Technology Reports

NVAS Implemented in NMN —Service Prototype and Enabler Use Cases—

NVAS build on core services in the operator networks to deliver a diverse and enhanced service experience to the user. This may include for example personalized multimedia enhancements of conventional text messaging services such as online chat and instant messaging, and new mash-ups of mobile services with the popular social Web. The general expectation is that NVAS will be massively spurring subscribers to use their phones more often, thus allowing operators to substantially drive up their ARPUs. In this article, we present an overview of the testbed we produced as a practical use case implementation of our NVAS research at DOCOMO Euro-Labs.

DOCOMO Communications Laboratories Europe GmbH

Michael Fahrmair
Matthias Wagner
Roman Katz
Rami Ajaj
Tim Weiss

1. Introduction

The mobile communication market is remarkable in its rigorous demands for new services in a highly competitive and dynamic environment. Driven by such high demands, our research related to Network Value-added Services (NVAS) at DOCOMO Euro-Labs is focused on multi-purpose technologies and enablers[*1] for the rapid creation of novel mobile services on top of the NTT DOCOMO network. Our goal is to develop unique enabling technologies that would add essential value to established products and services while at the same time accelerating the introduction of new services to the customer. To this end, DOCOMO Euro-Labs is implementing various enabler technologies in a generic Service Enabler Network (SEN), and is studying their integration with existing platforms such as IP Multimedia Subsystem (IMS)[*2] and Web services [1].

Stimulated by the work of the GSM Association (GSMA)[*3] working group on the Rich Communication Suite (RCS)[*4] [2], we regard NVAS as services that build on core services in the operator networks to deliver an outstanding service experience to the user. NVAS also enables third-party developers the further utilization of distinct enablers and features in the NTT DOCOMO network. Our hope is that this will result in novel value-added services that will massively spur subscribers to use their phones more frequently, thus driving up Average monthly Revenue Per Unit (ARPU)[*5] and revenues.

Novel value-added services that are currently under investigation include rich communication services implemented by applying multimedia technology to text-based communication services such as online chat and instant messaging[*6] as well as mash-ups of

*1 **Enabler**: A constituent part or feature of a service that is used from multiple service scenario control units.

mobile services with the popular social Web.

In this article, we describe the rich communication and smart media processing aspects of NVAS research at DOCOMO Euro-Labs.

2. Rich Communication and Smart Media Processing

Figure 1 depicts the basic scheme of NVAS. Through the flexible combination of key service enablers, a unique and distinct service is created inside the operator's network. This service enabler network shall be generic for flexible and scalable development, yet at the same time comprising individual enabling technologies and features to add competitive extra value to our services.

Currently, our NVAS enablers are primarily geared towards adding value to communication services, but also allow extensions to other Web or Internet elements. To this end, our research is centered on Rich Communication and Smart Media Processing with a focus on multimedia technologies. In particular, through the enrichment and extension of multimedia services, we hope that users will benefit from being able to access services that are more intuitive and enjoyable, support natural communication, and are simpler and quicker to use.

Table 1 shows a selection of service enablers from the domains of Rich Communication and Smart Media Processing that have been integrated on the test-bed at DOCOMO Euro-Labs. These enablers are currently running on our NVAS testbed.

In particular, communication-based use cases are well-grounded in everyday life and society. Transferred to new technical media and given the right enabling technologies, they promise to improve the user experience while at the same time being easy and fun to adopt. At DOCOMO Euro-Labs, we are focusing our effort on enablers that will allow for a vast improvement of communication between users, such as augmenting text messaging with personal attributes like speech or handwriting. We also intend to reduce the total perceived pain of adoption for users in existing services like photo or video sharing by integrating new concepts such as content recognition and automatic image manipulation.

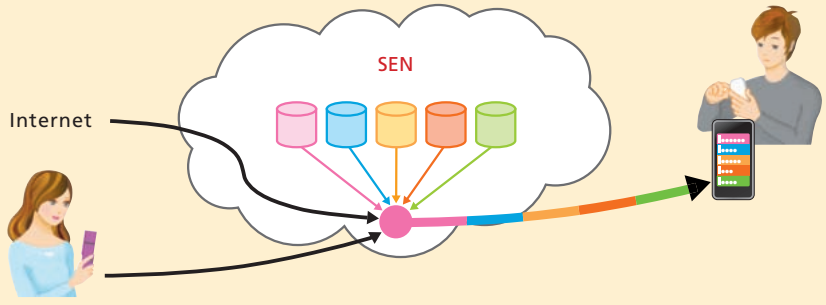

Figure 1 Basic configuration of NVAS

3. Rich Communication: NVAS Use Case

As an NVAS use case, **Figure 2** shows a Personal Video Message service that was presented at Mobile World Congress 2011. In this service, an enabler in an operator network is used to enrich simple text messages with added value, thereby providing users with content that has greater visual and emotional appeal. Before a message reaches its destination, three independent technologies are used to achieve this goal. In the first step, a Web based (external) translation enabler can be used to translate the original text message from English to Japanese or vice versa. Next, the message is personalized by a handwriting enabler. A user-specific handwriting font is selected so that the message can be displayed in a personal handwritten style. Finally, the message is delivered by the user's avatar[*7]. This avatar can, for example, be created from an existing photograph of the individual. 3D face modeling technology can be applied to animate the mouth movement and generate facial expressions.

*2 **IMS**: A communication system standardized by 3GPP for implementing multimedia services. IMS used IP and the SIP protocol used for Internet telephony to integrate the communication services of the fixed telephone and mobile communication networks.

*3 **GSMA**: World's largest industry association in the mobile communications domain.

*4 **RCS**: A concept of communication services realized in the IMS such as Presence, Instant Message and Video Sharing.

*5 **ARPU**: A useful indicator of an operator's profitability.

*6 **Instant messaging**: A service that allows users to send messages to other users on a network.

*7 **Avatar**: A character used a an on-screen representation of a person.

Table 1 Examples of service enablers integrated on the DOCOMO Euro-Labs testbed

Enabler name	Description	Example use case
Personal handwriting	Synthesis of personal handwriting font from a few character samples	Creation of handwritten messages from standard text (e.g. e-mail) to add a personal touch
Face & object detection	Detection of faces in images and video	Tracking of individuals in collaboratively collected videos, automatic tagging of video content
Video rendering	Rendering of single video stream from multiple sources, generating text, image and music video snippets	Conversion of standard text messages into an animated personal video messages
Semantic text analysis & mood detection	Classification of text into semantic categories based on its content, detection of writer's mood	Classification of text messages to trigger complementary actions, e.g. playing music to match the message content
Personal voice	Synthesis of personal voice from a few voice samples	Creation of personally spoken messages from standard text, e.g. e-mail, to add a personal touch
Voice search	Voice search based on voice samples	Retrieval of similar sounding clips from large multimedia collections, e.g. to compile similar audio clips associated with a single event
Video editing	Automatic non-linear video editing (movie creation) from individual video clips	Automatic creation of a single movie from a collection of video clips and photos taken at a specific event, e.g. a birthday party or a vacation
Video sync (time)	Synchronization of multiple video clips (purely based on video content; no clocking details required)	Composition of collaboratively collected videos into a single viewing experience
Video sync (place)	Synchronization of multiple video clips by using sensor and positioning data	Composition of collaboratively collected videos into a single viewing experience with 3D information such a height and direction
Audio sync	Synchronization of multiple multimedia clips (purely based on audio content)	Composition of collaboratively collected multimedia clips into a single viewing/listening experience
3D face animation	Generation of a full 3D model of a human face from one single portrait image	Creation of a life-like animated avatar of a user from a single camera phone snapshot
Digital make-up	Use of 3D face model for subtle enhancements of a portrait photo, e.g. adding a smile	Creating a more impressive profile image that can be modified according to circumstances
Background subtraction	Subtraction of background detail in images and videos	Virtual placement of video chat participants against a sophisticated or privacy-safe backdrop
Video summarization	Content-based summarization of video clips	Automatic editing of video content into shorter clips, e.g. showing only the highlights of a lengthy event
Fast stream switching	Seamless and ultra-fast client side switching between several video streams without the need for pre-buffering	Instant switching between streams showing multiple perspectives of the same event, or between mobile TV channels

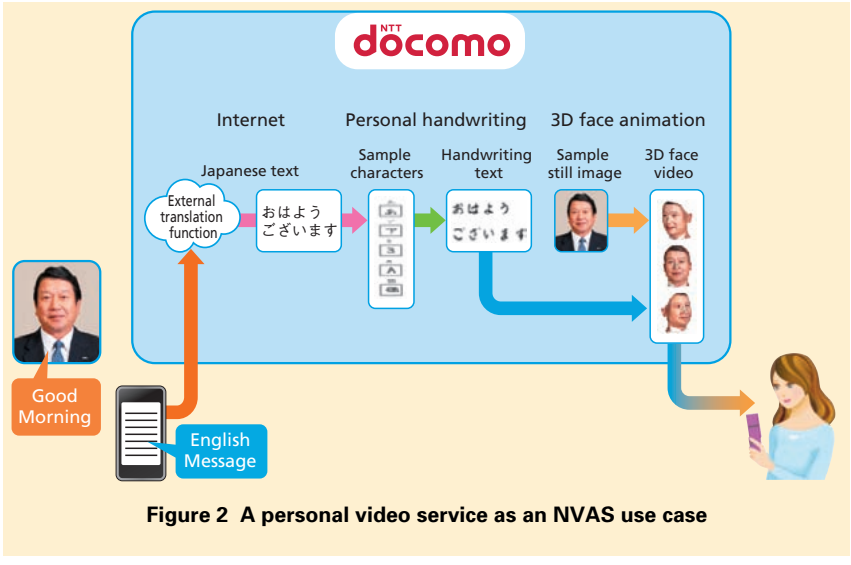

Figure 2 A personal video service as an NVAS use case

We developed a prototype application to demonstrate the use of this service. This application allows the user to individually choose service enablers to enrich the content of messages. A screenshot of the personal video message application is shown in **Figure 3**. At the lower right of each client's window, it is possible to select service enablers. Simple text messages are exchanged via the message board. If "Handwriting" is selected, message text is displayed in each user's personal font. If "Translation" is selected, mes-

Figure 3 Screenshot of the personal video messaging application

sages are translated within the network according to the languages used by each client. When "Face Animation" is used, the facial expressions of the person in the picture are changed according to the content of the message.

Before using the handwriting enabler, the customer needs to register his or her personal font by filling out at least five sample characters on a sheet of paper. A photograph of the character sheet is then taken with a mobile terminal or tablet computer, and the image is uploaded to a server that generates the customer's handwriting font. For the generation of 3D face models, the approach is quite similar. A single 2D portrait photo is taken and uploaded to a server, and the 3D-model is automatically generated. The 3D data is used to make the avatar look more realistic and natural by animating the mouth movements and changes of facial expression. In our prototype implementation, facial expressions are controlled by emoticons[*8] (e.g. happy, sad) parsed from the text content of the chat message.

By adding more service enablers it would be possible to enhance the existing service even further. For example, by adding a text-to-speech engine and a lip synchronization enabler, the personal avatar could easily be made to speak out the message in another language.

The Personal Video Message service is an example of how several service enabler technologies can be dynamically combined in the network of the service provider. It illustrates how the SEN lead to new service concepts and faster application development. Processor-intensive parts, like image processing and 3D model generation, are assumed to be executed within the network of the operator. Hence, the customer can enjoy exclusive services without having to worry about underlying technologies or drainage of battery power from the mobile terminal.

4. Smart Media Processing

At DOCOMO Euro-Labs, we are currently looking at further expansion towards smart media processing with a particular focus on multimedia technology. Specifically, in the multimedia field, we are interested in the growing popularity of user-generated content captured through mobile terminals, and on the increasing availability of more powerful and better equipped devices such as smartphones [3]. For instance, we are looking at expanding into new fields of technology such as multimedia and computer vision[*9] that originally grew from fields outside telecommunications (e.g. robotics, machine perception or video editing). Although additional research is needed to clarify the new issues and conditions that must be considered in applying these technologies to our telecommunication service scenarios, it is possible that NVAS will still be able to operate in these scenarios.

For example, our experience in the CoopMedia joint research project [4] indicates that telecommunication-specific use cases can—with the same specific requirements and technical con-

*8 **Emoticon**: A picture of a face made by combining text characters.

*9 **Computer vision**: Using computers for the acquisition, recognition, identification and processing of real-world video information.

straints—lead to completely new solutions outside the telecommunications field. Here, we have developed a new approach for video synchronization [5]. This synchronization enabler is uniquely based on video content without the need for additional clocking information. This means it can not only be used to synchronize several video streams at the same time, but can also be used sequentially so that new content uploaded to a portal can be resynchronized to the existing content that has already been uploaded. Besides its direct use this enabler is also a fundamental component for related multi-perspective video editing and digesting enablers. Driven by this work on video technology, we have extended our multimedia research into audio-based synchronization, video editing/digest creation, visual perception and visual localization services. Following the NVAS concept, these novel enablers can easily be combined to create value-added services in a highly dynamic fashion.

5. Conclusion

In this article, we have introduced a prototype system that integrates our infrastructure for rich communication with a personal video messaging service[*10] as a specific use case of NVAS. In the future, to complement the current application fields, we plan to conduct research aimed at context-aware and ambient services[*11] by building on more advanced multimedia technologies such as augmented reality. DOCOMO Euro-Labs will continue to build and support a sustainable service enabler network as well as innovative mobile applications in these fields, and will press ahead with research as a pioneer of new services while adapting flexibly to market trends. In the next issue, we will describe the Optical Mobile Network (OMN), which is one of the constituent elements of the Next-generation Mobile Communication Network (NMN).

References

[1] H. Berndt, et al.: "Requirements and Architecture for Mobile Networking—Pursuing Smart Innovation: HEART," NTT DOCOMO Technical Journal, Vol.13, No.3, pp.30-36, Dec. 2011.

[2] GSMA RCS Working Group: "Rich Communication Suite (RCS) & Rich Communications Ecosystem (RCE)." http://www.gsmworld.com/our-work/mobile-lifestyle/rcs/

[3] S. Hartwig, M. Luck, J. Aaltonen, R. Serafat and W. Theimer: "Mobile multimedia—challenges and opportunities," IEEE Trans. on Consumer Electronics, Vol.46(4), pp. 1167-1178, Nov. 2000.

[4] F. Schweiger, E. Steinbach, M. Fahrmair and W. Kellerer: "CAMP: A Framework for Cooperation Among Mobile Prosumers," In Proc. of the IEEE ICME Workshop on Community driven Mobile Multimedia, Jun. 2009.

[5] G. Schroth, F. Schweiger, M. Eichhorn, E. Steinbach, M. Fahrmair and W. Kellerer: "Video Synchronization Using Bit Rate Profiles," In Proc. of the ICIP '10, Sep. 2010.

*10 **Personal video messaging service**: A service exhibited by NTT DOCOMO at the Mobile World Congress 2011. This is one example of collaboration between DOCOMO Euro-Labs and the Network Development Department.

*11 **Ambient services**: Services that are absorbed into the user's surroundings to provide the user with information and other resources as and when necessary.