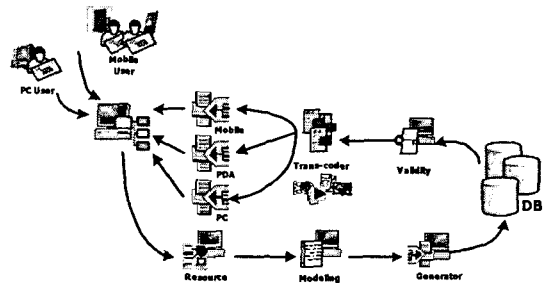


그림 1. Trans-Gate 시스템 구조
Fig. 1. Trans-Gate System Architecture

웹 콘텐츠는 사용자 정보를 기반으로 변환하고 서비스하기 위해서 Resource, Modeling, DB, Validity, Trans-coder로 구성된다. 이는 접속자의 정보를 기반으로 요구사항을 정하고, 요구사항에 맞게 서비스하기 위해서 HTML 코드를 변경하여 모바일 페이지를 생성하고, 웹 페이지에 있는 이미지를 변경하여 DB에 저장 한다. 위에 저장된 이

미지는 사용자의 요구가 발생하면 그때마다 수시로 재 변환되어 서비스 된다.

접속한 사용자의 정보를 기반으로 DB에 기록되어 있는 내용을 우선 인터넷 문서로 생성하여 서비스 하게 되는데 이때, 사용자의 플랫폼에 관한 정보와 하드웨어적인 지원 사항을 알 수 있다. 예를 들어 하나의 웹 페이지를 모바일 단말기의 SPEC에 맞도록 재구성 되어져 서비스 된다. 즉 WAP 기반 "LCD SPEC(2.0)260K TFT LCD (176*220)"이면 하나의 웹 페이지를 176*220으로 잘라서 보여지게 된다. (그림 2)는 CNN 페이지를 모바일 웹 페이지로 변환했을 때 하나의 페이지가 모바일 LCD에 맞게 다수의 페이지로 변환되어 계층적으로 서비스되는 것을 볼 수 있다.[6]



그림 2. CNN 웹 페이지의 모바일 서비스
Fig. 2. CNN Web & Mobile Web Service

2.2 이미지 픽셀

모바일 단말기의 디스플레이 발전으로 기존의 이미지 중심의 모바일 웹에 멀티미디어 요소 기술을 기반으로 하는 서비스가 적용되고 있다. 웹 페이지를 구성하는 다수의 이미지를 모바일에서 재생하기 위해서는 이미지를 읽어 재생하고자 하는 이미지의 크기와 포맷으로 변환 서비스되어야 한다.

서버의 응답시간을 결정하는 디스크 I/O를 최소화하기 위해서 Human Factor 중 인간의 시야(Human vision)를 고려한 인지층(Cognitive Layer)을 갖는 데이터가 중요 이미지이다. 이에 서버의 응답속도 및 서버의 기록용량 측면에서 서버의 역량을 확대하기 위해서 표시되는 화면의 크기(Display Pixel Rate)를 고려하여 이미지를 선정하고자 한다. 이를 위한 이미지 선정 정책은 다음과 같다.

$$tS = tR + t(WLsrc(1 + WRfirst * (1 + WRsecond))) + tW$$

..... (식 1)

tS : 이미지 변환 서비스 시간

tR : 디스크에서 Resource를 읽는 시간

WLsrc : Resource를 변환하는 시간

WRfirst, WRsecond : Resource 변환 필터

tW : 디스크에 기록하는 시간

웹 문서는 TEXT와 이미지를 포함하여 사용자에게 정보를 빠르게 전달하고자 한다. 이와 마찬가지로 모바일의 작은 디스플레이 장치에서 효과적으로 내용을 전달하기 위해 이미지는 매우 중요한 구성요소이다. 따라서 작은 화면에서 보다 효과적으로 내용을 전달하기 위해 웹 문서 내에서 중요 이미지를 찾고, 모바일 시스템에 재생 가능한 이미지 파일로 만들어야 한다. 그러나 웹 페이지에 있는 광고나 중요도 낮은 이미지를 제외하고 중요도가 높은 이미지를 찾아서 모바일 페이지로 제작하여 서비스 정책은 다음과 같다.

정책 1. Display Pixel Rate를 고려한 중요 이미지 선정 모바일 서비스

Policy 1. Policy for Selecting Priority Images by Display Pixel Rate

- step 1. 모든 이미지 정보를 검색한다.
- step 2. 이미지 정보를 분석한다.
- step 3. 재현율이 높은 이미지를 선정한다.
 - 01 이미지 정보를 이용하여 리스트 구성
 - 02 이미지의 표현되는 픽셀 정보를 기반으로 정렬
 - 03 정렬된 이미지 순으로 우선순위를 부여한다.
 - 04 이미지를 변환하는 정보인 XML의 attribute에 우선순위 정보를 기록한다.
- step 4. 선택된 이미지를 모바일 사용자에 맞게 크기, 포맷 변환한다.

이 정책은 모바일 사용자가 히트한 웹 사이트를 사용자의 모바일 단말기에 맞도록 변환시켜 서버에 저장하는데 걸리는 총 시간을 tS라 할 때, 다음과 같은 식을 얻을 수 있다.

또한 디스크의 응답시간에 대한 문제 연구 중에 Rousskov와 Soloviev의 히트 되었을 때의 응답시간이 30% 정도의 갖는다는 연구이다[5]. 즉 tR과 tW를 최소화하면 변환 서버의 처리 속도를 향상시킬 수 있다. 이를 위해서 다수의 이미지 중에서 중요 이미지를 선정하여 변환하는 정책을 제안한다.

2.3 크기 정보

정책 1에서 제안한 방법을 기반으로 CNN 페이지를 분석하면 다음과 같은 도표가 만들어 진다. 웹 페이지에 포함되어 있는 이미지 정보를 분석하면 (그림 3)과 같은 결과화면을 얻을 수 있다. HTML 소스 코드에 표현되어 있는 이미지에 관한 정보를 분석하여 현재 화면에 표현되는 이미지 픽셀로 정렬 한 모습이다.

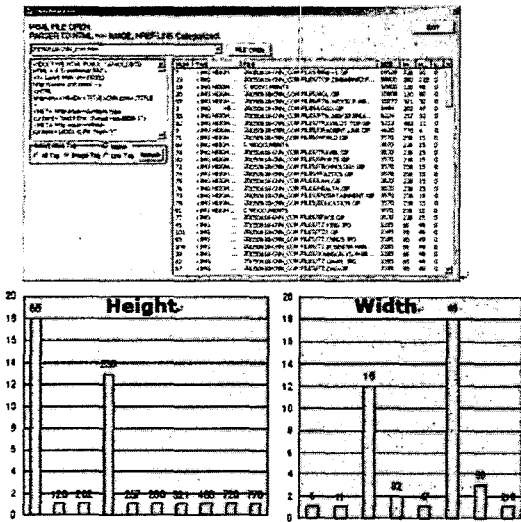


그림 3. HTML 코드 분석 데이터
Fig. 3. Data of HTML Code Analysis

그림은 CNN 페이지의 HTML 페이지를 분석한 것으로 이미지 Tag를 중심으로 분석한 것이다. 대부분의 웹 페이지는 특정 디자인을 적용하여 페이지를 제작하는 경우가 많아서 분석표에서 보듯이 특정한 해상도를 갖는 이미지들이 많이 분포되어 나타나는 부분이 있음을 알 수 있다. 이는 해당 웹 페이지에서 중요성을 나타내기 위한 일정 표현방식이라고 할 수 있다. 따라서 이미지의 반복된 크기 정보는 중요 이미지로 선정될 가능성이 높다.

2.4 위치 정보

일반적인 웹 페이지는 2-3개의 레이아웃으로 구성되는 경우가 많다. 중요 이미지 정책을 적용 할 때도 웹 페이지를 구성하는 레이아웃을 고려하여 모바일 페이지를 생성하여야 한다. 즉, 다른 레이아웃에 존재하는 이미지를 중요 이미지로 선정하는 경우가 발생할 수 있다. 따라서 중요 이미지 선정 시 웹 페이지에 표현되는 이미지 영역을 고려하여 중요 이미지 위치 선정 정책을 제안하고자 한다.

정책 2. 위치 선정 정책

Policy 2. Policy for Selecting Priority Images by Positions

- step 1. 모든 이미지를 정보를 검색한다.
- step 2. 정보를 분석한다.
- step 3. 위치 정보를 고려하여 이미지 선정한다.
 - 01 이미지 정보를 계층구조를 구성한다.
 - 02 Breadth First Search 방법을 이용하여 최상의 영역의 이미지를 선정한다.
 - 03 정렬된 이미지 순으로 우선순위를 부여한다.
 - 04 이미지를 변환하는 정보인 XML의 attribute에 우선순위 정보를 기록한다.
- step 4. 선택된 이미지를 모바일 사용자에게 맞게 크기, 포맷 변환한다.

웹 페이지는 페이지의 상단에 해당 웹 사이트의 로고 및 중요 광고를 담는 경우가 많다. 이러한 이미지를 이미지의 크기만을 고려한다면 우선순위가 매우 낮아질 수 있다. 이를 보완하기 위해서 웹 페이지의 상위에 표시되는 이미지를 조사하고, (그림 4)에서처럼 이미지 선정 정책과 함께 위치 선정 정책도 적용하고자 한다.

변환 서버를 이용하여 웹 데이터를 모바일에 전송하는 서버의 성능을 향상시키기 위한 이미지 선정 정책(Policy)을 유명 웹 페이지의 모바일 서비스에 적용한다. 시스템에 적용 전과 후의 실험을 통해서 이미지 선정 정책의 효과를 찾고자 한다.

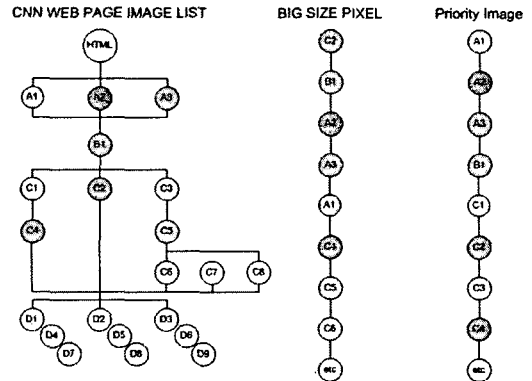


그림 4. CNN 웹 페이지의 이미지 리스트 계층도
Fig. 4. The Image List Level Chart of CNN Web Page

III. 실험 및 성능평가

3.1 실험 시스템

이미지 선정 정책의 성능 평가를 위해서 13개의 유명 웹 사이트를 분석하고 이미지 선정 정책을 적용하고자 한다. 표1은 현재 웹 서비스를 하는 인터넷 웹 페이지에 대한 정보로 웹 사이트 명과 메인 페이지의 크기, 이미지 수, 이미지들의 총 메모리 크기, 정보를 갖는 이미지에 수에 대해서 열거하고 있다.

표 1. 실험에 사용한 웹 페이지 분석 데이터
Table 1. Web Page Analysis Data Used in the Experiment

NO	웹 페이지	파일 (K)	이미지 수	용량 (K)	중요
1	CNET	158	78	161	35
2	Shutterfly	7	8	72	8
3	ESPN.com	63	38	109	14
4	National Geographic	44	26	107	11
5	Evite	90	49	58	9
6	Washington Post	129	38	63	23
7	CNN	68	72	173	13
8	LA Times	129	88	170	25
9	Empas	160	90	150	20
10	Lycos	77	23	77	15
11	Yahoo	85	29	32	11
12	NAVER	75	38	50	18
13	NASA	84	61	123	22

3.2 실험 결과

웹 페이지를 구성하는 이미지들의 분포율을 분석하고자 한다. 선택된 이미지를 변환하여 서비스하고자 할 때 기존의 웹 페이지의 어느 정도의 재생 율을 갖는지 분석하였다.

일반적으로 하나의 웹 페이지는 압축된 이미지 용량으로 110K 정도를 갖고 있으며 이는 평균적으로50개 내외의 이

미지로 구성된다. 큰 영역을 갖는 이미지를 기준으로 정렬 후 각각의 DPR을 보면 (그림 5) ①과 같이 나타난다. 즉 웹 페이지는 몇 개의 대표되는 이미지를 갖는 것을 볼 수 있다. 이미지의 DPR을 기준으로 정렬하여 누적하여 화면의 구성을 보면 ②와 같은 그래프를 갖는다. 대부분의 웹 사이트들에서 11개의 선정된 이미지의 화면 표시 율이 전체 표현하는 이미지들의 합에 비하여 70%이상 갖는 것을 볼 수 있다.

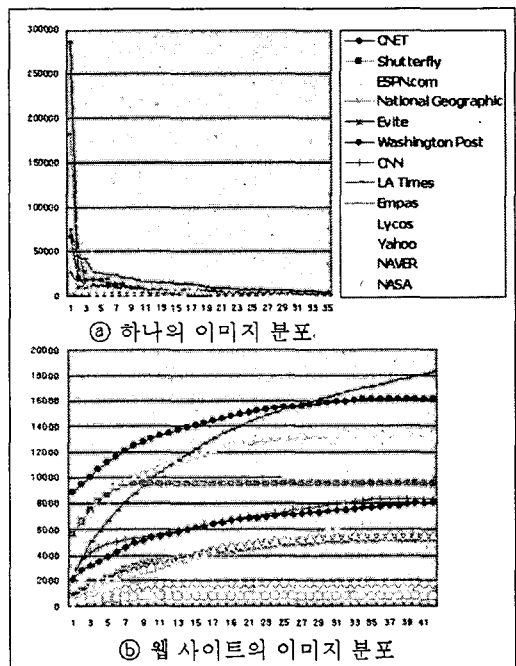


그림 5. 웹 사이트의 이미지 분포
Fig. 5. Image Distribution on Web Sites

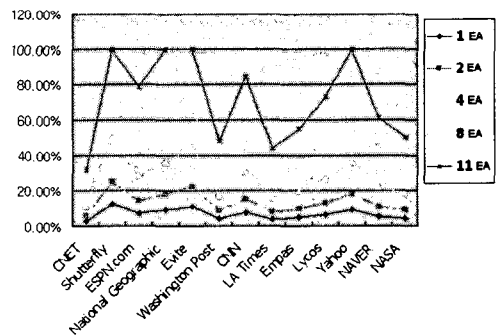


그림 6. 13개의 웹 사이트의 이미지 선정에 따른 재현율
Fig. 6. Recall Ratios for Selecting Images from 13 Web Sites

8개를 선정할 경우 1개의 웹 사이트는 100% 재현할 수 있다. 13개의 웹 사이트에서 평균 55%를 만족하는 것을 볼 수 있다. 이미지 선정 정책에 의해서 선택된 11개는 전체 웹 페이지를 구성하는 이미지 데이터의 31%에 해당하지만 전체 영역의 71%를 재생하므로 대단한 성능을 갖는 것을 확인 할 수 있다. 71%의 이미지가 웹 페이지에서 제공하고자 하는 콘텐츠의 90%를 갖는 것을 확인 할 수 있었다.

결론적으로 100K 내외의 HTML파일을 분석하여, 분석된 내용을 기반으로 Trans-coding(모바일에서 재생하기 위한 이미지 변환작업)을 하면 기존의 웹 페이지에 존재하는 모든 이미지를 모바일에서 재생 가능한 데이터로 변환하는 것에 비해서 4.5배 이상 효율적으로 작업이 이루어진다.

IV. 결론 및 향후 과제

유비쿼터스 환경이 되면서 컴퓨터를 이용하여 사용하던 인터넷 콘텐츠의 활용도가 점점 증가하고 있다. 즉 유무선 통합 서비스를 하기 위한 프레임워크인 Trans-Gate System이 많은 요청을 갖게 될 것이다. 본 논문의 프레임워크를 이용하여 다양한 플랫폼간의 데이터를 변환하여 서비스하는 가능하다. 즉 하나의 콘텐츠를 사용자에게 맞게 재구성하여 생산된 페이지를 서비스하는 것이다. 이때, 모든 웹 문서를 모든 플랫폼에 대해서 보관하면 서버의 부하가 커지고, 일부를 삭제하면 서비스의 양적 질적으로 손실을 갖는다. 이를 DPR과 PBFS 정책에 기반하여 선정된 이미지만을 서비스 하고자 한다.

많은 사용자가 접속하는 13개 웹 사이트의 메인 페이지에 대해서 실험해본 결과 평균 50개 이상의 이미지로 구성된 웹 사이트에서 11개를 선정하여 사용자 페이지를 구성하였을 때, 전체 이미지를 변환 한 것에 비하여 71% 분포를 갖는다. 또한 이용자에 대한 평가를 하였을 때 90%이상 필요정보를 찾은 것을 볼 수 있었다. 정리하면 서버의 성능은 4.5배 향상시키면서 기존의 모든 이미지를 변환한 것의 콘텐츠에 비해서 10% 정도 손실이 있는 것으로 볼 수 있다.

참고문헌

- [1] J. Wang and Dong Li. A Light-weight, Temporary File System for Large-Scale Web Servers. MASCOTS(2003).
- [2] K.Kant and P.Mohapatra. "Scalable Internet Servers : Issues and Challenges." ACM SIGMETRICS Performance Evaluation Review. Vol.28(2000).
- [3] In-Ho Lin and Bih-Hwang Lee. "Synchronization Model and Resource Scheduling for Distributed Multimedia Presentation System," IEICE TRANS. INF. & SYST., vol. 83-D, no. 4 (April 2000).
- [4] Chua H. N., Scott S.D., Choi Y. W., and Blanchfield P. "Web-Page Adaptation Framework for PC & Mobile Device Collaboration", IEEE AINA, 2005
- [5] Yonghyun Hwang, Jihong Kim, and Eunkyong Seo, "Structure-Aware Web Transcoding for Mobile Devices", IEEE INTERNET COMPUTING(2003)
- [6] A.Rousskov. and V.Soloviev. On Performance of Caching Proxies. In Proceedings of the ACM SIGMETRICS Conference(1998)
- [7] Chan Young Kim, Jae Kyu Lee, Yoon Ho Cho, Deok Hwan Kim. "VISCORS: A Visual-Content Recommender for the Mobile Web," IEEE Intelligent Systems, vol. 19, no. 6, pp. 32-39, November /December(2004).
- [8] Bo Shen, Ishwar K. Sethi, and Bhaskran Vasudev, "Adaptive Motion-Vector Resampling fo Compressed Video Downscaling", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 9, No. 6, September 1999.
- [9] Building Web Sites with XML, Michael Floyd, Prentice Hall PTR(2000).

저 자 소개



박 대 혁
 2004년 ~ 현재 : 숭실대학교
 미디어학과 박사과정
 <관심분야> 멀티미디어 기술, 무선
 인터넷, 시스템 소프트웨어



홍마리아
 2001년 숭실대학교 정보미디어학과
 공학석사
 2004년 숭실대학교 컴퓨터학과
 공학박사
 2004년 ~ 현재 : 안양대학교
 디지털 미디어공학과 교수
 <관심분야> 멀티미디어 스트리밍,
 MPEG21, 유비쿼터스, DMB



서 정 만
 1988년 ~ 1993년 엘지전자
 컴퓨터연구소 주임연구원
 1993년~1999년 삼성중공업
 중앙연구소 선임연구원
 2003년 충북대학교 컴퓨터공학과 박사
 2002년 ~ 현재 : 한국재활복지대학
 컴퓨터게임개발과 교수
 <관심분야> 컴퓨터게임,
 데이터베이스, 가상현실,
 실시간처리



이 근 수
 1983년 숭실대학교 전자계산학과
 공학사
 1988년 숭실대학교 전자계산학과
 공학석사
 1993년 숭실대학교 전자계산학과
 공학박사
 2003년 3월 ~ 2004년 2월 미국
 George Mason
 University, 전자계산학과
 객원교수
 1989년 ~ 현재 : 한경대학교
 컴퓨터공학과 교수
 <관심분야> 패턴인식, 퍼지이론,
 컴퓨터비전, 지식기반 시스템,
 동작이해, 비디오검색 등임.



임 영 환
 1985년 Northwestern University
 전산학과(박사)
 1979년 ~ 1996년 한국전자통신연구원
 책임연구원
 1996년 ~ 현재 : 숭실대학교
 미디어학부 교수
 <관심분야> 멀티미디어 기술,
 유비쿼터스, 모바일 분야