Cognitive Lattice for Service Adaptation on Ubiquitous Computing

Svetlana Kim and Yong-Ik Yoon
숙명여자대학교 멀티미디어과학과, 분산시스템 연구소
xatyna@nate.com._yiyoon@sookmyung.ac.kr

Abstract Lately the usage of multimedia equipment with small LCD displays is rapidly increasing. Although many people use devices like cell phone and PDA, videos intended for TV or HDTV are sent to these mobile devices. Therefore, situation 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 are suggesting new motion how to let the user choose an area of interest based on Cognitive lattice. And present then sending it to the user in a way based on FocusOfChoice that fits the device.

1. Introduction

Users of Ubiquitous Computing are not concerned about time and place. At the same time, they expect to be provided with services that grant access to information. The service might be due to the increased accessibility of technologies on a broad level. It confronts service providers with the need to customize those services according to the user's receiving device which could be a handheld, or a wide screen terminal. Otherwise the provided information could be inadequate. If a minimal standard is applied users at a wide screen device would suffer under-represented information while with a maximum standard users of handheld devices might not be able to properly view or use the service at all. It should therefore be clear that a way to optimize is needed, a way of adapting the provided service to match the device in use to maintain or even enhance user-friendliness. This is the only way to cope with the technological convergence of devices which would grant access to the same services.

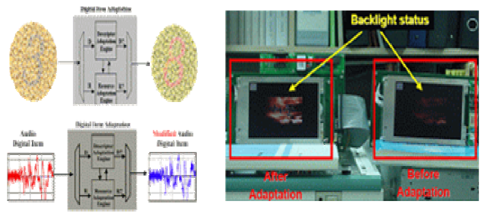
Within this paper it is our aim to probe into the use of dynamic components which make up Ubiquitous Computing and to actively reflect upon the users' needs in order to be able to provide flexible context-oriented services. A cognitive lattice system as well as service adaptation algorithms will be looked into in order to facilitate service adaptation. These are needed to recognize and adapt to changes of the multimedia context (using video streams). Ultimately a middleware model based on a cognitive lattice device for service adaptation is to be developed. This model's agenda is to overcome the limitation of current persistent services and allow for well-balanced operation between service and applications employing several adaptation levels as well as providing customized services for video stream changes on user demand.

2. Related work

Our research focuses on a service adaptation model so users in a ubiquitous environment can get the wanted information on their device in the desired way. To offer adapted services in a ubiquitous environment the following technologies are needed: MPEG-21 DIA (Digital Item Adaptation) and Multimedia Mining. We will describe each technology, how they are currently used and how we will use them in our research. Also, if there are problems using these we will address them.

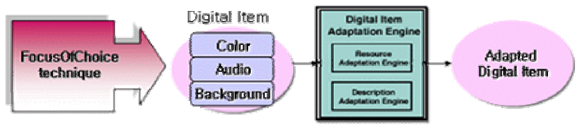
2.1 MPEG-21 DIA

MPEG-21 DIA is useful to us because it takes into account the user's characteristics, context information, and network or terminal characteristics [3]. Using this detailed data it defines a systematical standard. With this standard it is possible to offer widespread multimedia access to content consumers. Currently, MPEG-21 is used to convert images. For example <Figure 1>, image size, quality, or audio qualities are changed.



<Figure 1>MPEG-21 DIA for color, audio and backlight

This functionality is similar to what we want to accomplish. Therefore we will see how we can build on MPEG-21 DIA<Figure 2> and extend it to offer better service adaptation to users.



<Figure 2> FocusOfChoice Technique build on MPEG-21 DIA

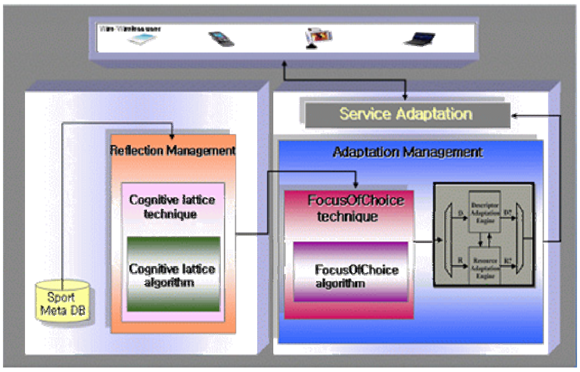
FocusOfChoice technique that is able to send the user's area of interest to the user in a way that retains quality and fits the device. But, it uses a MPEG-21 DIA, more it will be able to deliver the service which is quality. There will be like color vision deficiency and audio preference function to MPEG-21 DIA, when live broadcast delivering the user is convenient more and it will be able to use.

2.2 Multimedia Mining

Multimedia mining enables us to analyze multimedia context according to parameters like image size, color histogram, descriptor, shape texture and extract this information. While this is already very useful for our purpose we will add functionality. Splitting video contents into frames and dividing those frames into lattices to use multimedia mining's innate functionality of image awareness.

3 Cognitive Lattice for Service Adaptation

One of the major constraints of the access ways to support present adaptation is that the application is required to directly execute an adaptation mechanism when the lower level informs about an occurred change [8]. Even if such an application is able to catch the change of the running environment, it is more efficient if the middleware manages such adaptation mechanisms. The middleware structure for ubiquitous computing has to be recognizable, and the applications executed on middleware have to be adaptable to the context change by using diverse logics.



<Figure 3> SAM based on cognitive lattice architecture

Show in <Figure 3> Middleware Model based on a cognitive lattice for service adaptation is model of cognitive technique and focusofchoice technique. A cognitive technique and focusofchoice technique that is able to send the user's area of interest to the user in a way that retains quality and fits the device.

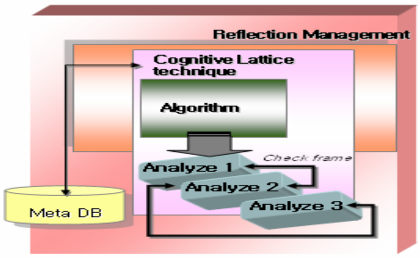
Ubiquitous computing allows application developers to build a large and complex distributed system that can transform physical spaces into computationally active and intelligent environments [6]. Ubiquitous applications need a middleware that can detect and act upon any context changes created by the result of any interactions between users, applications, and surrounding computing environment for applications without users' interventions [4].

This paper delves into dynamic components forming ubiquitous computing contexts to actively reflect on users' needs and provide flexible services for users.

4 Reflection Management

4.1 Cognitive Lattice technique

When looking at user in today's ubiquitous environment the amount of small devices is growing very fast [7]. The number of users watching live broadcasts on these devices is steadily growing. However, most services do not take the devices' display size into account and simply send content meant for TV or HDTV after reducing the size [5][6]. Because the displays are so small, many users have trouble viewing the content. For example, viewing soccer matches recorded with a longshot camera it is very hard to see the ball and players because they are extremely small [9]. Therefore it is necessary to develop a way of sending the part that interests the user to his device in a fitting way. So far, the user could choose an area of interest but this area was simply enlarged. Therefore there was the problem that the quality was low and the image blurry. To resolve this problem we propose our cognitive model. The cognitive model divides the frame into a lattice using multimedia mining show in <Figure 4>.



<Figure 4> Reflection Management

For instance, when viewing a soccer game that lattice is used to recognize the location of the ball. Using that location as the center only a part of the frame is selected and sent to the FocusOfChoice Model.

4.2 Defining the area of interest

The screen of soccer game consists of *long shot and not long shot.* First of all, we will divide the frames of the video and number each frame.

$$Video =< Shot_0,..........Shot_{n-1} >$$

$$(\mathbf{N}: \text{number of video frame}) \qquad (1)$$

Shot is an image sequence acquired through a single camera recording process. So we can describe Shot using a sequence of frames.

$$Shot_n =< F_n^0,.......F_n^{k-1} > \qquad (2)$$

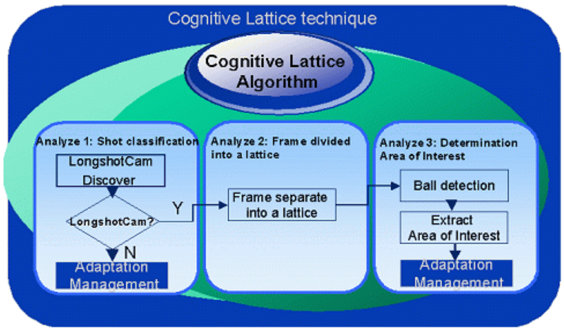
Here, F^0 indicates the i th frame of a Shot, k in F^{k-1} is the number of frames in a Shot. Also from now on, we will represent the i th frame of the whole video a F^i.

We can represent the area of interest in the i th frame of N th shot as *Object* (f^i_n). Based on the type of Shots, *Object* (f^i_n) can be an entire frame or part of it. Mathematically expressed,

$$Object(f_n^i) \subset f_n^i, ifClass(Shot_n), isaLongshotCam$$

$$Object(f_n^i) = f_n^i, ifClass(Shot_n), isaNotLongshotCam$$

This means that if the image frame is not a long-distance shot, *Object* is defined to be the same as the original image, while if it is a long-distance shot, Object is defined as a part of the image frame. Area of interest takes a small portion in a frame for a LongshotCam case, whereas area of interest is determining identical to the original image otherwise. The diagram of the proposed algorithm is <Figure 5> shows an example.



<Figure 5> Diagram of the Cognitive Lattice algorithm

In case of soccer video analysis, the ground color plays an important role, especially in distinguishing shot classes. We used color descriptors to differentiate *LongshotCam* and other shots. By classifying the colors of the frame, determine L with the greatest MGS (Much Green Segment). By measuring the amount of MGS in a frame, we can find out the type of Shot. If MGS appears to be less than Θ_L we determine the Shot to be something other than LongshotCam.

After the Shot ending decision each frame is divided into either LongshotCam frame or NotlongshotCam frame. For NotlongshotCam frame, there is no need to establish an area of interest smaller than the entire image frame. It is sufficient for the user to show the entire frame. But for LongshotCam frame, there is a need to magnify and replay the area of interest, the location of the ball. Therefore, in order to find the location of the ball, we divide the frame into lattices.

$$v\_num[k] = \sum (x_1)V_i[k] \quad (7)$$

$$h\_num[k] = \sum (x_2)H_i[k] \quad (8)$$

Here, we divide the frame vertically and horizontally. Next, we have to number each lattice. Vertical lattice is assign the value i, horizontal lattice the value j.

$$if\,(OCL[i][j], contains(shape, color, form))$$

$$if\,(analizeColor, shape, form \sum \frac{Vi[k]}{Hi[k]} \{PS, S, L, B\})$$

Since there could be several objects in each lattice, we must analyze them one by one.

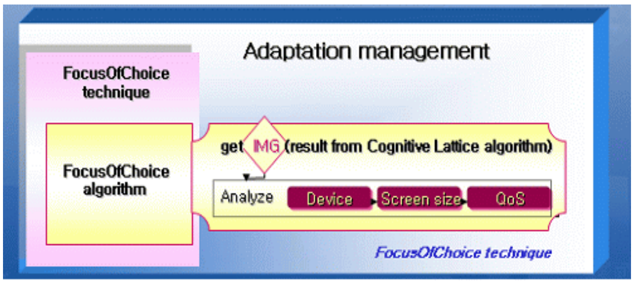
$$found\_x = j; \quad (15)$$
$$found\_y = i; \quad (16)$$
$$Coords.x = getx(Object[found\_y][found\_x])$$
$$Coord.y = gety(Object[found\_y][found\_x])$$

Here, Coords indicates the value of ball location. Find the value of the location by analyzing each lattice with the help of ball's fixed data based on multimedia mining. Next, send the value of Coords to Adaptation management and display it on the viewer's device through FocusOfChoice technique.

5 Adaptation Management

5.1 FocusOfChoice technique

In this paper we suggest a FocusOfChoice technique to supply users with services that are adapted to dynamic situations in a ubiquitous environment. After the requested data is analyzed using the cognitive lattice model, it should be adapted to fit the user device [1][2]. For this, we propose our FocusOfChoice technique. The FocusOfChoice technique's algorithm receives the cognitive lattice model's output and analyzes the user's context data, device information such as screen size and QoS. For example, let's say the cognitive lattice analysis returned the ball's location. Now a high quality adaptation is needed before the data is sent to the user. For this, we suggest our FocusOfChoice algorithm <Figure 6> shows an example. After determining the user's environment, the adapted frame is sent to the user.



<Figure 6> FocusOfChoice technique

5.2 FocusOfChoice Algorithm

Since the FocusOfChoice algorithm best satisfies user demands that are expressed by way of preferences and also actively chooses the most suitable service application to send, depending on the context, it is an algorithm that could possibly influence adaptation. The FocusOfChoice algorithm delivers user's correct service adaptation course after analyzing the device, screen size, QoS. For this, it uses the MPEG-21 DIA technology. Using the cognitive lattice algorithm's output it crops out the desired part of the frame in a way that best fits the user's device. The FocusOfChoice algorithm is shown in below <Algorithm 1>

Step 1 : (UD :User Device (Device information was formed color, size, height and width))
INPUT : Value of UD

OUTPUT : true or false

```
size            =  UD. size;
height          =  UD. height;
width           =  UD. width;
colorDepth      = UD. colorDepth
```

Step 2 : analyze screen size to send correct information. (Screensize : size, height, width, colorDepth))

INPUT : value of Screensize,Coords,Time stamp OUTPUT : New image

```
image = get image(time stamp)
center_x = coords.x
center_y = coords.y
Focus_i = Coords_i
Dist_i = 0 ,
Diff = Coords_i (coords.x, coords.y)
- Focus_i-1
New image = grop image (size,
height, wight, colorDept,
center_x, center_y)
0 <= diff <= dist_i
Focus_i = Focus_i
dist_i < diff <= frame.size / 2
Focus_i = Focus_i-1 + dist_i
```

Step 3 : analyze QoS (Quality of Service)

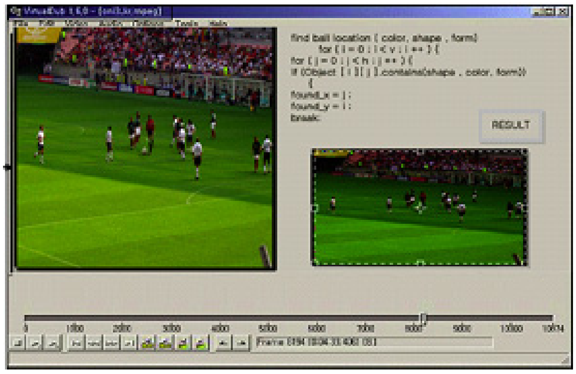
INPUT : value of QoS , New image

OUTPUT : adapt image

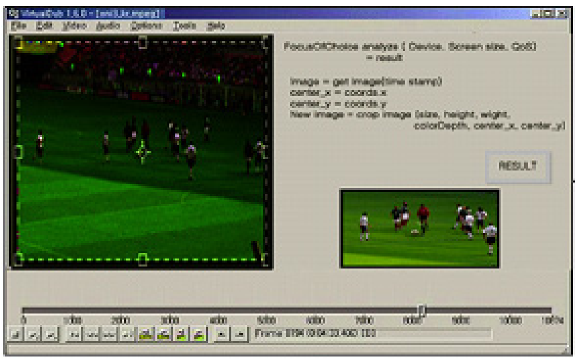
```
if (QoS < MaxQ ) {
img  = compressImage( MaxQ,QoS,
img);
img  = reduceColor( MaxQ,QoS, img);
}
```

6 Implementation Scenario

In <Figure 7> shows the implementation scenario structure for method which the display does the football screen. The soccer game screen is roughly composed of Longshot and shots that are not.



<Figure 7> (a) Before adapting Middleware model based on Cognitive lattice for service adaptation



(b) After adapting Middleware model based on Cognitive lattice for service adaptation

Longshot since Longshot frames have been filmed from a distance, when viewed in a small screen the size of the ball and the players is small. So, before transmitting the screen to the viewers one must check whether it is LongshotCam or not. Therefore for better display on a small screen, frame must be divided into two types and in the case of a long-distance shot, part of the entire screen must be shown.

The simplest way of establishing the area of interest is to place the center of the screen on the location of the ball. Based on multimedia mining, cognitive lattice technique divides the long shot frame vertically and horizontally, consequently finding the location of the ball from each lattice through comparing it with the saved data of the balls' form, shape, and color. After the location value is sought, it is sent to FocusOfChoice technique. This technique, after determining the position of display screen, must finally process and show it to the viewers.

7 Conclusion and Future Studies

Recently, mobile devices are used much like PCs. Viewing live video broadcasts on these devices is very popular. A lot of research is happening to provide these users with good service. However, there are still many things that should be improved. Though viewing shows like news and drama on a small display is not really a problem, watching live sports broadcasts poses many problems. In this research we Fited on sports and suggest a way of providing users with a way to comfortably view the content he wants even on a small display.

In the future, it is necessary to think of ways to support other kinds of sports. By researching which scenes are shown the most for any given kind of sport it should be possible to implement an algorithm that recognizes a standard view. For example, with baseball the classical views of the pitcher, the batter, or the scoreboard. Once those are established they can be used to automatically enlarge areas of interest like single players, for instance. In this way, every popular kind of sport could be covered and be suitably adapted for mobile device users.

References

[1] University of Illinois at Urbana-Champaign, "Framework for Ubiquitous Multimedia Applications," IL 61801. (2002).

[2] W. H. Cheng, "Automatic Video Region-of-Interest Determination Based on User Attention Model," In Proc. ISCAS (2005).

[3] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," In Proc. IEEE Tr. On Pattern Analysis and Machine Intelligence, pp. 1254-1259, (1998).

[4] K Wan, X Yan, X Yu, C Xu, "Real-time goal-mouth detection in MPEG soccer video," In Proc. of the eleventh ACM international conference on Multimedia, (2003).

[5] L Xie, P Xu, SF Chang, A Divakaran, H Sun, "Structure Analysis of Soccer Video with Domain Knowledge and Hidden Markov Models," In Proc. Pattern Recognition Letters, (2004).

[6] Y. Seo, S. Choi, H. Kim, K.S. Hong, "Where Are the Ball and Players? Soccer Game Analysis with Color Based Tracking and Image Mosaick," In Proc. of the 9th International Conference on Image Analysis and Processing-Volume II, (2005)

[7] F. Simone, F. Hanane, F. Frank, C. Kiho, P. Ramjee, "A User-Centric system, In Proc. CTIF,(2004).

[8] V. Anthony, Senior Member, IEEE, T. Christian, "Digital Item Adaptation," In Proc. Overview if Standardization and Research Activities, TR2005-056, (2005).

[9] Bill Schilit, Norman Adams and Roy Want, "Context-aware computing applications," In Proc. of IEEE Workshop on Mobile Computing System and Applications, (2002).