

모바일 환경에서 맞춤형 콘텐츠 전달을 위한 비디오 적응성 모델

김 스베틀라나[†] · 윤 용 익^{††}

요 약

이동하면서도 디지털 방송을 볼 수 있는 모바일 TV가 세계에 보급되기 시작하였다. TV 또는 HDTV와 같은 대화형 데이터 서비스를 제공하기 위해서 데이터 표현, 전송 및 처리가 아주 중요하다. 특히 제한된 대역폭과 낮은 성능의 단말과 같은 모바일 방송통신 환경에서 오디오, 비디오, 이미지 등의 다양한 멀티미디어 서비스를 제공 받기 위해서는 효율적으로 전송 및 적응형 데이터 서비스 기술이 필수적으로 요구된다. 본 논문에서 모바일 환경에서 사용자 중심으로 비디오서비스 대한 미들웨어 기술을 효율적으로 개발하기 위하여 사용자 환경상향인식(Usage Environment), 개인화(Personalization)를 위한 데이터 수정 및 처리 기술, 사용자 중심으로 맞춤형 적응 서비스 미들웨어 UP-SAM(User Personalization Context-Aware Service Adaptation Middleware) 모델을 제안한다.

키워드 : MPEG-21, 비디오 콘텐츠, 멀티미디어 서비스, 유비쿼터스 컴퓨팅, 미들웨어, Cognitive 콘텐츠, 개인화 콘텐츠

Video Adaptation Model for User-Centric Contents Delivery in Mobile Computing

Kim Svetlana[†] · Yong-Ik Yoon^{††}

ABSTRACT

Lately the usage of multimedia equipment with small LCD displays is rapidly increasing. Although many people use devices like this, videos intended for TV or HDTV are sent to these mobile devices. Therefore cases where it is hard for the user to view the desired scenes are growing more frequent. Currently, most services simply reduce the size of the content to fit the screen when they offer it for mobile devices. However, especially with sports broadcasts, there are many areas that cannot be seen very well because it was simply reduced in size. We therefore consider this weakness and are researching how to let the user choose an area of interest and then sending it to the user in a way that fits the device. In this paper, we address the problem of video delivery and personalization. For the delivered video content, we suggest the UP-SAM (User Personalized Context-Aware Service Adaptation Middleware) model that uses the video content description and MPEG-21 multimedia framework.

Keywords : MPEG-21, Video Content, Multimedia Service, Ubiquitous Computing, Middleware, Cognitive Content, Personalized

1. 서 론

유비쿼터스 환경에서 일반적인 단방향 서비스에서 벗어나 사용자의 의지에 따라 원하는 정보, 원하는 시간과 원하는 장소에, 원하는 형태로 서비스를 선택할 수 있을 뿐 아니라

필요로 하는 서비스를 요구 하지 않더라도 자신과 주변 환경의 변화에 따라 서비스들이 능동적인 의사결정을 하여 개인에 맞는 서비스를 제공해 주기를 원한다. 스트리밍 서비스에서 이런 기대를 만족시키기 위해서는 사용자의 이동에 따라 다양한 단말기로 자유롭게 이동할 수 있는 기술과 개인화된 기술이 필요하다. 개인화된 서비스 환경에서는 동일한 서비스라도 사용자 마다 서로 다른 콘텐츠를 사용하기 때문에 서비스를 실행하기 위한 상황인식 조건이 다르게 설정되어야 한다. 사용자의 요구상황에 맞는 서비스를 지원해 주기 위해서 우수한 성능의 적응 서비스 미들웨어가 요구되

* 본 연구는 숙명여자대학교 2008학년도 교내연구비 지원에 의해 수행됨.

[†] 준 회 원 : 숙명여자대학교 멀티미디어학과 박사과정

^{††} 종신회원 : 숙명여자대학교 멀티미디어학과 교수

논문접수 : 2008년 8월 6일

수정일 : 1차 2009년 6월 26일, 2차 2009년 7월 29일

심사완료 : 2009년 8월 3일

고 있다.

본 논문에서 기존 적응 서비스 미들웨어 보다 다양한 사용자의 환경 상황을 인식하여 사용자의 요구에 따라 맞춤형 개인화 서비스 구축을 위한 사용자 개인화 상황인식 적응성 서비스 미들웨어 UP-SAM (User-Personalized Context-Aware Service Adaptation Middleware) 모델을 제안한다.

2. 관련 연구

국외에서는 미국, 유럽을 중심으로 유비쿼터스 환경에서 사용자 중심 모바일 서비스 기술에 대한 연구 프로젝트가 이미 수행하고 있다. 영국의 DSG(Distributed System Group)에서는 CAMS(Context-Aware Mobile Service), CARMEN(Context-Aware Multimedia Environment for Narrative) 등의 프로젝트를 수행하고 있다.

국내에서 실제 유비쿼터스 환경에서는 일반적인 영상 데이터를 전송이 아닌 사용자 환경 상황과 요구 상황에 따라 다양한 서비스를 제공하기 위한 기술의 연구가 시도되고 있다. 또한, 여러 가지 예측하지 못한 사용자환경 상황에서도 실시간으로 적절하게 대처할 수 있는 적응형 서비스 미들웨어 기술의 연구가 국내에서 태동하고 있는 단계에 있다.

본 연구에서 사용자 중심 모바일 비디오 서비스를 포함한 멀티미디어 서비스를 위한 미들웨어 기술 개발을 크게 세 가지 주요 기술에 재한 연구이다. MPEG-21 프레임워크를 이용하여 네트워크 성능 및 사용자의 환경상황인식 기술에 대한 연구, 사용자의 요구 사항을 처리 하기 위한 개인화 기술에 대한 연구 그리고 개인화된 콘텐츠를 사용자 중심으로 전달을 하기 위한 맞춤형 적응 서비스 미들웨어 기술에 대한 연구이다.

본 장에서 서비스 적응성을 지원하기 위해 필요한 각 기능에 대한 기술과, 현재 연구 정도 그리고 문제점들을 도출한다.

2.1 사용자 환경 인식 기술

지능적 환경을 구현하기 위해서는 각 시스템들과 서비스들이 주어진 상황에 빠르게 대처하는 user environment context-aware 기술을 탑재하여 다양한 디바이스, 이동 사용자 및 빠르게 변환하는 상황(context) 정보를 적응적으로 서비스에 반영하는 기능이 요구되고 있다. 이러한 요구를 지원하기 위해서 context-aware 미들웨어 기술이 필수적이다.

그러나 현재까지 국내·외에서 개발 된 user environment context-aware 서비스를 지원할 수 있는 미들웨어 기술은 분산된 에이전트간의 상호작용 시 "Syntactic and Semantic Interoperability" (즉, 공통된 개념 또는 기준)를 제시하여 주지 못하는 문제점을 지니고 있다[2]. 이는 서로 다른 콘텐츠 처리 할 수 있는 미들웨어가 없기 때문이다[3, 4].

따라서 이러한 문제점을 해결하고 사용자의 요구를 주어진 상황에 최적으로 만족시켜 주기 위해서 MPEG-21 기술의 네트워크 환경 및 성능, 사용자 프로파일 등을 능동적으

로 지원하는 사용자 환경상황 인식 기술과 MPEG-7 User Preference내에 사용자 요구 상황 인식용 메타 정보 관리 기술을 포함하여 정확한 사용자의 요구 상황에 대한서비스를제공하는 모델의 연구가 필요하다[1, 6].

2.2 연구의 차별성 및 독창성

사용자 중심 서비스를 포함하는 멀티미디어 서비스를 지원하는 미들웨어 기술개발을 위한 세 가지 주요 기술이 필요하다. 사용자의 환경상황인식 기술, 사용자의 개인화 및 사용자의 요구사항을 처리 기술 그리고 서비스 전송 기술과 적용된 서비스를 사용자 중심으로 지원하는 맞춤형 적응 서비스 미들웨어 기술이다.

<표 1>과 같이 기존 연구 방법과 제안하는 논문의 차별성 및 독창성을 보여준다.

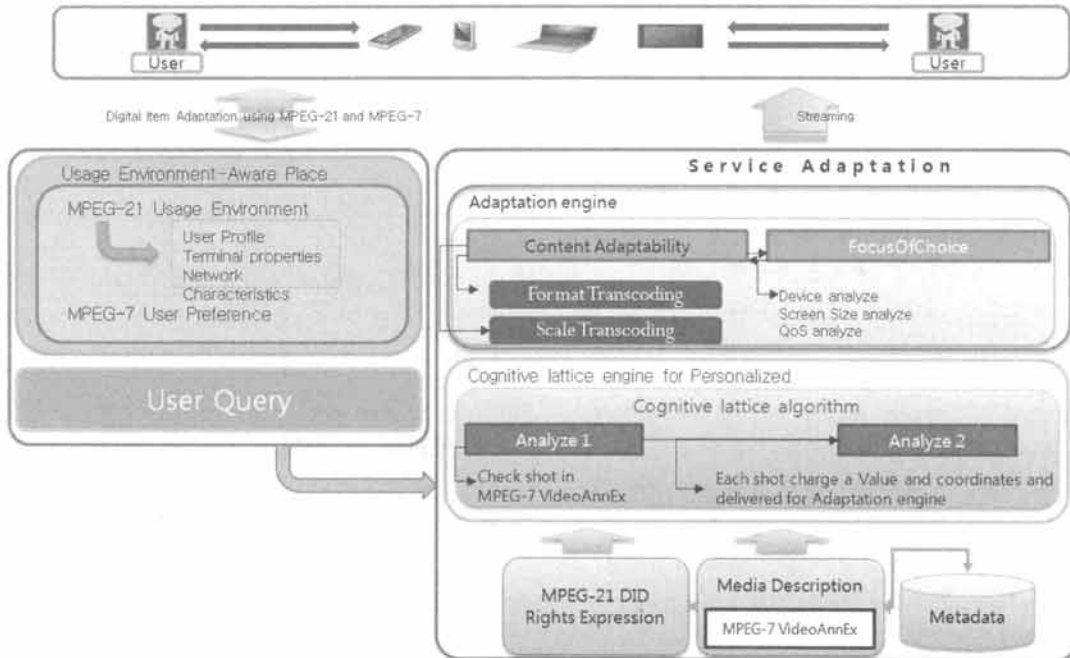
본 논문에서는 기존 적응성 서비스 미들웨어 보다 다양한 사용자의 환경 상황을 인식하여 사용자의 요구에 따라 맞춤형 개인화 서비스 구축을 위한 사용자 개인화 상황인식 적응성 서비스 미들웨어 모델을 제안한다.

<표 1> 관련 연구 동향 및 차별성

기존 방법	차별성 및 독창성
사용자 환경상황 인식 기술[5] Tsang,B.L.	
MPEG-21 DIA(Digital Item Adaptation)기술을 이용해 사용자의 환경상황의 대한 단순 인식	MPEG-21 프레임워크 이용하여 네트워크 성능, 사용자프로파일 및 디바이스 환경 등을 능동적으로 지원하는 사용자 환경인식 기술
사용자 중심 서비스[4, 5] Tsang,B.L.	
일반적인 단방향 서비스 사용자의 요구에 따라 확장 축소 할 수 있는 적응성 서비스 지원 안 함	사용자의 의지에 따라 원하는 정보, 원하는 시간과 장소에, 원하는 형태로 서비스 선택 가능 필요한 서비스를 사용자 주변 환경에 따라 개인에 맞는 서비스 실시간으로 제공
실시간 적응형서비스 미들웨어[5, 7] Tsang,B.L.,D.Tian	
일반적인 영상 데이터 전송 적응형 미들웨어 특성을 반영하지 못함	외부 환경 변화에 대한 적응 기술 사용자 중심으로 데이터 표현 및 전송 사용자의 요구에 따라 맞춤형 적응 서비스 미들웨어 제공

3. 사용자 개인화를 위한 상황인식 적응성 미들웨어 (UP-SAM) 모델

모바일 컴퓨팅 환경의 발전으로 언제, 어디서, 어떠한 단말기 상에서 인터넷과 접속이 되면 사용자가 필요로 하는 멀티미디어 정보를 정확하게 전달하는 시스템이 다양한 응용 분야에서 요구되고 있다[2, 3]. 본 연구는 유무선 통합환경에서 적합한 멀티미디어 콘텐츠 전달을 하기 위한 사용자 중심 모바일 멀티미디어 서비스 미들웨어의 연구 제안한다.



(그림 1) UP-SAM 모델 구조

사용자 환경 상황정보 인식 기술, 서비스 적응성 미들웨어 기술, 서비스 전달 기술, 리치 미디어 처리 기술 등을 습득하여 모바일 환경에서 최적의 멀티미디어 정보들을 최종 사용자에게 서비스 하기 위한 기반 기술을 전체적으로 도시한 것은 (그림 1)과 같다.

3.1 개인화를 위한 Cognitive Lattice 엔진

유비쿼터스 컴퓨팅의 특징 중 하나는 개인화된 서비스를 제공하는 것이다. 개인화된 서비스를 제공하기 위해서는 동일한 스트리밍 서비스라도 사용자 마다 각기 다른 요구사항에 대한 서비스 제공이 이루어져야 한다. 또한 사용자의 환경상황과 요구에 맞게 미디어 데이터를 편집하여 저작된 미디어 콘텐츠를 실시간으로 재생이 처리하도록 미디어 관리 기술과정들이 요구된다. 이러한 영상 미디어 데이터 관리를 위해서 Cognitive lattice 알고리즘과 MPEG-7 VideoAnnEx 편집 프로그램을 이용한다. Cognitive Lattice 엔진은 크게 두 아날라이즈로 구분하여 실행 한다.

[아날라이즈 1] MPEG-7 VideoAnnEx 이용하여 전송할 비디오의 프레임이 몇 샷으로 이루어져있는지 분석한 후 클로지업 샷과 롱샷으로 구분을 한다. 원본 영상 미디어를 샷으로 나누는 이유는 바로 클로지업샷 와 롱샷을 추출하여 더욱 정확한 화면을 제공해 주기 위해서이다. 샷을 구분하기 위해서 모든 카테고리 속성을 N 으로 정한다. 사용자의 preference vector은 $\vec{P}=[P_1, P_2, \dots, P_n]^T$, 여기서 p_i 속성의 $i, 1 \leq i \leq N$ preference weighting 표시한다. 모든 샷을 M 으로 표시하며, $\vec{S}=[S_1, S_2, \dots, S_M]^T$, 여기서 s_i 은 각 샷 넘

버로 $i, 1 \leq i \leq M$ 이다.

$$\vec{W} = \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_m \end{bmatrix} = \begin{bmatrix} cor_{1,1} & cor_{1,2} & \dots & cor_{1,n} \\ cor_{2,1} & \dots & \dots & \dots \\ \dots & \dots & cor_{i,j} & \dots \\ cor_{m,1} & \dots & \dots & cor_{m,n} \end{bmatrix} * \begin{bmatrix} p_1 \\ p_2 \\ \dots \\ p_n \end{bmatrix};$$

w_i 는 아날라이즈2에서 각 샷에 대한 값과 좌표를 매기기 위해서 구분을 하는 것이다.

[아날라이즈 2] 사용자에게 데이터를 더욱 효과적으로 제공하기 위해서 각 샷의 시간과 좌표 값을 아날라이즈를 한다. 롱샷의 관심영역을 찾기 위해서 반드시 좌표 값이 필요한 것이다. 아날라이즈 1을 통한 영상미디어 샷 패키지가 넘어오면 각 샷을 가로/세로 lattice로 나눈다:

$$v_num[k] = \sum (x_1) V_i[k],$$

$$h_num[k] = \sum (x_2) H_i[k].$$

이 과정을 통해서 각 샷에 들어있는 lattice에 대한 번호/값이 들어간다. 세로 lattice 대한 값은 i , 가로 lattice에 대한 값은 j 로 지정한다. For shot:

$$s_i = \sum_h^w k_i \sum_{v \in (i,j)} w_1, h_i[k_{i,j}]$$

Coords.x=getx(Object[found_y][foung_x]),
Coords.y=gety(Object[found_y][foung_x]).

Coords는 lattice가 들어있는 숫에 대한 시간과 좌표 값을 갖고 있으며, 한 오브젝트를 선택하여 관심영역을 선택을 할 수 있다. 완성된 작업을 사용자에게 전달을 하기 위한 적응성 서비스 엔진에 전달을 하게 되는 것이다.

3.2 적응형 서비스 미들웨어 엔진

적응형 서비스 미들웨어 엔진은 미디어 콘텐츠 특성을 분석하여, 사용자 중심으로 변환 및 특정부분에 맞춤형 서비스를 제공을 위한 적응형 알고리즘을 제안한다. 미디어 데이터를 사용자 중심으로 제공하기 위해서 Format transcoding 처리 기술, Scale transcoding 기술과 Presentation composer 등에 대한 전체적인 방향 계시를 한다. 실행 중에 있는 서비스의 외부 환경이 동적이기 때문에 모든 상황을 예측해야 한다[1, 6, 7]. 따라서 실행 중의 서비스는 외부 환경에서 발생하는 예측하지 못한 상황들에 대해서 대처하기 위한 기법이 필요한 것이다. 이를 위해서 미디어 콘텐츠를 사용자 중심으로 정확하게 효율적으로 전달을 하기 위해서 적응형 서비스 미들웨어 내에 콘텐츠에 대한 트랜스코딩을진행하는 FocusOfChoice 엔진을 제공한다.

적응형 알고리즘1단계에서 Content adaptability 기술을 포함한다. Content adaptability는 Format transcoding과 Scale transcoding 기술을 제안한다. Format transcoding 기술은 Cognitive lattice 알고리즘에서 받아들인 멀티미디어 콘텐츠를 비디오, 오디오, 이미지, 텍스트와 같은 구분을 하여 다음 단계로 전달을 한다. Scale transcoding 기술은 구분할 멀티미디어 콘텐츠를 사용자에게 디바이스에 맞게끔 비율과 사이즈 변경 및 영상이나 오디오 선명도 컬러 밝기 등을 증대시키는 기술의 신호 처리를 한다.

트랜스코딩 부분을 진행하는 FocusOfChoice 엔진 내의 적응형 알고리즘 마지막 단계에서 사용자 단말기에 맞춤형 적용 서비스를 제공하기 위해서Presentation Composer 제안한다. 정확한 사용자의 데이터를 알기 위해서 디바이스 아날라이즈, 스크린 아날라이즈, 및 QoS 같은 분석 기술을 포함되어 있다. Presentation Composer 의 콘텐츠 값을 V 사용하여 모든 콘텐츠를 트랜스코딩 구성을 C: C={I, r} 한다. 여기서는 $I \subset \{1,2,\dots,n\}$. 비디오의 클래스를 표현하여 r 는 축소된 트랜스코딩 비디오 값을 의미한다. 콘텐츠 값이V 는:

$$V = C(I, r) = \sum_{i \in I} V_i(r) = \sum_{i \in I} (s_i \cdot u(r - r_i)) \text{ 여기서}$$

$$u(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{elsewhere} \end{cases}$$

사용자의 디스플레이 사이즈를 표시하는 W(width) 와 H(height)는 각각 적용할 콘텐츠 리소스는 전달을 하기 위해서는:

$$\begin{aligned} &\text{Maximize } C(I, r), \\ &\text{such that } \begin{cases} r|x_u - x_l| \leq W \\ r|y_u - y_l| \leq H \end{cases} \text{ and,} \end{aligned}$$

모든 I maximum resolution factor 은 :

$$r_{\max}^j = \frac{\min r_{ij}}{i, j \in I},$$

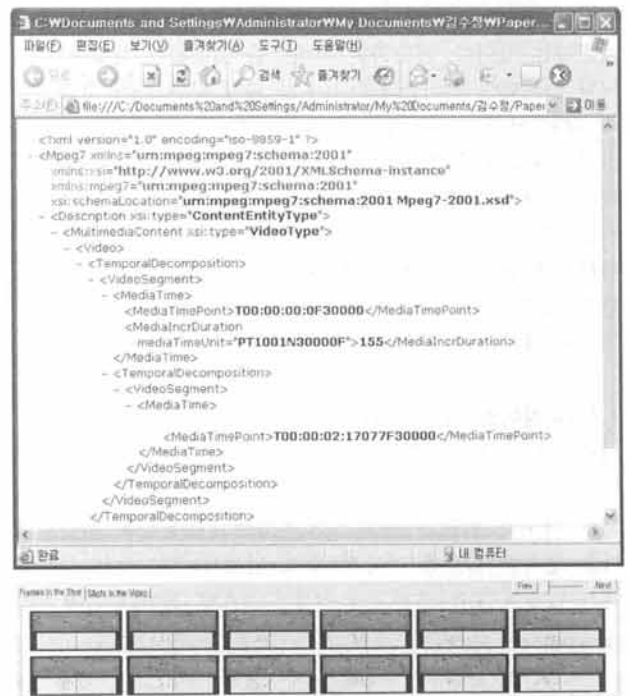
$$r_{i,j} = \min \left(\frac{W}{|x_i - x_j|}, \frac{H}{|y_i - y_j|} \right)$$

따라서, $C = \{I, r_{\max}^j\}$ 인 경우에는 I의 maximum 값이 될 수 있다. 그러므로, 멀티미디어 콘텐츠를 전달을 하기 전에 사용자의 디스플레이 사이즈의 맞는 콘텐츠 트랜스코딩 하게 되면, 다음 콘텐츠를 전달을 할 때 반복적인 디스플레이 분석이 할 필요가 없게 된다.

4. 구현과 및 검토

(그림 2)와 같이 비디오 콘텐츠를 MPEG-7 XML 스키마로 변환을 하여, 각 프레임을 샷으로 나누는 결과 화면을 보여주고 있다.

대부분 축구 같은 비디오 경우에는 클로즈업 샷으로 보여주기 때문에 선수들의 위치를 파악 할 수 없다. 그러한 단점들은 보완하기 위해서 가상 축구장을 만들어 lattice 값을 갖는 오브젝트의 위치를 파악을 할 수 있는 기능을 추가를 했다. 다음 (그림 3),(그림 4)와 같이 왼쪽에 MPEG-Player에서 적용할 비디오를 실행하는 화면이고, 오른쪽에는 FocusOfChoice 알고리즘을 통한 모든 디바이스에 비디오를 적용 수 있게끔 파일을 mp4, H264 또한 녹음 된 오디오 파일을 Mp3, Wma, Ogg, Mpeg-1, Mpeg-2로 변환시켜 주어 보여주는 결과 화면이다.



(그림 2) 프레임 샷으로 나누는 과정

5. 결론 및 향후 연구

현재는 제공되고 있는 다양한 모바일 서비스는 사용자의 요구를 반영하며 끊임없이 진화하고 있다. 차세대 이동통신 망 중심의 융합과 이동통신 단말의 컴퓨팅 능력의 고도화가 이루어지면 사용자가 언제 어디서나 네트워크에 접속할 수 있는 유비쿼터스 환경이 형성될 것이며, 향후의 이러한 환경에서의 모바일 서비스는 이동 컴퓨팅 및 네트워킹 환경에서 수시로 바뀌는 사용자의 상황정보를 인식하여 특정 시점에 가장 적합한 정보나 서비스를 제공해주는 특성을 가지게 될 것이다.

본 논문에서는 모바일 서비스의 현황 및 사례를 중심으로 모바일 서비스의 동향을 분석하였으며, 향후 유비쿼터스 환경의 주요모바일 서비스로 상황인식 서비스를 제시하고 가능한 응용 서비스를 제안하였다. 또한 모바일 리치미디어 서비스 기술은 이동멀티미디어 방송 환경에서 오디오/비디오 콘텐츠를 사용자의 환경상황과 요구에 맞게 미디어 데이터를 편집 저장하고, 저장된 미디어 콘텐츠를 실시간으로 재생 처리하도록 미디어 관리 기술과정들이 포함하였다. 본 논문에서 UP-SAM(User Personalization Service Adaptation Middleware)모델은 현재 사용자가 원하는 서비스를 제공해 주기 위해서 기존 적응 서비스 미들웨어 보다 사용자의 환경 상황을 인식하여 사용자의 요구에 따라 맞춤형 콘텐츠 전달을 위한 비디오 적응성 서비스를 효과적으로 제공할 수 있는 핵심 기능을 제안한다.



(그림 3) 비디오 트랜스코딩 결과 화면



(그림 4) 모바일 환경에서 비디오 트랜스코딩 결과 화면

참고 문헌

- [1] Lan S.Burnett, Fernando Pereira, Rik Van de Walle, Rob Koenen (ed), The MPEG-21 book, John Wiley & Song.Ltd (2006).
- [2] W.H. Adams, G. Iyengart, C.Y.Lin, M.R. Naphade, C.Neti, H.J. Nock, J.R. Smith, Semantic indexing of multimedia content using visual, audio and text cues, EURASIP J.Appl. Signal Process.2003(2) (February, 2003).
- [3] Belle L.Tseng, Ching-Yung Lim, John R.Smith, Video personalization system for usage environment, Multimedia Database Management System 15 (September, 2004).
- [4] Tseng, B.L., Lin, C.-Y., Smith, J.R., 2002. Video summarization and personalization for pervasive mobile devices. In: SPIE Electronic Imaging 2002—Storage and Retrieval for Media Databases, San Jose, (January).
- [5] Tseng, B.L., Lin, C.-Y., 2002. Personalized video summary using visual semantic annotations and automatic speech transcriptions. In: IEEE Multimedia Signal Processing MMSP 2002, St. Thomas, (December).
- [6] Fernando Pereira, MPEG Multimedia Standard: Evolution and Future Developments, MM'07, Augsburg, Bavaria, Germany September, (2007).

- [7] D. Tian, M. Hannuksela and M. Gabbouj, Sub-sequence video coding for improved temporal scalability, Proc. of IEEE Int. Symposium on Circuits and Systems (ISCAS), Kobe, Japan (2005).



김 스베틀라나

e-mail : xatyna@nate.com

2004년 숙명여자대학교 멀티학과(이학사)

2007년 숙명여자대학교 멀티미디어과학과
(이학석사)

2007년~현 재 숙명여자대학교 멀티미디어
과학과 박사과정

관심분야: 유비쿼터스 컴퓨팅, 분산 미들웨어, 모바일 에이전트,
MPEG-21



윤 용 익

e-mail : yiyoon@sm.ac.kr

1983년 동국대학교 통계학과(이학사)

1985년 한국과학기술원 전산학과(공학석사)

1994년 한국과학기술원 전산학과(공학박사)

1998년~현 재 숙명여자대학교 멀티미디어
과학과 교수

관심분야: 유비쿼터스 컴퓨팅, 멀티미디어 시스템, 분산시스템,
실시간 처리시스템, 미들웨어, 실시간 OS/DBMS, 실
감 미디어