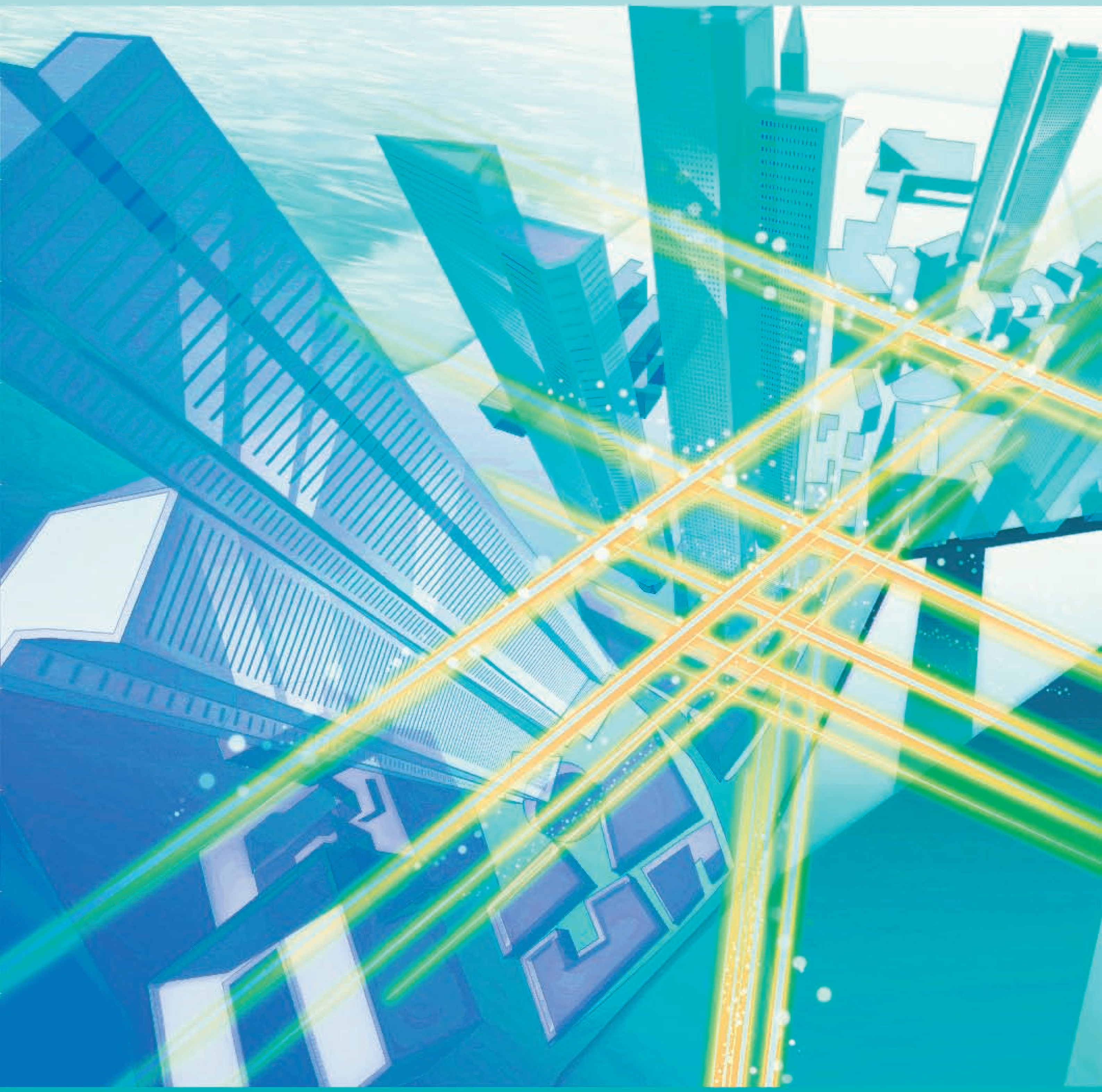
NTT Technical Review

5
2014



May 2014 Vol. 12 No. 5

NTT Technical Review

May 2014 Vol. 12 No. 5

Feature Article: NTT R&D Forum Workshop

Feature Articles: Video Technology for 4K/8K Services with Ultrahigh Sense of Presence

Regular Articles

Global Standardization Activities

New NTT Colleagues

Papers Published in Technical Journals and Conference Proceedings

NTT's Open Innovation in Silicon Valley

Srini Koushik
President and CEO,
NTT Innovation Institute, Inc.

Mayan Mathen
SVP and Chief Technology Officer,
NTT Innovation Institute, Inc.

Abstract

This article introduces the activities of NTT Innovation Institute, Inc. (NTT I^3), which recently marked the first anniversary of its establishment in Silicon Valley, California. This article is based on a speech given by Srini Koushik, President and CEO of NTT I^3 and Mayan Mathen, SVP and Chief Technology Officer of NTT I^3, at NTT Forum on February 13, 2014.

Keywords: open innovation, cloud, security

Speech given by Srini Koushik

1. Overview

My objective is to introduce NTT I^3 by explaining who we are and what we do and then to briefly discuss the nature of innovation. The whole theme of this conference is "co-innovation," and there has also been discussion on the concept of "open innovation," so it is crucial to cover precisely what innovation means and how we are trying to enable it within Silicon Valley.

2. Characteristics of NTT I^3

We previously heard about the four steps of open innovation: speed, working together with NTT Group operating companies, remaining mindful of the ecosystems of our partners, and working with start-ups and other companies that have the skills and technologies we need to bring our products to market faster. NTT I^3's mission includes these four steps. Although a lot of attention is now being focused on B2C (business-to-consumer) and B2B (business-to-business), we are concentrating on the enterprise customers. Specifically, we are building solutions that the operating companies can take to market within the next 12 to 18 months. Currently, the opportunities in terms of the cloud and security are huge, and we are well positioned within the U.S. marketplace as a global player to take advantage of that. Everything that we do, we are working on quickly. In fact, we released one of our first products to market in March of this year, and considering that we just launched NTT I^3 in June of last year, this represents an incredibly rapid way of bringing products to market. Our colleagues have talked about the importance of balancing engineering that includes many features

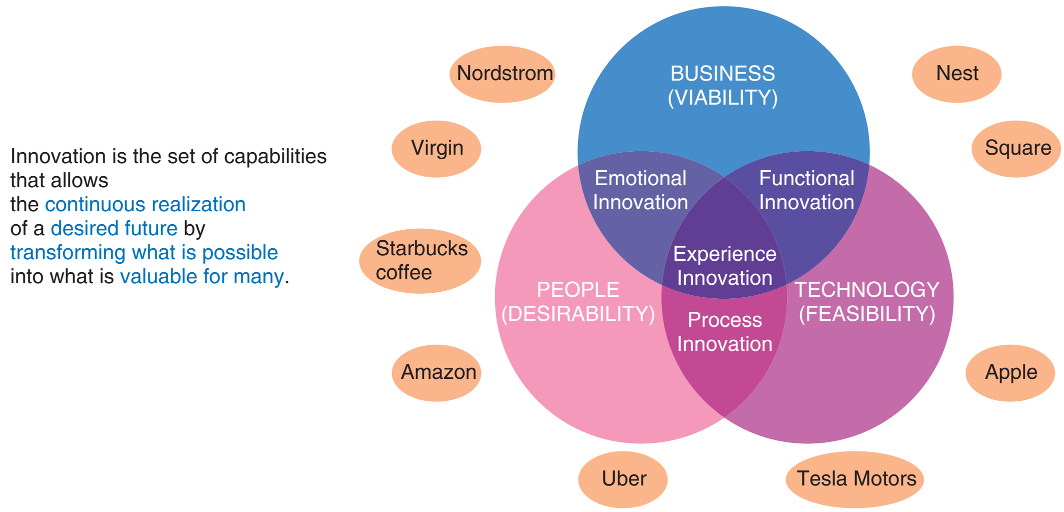


Fig. 1. Innovation requires human-centered design.

with speed-to-market, and of course, this is what we are striving for on a daily basis. But for me the key idea that we are focusing on is innovation, especially open innovation.

3. What is innovation?

So, what exactly is innovation? Within the framework of human-centered design, innovation and how to create innovation is something that a lot of people are spending time on (**Fig. 1**). As a research and development (R&D) community, it is crucial for us to keep in mind that humans are using our solutions. All too often the focus is mainly on the technology/feasibility side of the design process. For example, we look at a new technology and are very excited about it and want to build something with it. However, true innovation comes from being able to balance technology—which encompasses the feasibility, the "Can I build it?" side of things—with a financial model, i.e., "If I build it, will anyone buy it?" And, to add a third element that is equally important, as companies such as Apple have so eloquently shown us: It's all well and good if you build it and people are prepared to buy it, but people should want it. They should want to buy from us.

How many people know what a "Microsoft Zune" is? Maybe a few people. The Microsoft Zune was the competitor to the iPod that Microsoft brought out. It was very well designed, utilized very good technology, and was something that people would pay money for. However, people did not want it, because quite simply, the design was not easy to use. That was its downfall. What we are striving to do in NTT I^3, with all of our products, is to hit that middle ground, to come up with ideas for products that we can build, that we can take to market through our operating companies, and that can be differentiated from other products in the marketplace. So when we go talk to our enterprise customers, they can see the difference that we're bringing to the table. This high level of perfection and excellence is the core concept of our idea of innovation.

4. The seven components of innovation

I consider innovation to be made up of seven main building blocks. Here, it might be helpful to think about children and the way they interact with the world. One of the things that you will notice about children is that they are the most innovative people you can find. When kids are young, e.g., five or six years old, and you give them a stick, within five minutes they will have found about ten uses for the stick. They might use it as a musical instrument or as a tool; they might bend it and use it for something else. They typically have no problem coming up with at least ten different uses for that stick. Why is that? If you think about it, children are experimenting all the time. If you give them something, they're always experimenting with it, trying to see how it works, what it can do. This is the first building block of innovation: experimentation.

Second: children are not afraid of failing. If they try something and it doesn't work, they just get up and try something else. They keep working on something until they like it, until it does what they want. Therefore, the acceptance of failure is very important.

A third component—one that Google and others talk about, and which is critical—is having sufficient time to think. For example, if we tell our researchers to be innovative, it is imperative that we give them time to mull over their ideas. It can be quite difficult in our daily lives, as we tend to our daily tasks, to be creative and come up with different ways of solving a problem. In my own case, there are many times when I have been working on a very tough problem, and having been unable to find a solution, I have finally given up and gone home to bed. Then the next morning when I wake up and get in the shower, I suddenly have three solutions in my head, three innovative ways of solving that problem. I've had sufficient time to think about the problem, to let it percolate, and to move around and change locations to a completely different environment, and as so often happens in this type of situation, that is when I'm finally able to come up with a new idea. This third component, time to think, is therefore crucial.

The fourth component of innovation is creativity and playfulness—again, traits that are abundantly evident in children. That's what actually makes children creative, being playful. When you look at campuses such as the ones that Google has built, or that Facebook is trying to build, and others, the reason they have all those very bright colors and fun ways to interact is because they're trying to stimulate creativity and playfulness. When you're playing with something, it's much easier to think of new ideas. When you're under stress, the opposite is true, and it becomes harder. When you're playing, when you're having fun with others, that is when your creativity really starts to improve.

The fifth component of innovation is having effective leaders. Innovative organizations have leaders who inspire. They don't direct or order people to do things. You can't order someone to become innovative. You have to allow them to do their job and help them when they need obstacles removed. That's critical.

The sixth component relates to diversity of thought. Lateral thinking serves as a good example of this in that many of the most innovative ideas come from taking something that worked in one domain and applying it somewhere else where nobody thought it was applicable. This is why we've seen success here at NTT I^3 by combining teams, putting together labs that mix different domains, which is an excellent way to stimulate ideas and come up with innovative new plans.

All of these components are important, but perhaps this last one is the most crucial: willingness to change. If you have the other six components in place but you're not willing to change, you might come up with two or three good ideas, but then it'll all fizzle out.

I described these seven components in detail because it's not enough to simply talk about innovation. We also have to understand it. We have to be able to build innovative products. As the CEO of NTT I^3, it is imperative that I set up an environment like this, because otherwise, the team as a whole will be unable to come up with many ideas. We spend a lot of time at NTT I^3 trying to build this type of culture, this type of idea. We like our employees to be creative and playful, to joke around and to play. We encourage that, and a lot of other executives are starting to do the same thing. As part of our team, we have senior executives who aren't afraid to joke around and dress up in a Halloween costume for an office party. We have 20 engineers from R&D working in Silicon Valley, and when they see a senior executive do something like that, they know it is OK for them as well. The result is they are more fully engaged and better able to help us innovate. This is one of the primary reasons that, although it is not technical, I wanted to clarify my views on innovation. Innovation is a passionate subject for me, and it's critical that we build an organization that supports it.

5. Research focus of NTT I^3

Here are the five major areas on which we're focusing our research (**Fig. 2**).

First, everything that we're doing is centered on security, because whether we're talking about wearable computers, big data, or the cloud, security is at the core of everything that we do.

The nature of security threats is changing. It has gone from being rather general to something very specific. Now we have attackers who will target a specific company to damage its reputation, or we have countries that are actually sponsoring people to come in and attack. The nature of attacks is becoming different, and so we're spending a lot of time working on security.

A second area on which we are focusing is the "asset-light generation." This term, which has recently become more and more common in the U.S., refers to the next generation of consumers currently out

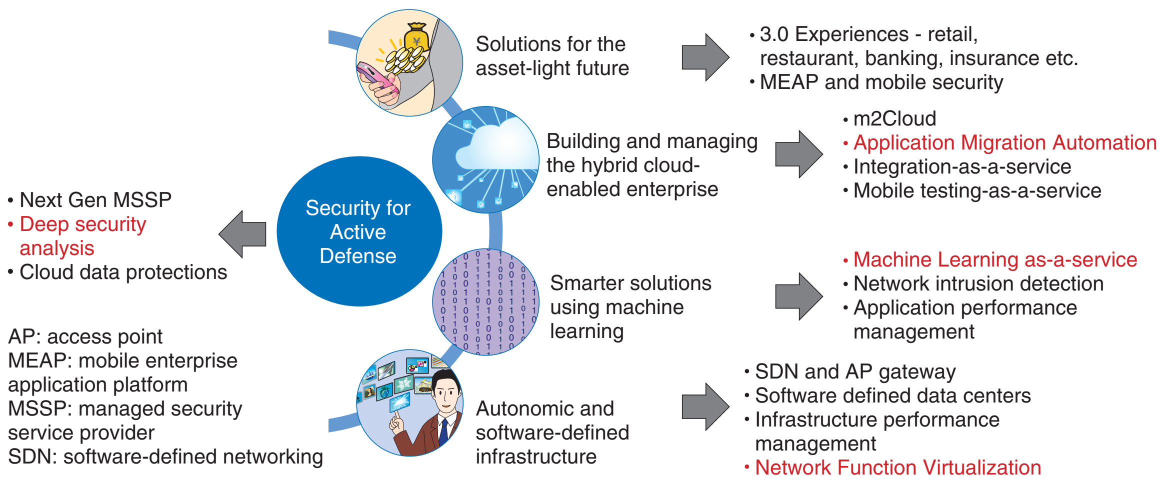


Fig. 2. At NTT I^3 our R&D focuses on cloud-enabled enterprises.

there in the marketplace. This generation of consumers does not buy things, as previous generations did. When I was young, I bought record players; I bought CDs; I bought DVDs and the like. People don't do that today; rather, they stream things. The CDs and DVDs of the past have been disrupted by iTunes, and iTunes has been disrupted by SoundCloud and Pandora and others, and nobody buys things anymore, not even a 99 cent song, because why buy something when you can stream it, on demand? So these things we used to own, these assets, are becoming rare as more and more people join this asset-light generation. When this type of generation forms our client base, it changes how we view problems and how we build solutions. We are intently focused on this generation and coming up with ways of addressing the unique needs it presents.

Third, everyone these days is talking about the cloud, and how we need to help enterprises get to the cloud. From my point of view, in terms of our customers, it's important that somebody be there to help them when they get there. The internal workings of companies are becoming more and more complex as an incredible number of different systems are integrated and then are constantly and sometimes erratically updated. Let's say you're in an IT (information technology) environment where your mainframe system is changing once a month, and your sales force system is changing once or twice a week. It becomes extremely difficult to maintain a certain way of testing code and to make sure that you can do and undo connectivity. In response to this, at NTT I^3, we're working on solutions involving application migration automation and how to enable DevOps (development and operations) in the cloud. This is one of our first products that will be coming out next month.

A fourth area on which we focus is machine learning, which is very important to us and which we are using in several different areas. For example, we took Jubatus and deployed it on our cloud offering, put it on Dimension Data's cloud, and built a service on top of it that can solve specific problems without the help of data scientists. This is just one example, but our idea is to see if we can solve 70–80% of our problems in this way, package the process, and then take it to our mid-sized customers and others. We believe it is possible to do this by working closely with start-ups such as BigML, SnapLogic, and others so that we can bring services to the marketplace very quickly.

Finally, because we are an infrastructure company, a networking company, as such, our very core is network function virtualization, so we are focused on everything starting from the AP (access point) gateway and SDN (software-defined networking) gateway all the way to the network function gateway and network function virtualization that we've got set up. We are eager to start releasing products in this sphere in the coming weeks and months.

The five areas above are our primary focus. This enables us to actually talk to customers and demonstrate to them that we are looking at the end-to-end spectrum of how to manage and build cloud applications.

Speech given by Mayan Mathen

1. Innovation and collaboration at NTT I^3

There has been a lot of discussion at this conference about co-innovation and collaboration. In my opinion, collaboration and innovation go together very closely. One of the main areas that we focus on at our Silicon Valley branch is working with the NTT operating companies. Without them, whatever we're doing wouldn't be as effective, and we wouldn't really have a meaningful mission.

One example is the case of the NTT operating companies initiating a plan to work together in the security domain. We now have NTT R&D, Dimension Data, NTT Com Security, and NTT DATA all coming together on a regular basis hosted by NTT I^3 to work together towards a common mission in the security domain.

In the very short term, we're looking at a standard architecture for security as a group and go-to-market technology that has successfully positioned NTT Group both globally and as a leader in the Gartner Magic Quadrant for security. This is something of which we should be quite proud. Moreover, as of recently, from our visit to Tokyo, we have started launching the next wave of collaboration across the NTT Group operating companies with a focus on the cloud infrastructure domain. In April, we'll be starting the same type of collaboration in another domain, and we'll continue through 2014.

2. Working with start-ups and academia

We're working with various start-ups, which is a lot of fun because these people have bright ideas. They don't think in a box; they think out of the box, and they seize every chance to shake things up. We both see it as a great opportunity to join forces and collaborate using various assets that we both possess. As an example of application migration to the cloud, we have worked with a great company called CliQr that helped us identify ways we can help clients move onto Dimension Data's and NTT Communication's cloud business units. There are many others, too.

We have also been working with academia and standards bodies. Many outstanding individuals have been dispatched from Tokyo to NTT I^3, and everywhere they go—meetings, conferences, etc.—they are held in extremely high regard. Everyone looks at them and thinks, "Oh, the NTT team is here." So we should be very proud of their achievements in that regard.

3. Developments at NTT I^3

Now, I'd like to provide you with two examples of what we're working on.

In Silicon Valley, it's necessary, as pointed out earlier, to work with others and at great speed. A friend of mine runs Plantronics, and they have a division called PLT Labs. My friend and I were having coffee and talking about an idea, one that we are both very passionate about—wearable technology—and we started discussing Google Glass, particularly its disadvantages: it's clumsy, it's got certain challenges, it sits on your face, so if you have prescription glasses, you can't use it effectively, etc. We both looked at each other and said, "Let's try to come up with something more interesting."

Specifically, we took a very common, readily available piece of technology (a little Bluetooth headset) and had the team at PLT Labs expose a few APIs (application programming interfaces). Within just a few hours, two engineers had written a little application that enables users of this headset to move their heads without wearing any obtrusive pieces of technology, and to run an application on their mobile device just with the motion of their heads. Users can tap the device and move it around to activate built-in sensors that detect the temperature. This gives users an entire platform with which to create countless applications with different uses. The best part is that it's inexpensive and can be used to do things very quickly. For example, if I wear the device while accessing Google Street View, I simply move my head around to move around the street. Now, just think about the applications of doing something like this. Think of the enterprise as a platform, as a global telecommunications provider and as a platform service initiative. Imagine if we allow enterprises to program onto this. Imagine buying inexpensive wearable technology and then exposing information and creating mash-ups everywhere. For example, if we could mash up information using cloud APIs. Picture somebody sitting in an office with wearable technology and taking a context-centered platform to the next realm, being able to offer services to a potential customer. There's an amazing range of applications in health care, in financial services, and more. This technology offers the cutting-edge ability to zoom around Street View and identify a potential customer.

The second example I want to share with you goes back to our collaboration with start-up companies. You may remember the movie "Minority Report" with Tom Cruise some years back. The way it showed them moving files around was very sci-fi. Recently, a Silicon Valley venture company gave the man who came up with the creative concept behind that movie money to actually fund his dream. So he formed Oblong Industries, which really excited us. Their headquarters in Palo Alto is very close to where our new office is located, so we went over to have a look, and the technology is extremely impressive. Moreover, what interested me and my team much more was the fact that it gave us ideas for an integration of the next generation of collaboration that is possible by leveraging our global networks, our telepresence network capability, and our audio and video conferencing capabilities. We are also excited by the fact that, because we are everywhere, because we are one of the largest carriers in the world, we can actually enable this sort of integration.

Our next step was to start discussing what we could do with this idea. In San Francisco, I met with a so called "cyborg anthropologist"—a person who looks at how technology and humans interact and how they evolve. He is now going to work with us to implement this Oblong technology in client experience data and to host client innovation workshops. Our enterprise clients will be able to sit in a room with us, experience the visualization exercise we offer, and will almost certainly view us as a remarkably innovative company.

Every day at NTT I^3, I get to experience this incredible technology, and it inspires me when our clients come in, see what it is that we do, and then leave with the confidence that they are working with the best and most innovative company in the world.

Conclusion

We have provided an overview of some of the technologies we are developing. These days, people are increasingly using phones and other mobile devices to access information, and our technology is unobtrusive, enabling users to view content on iPads, phones, etc., in a manner in which they do not have to physically have the devices in front of them as they are walking.

This type of technology is evolving at an incredible rate, and one of the nice things about being in Silicon Valley and working under the NTT umbrella is that we are ideally located to identify new trends, to work with up-and-coming designers and developers, and to build and help grow new technology. We can come up with a new idea and have a prototype built two weeks later. The possibilities for future applications are limitless.

Our strategy is to build a very quick prototype, and then see if it works, and if it works, to try to engage the NTT operating companies and say, "Can we actually make money on this?" If we can make money on it, we build it, and we build it quickly. In this manner we can deliver a minimum viable product, or MVP, that we can take to market and sell to customers. By working closely with our clients, we can improve the product and start adding more and more features as we move into the future.

Speakers' profile

Srini Koushik

He joined IBM in 1994 and became head of e-business development in 2000. He joined Nationwide Insurance in 2001 and became its CTO (chief technology officer) in 2009. In 2011, he joined HP and took up the post of chief development officer spearheading the development of global applications. He founded Right Brain Systems, a consulting company, in 2012, serving as president and CEO. He took up his present position at NTT I^3 in May 2013.

Mayan Mathen

He joined a technology integrator company in 1999 and then Dimension Data in 2004. He joined Converged Communications in 2004 and became its National Practice Manager. Prior to his current role as CTO at NTT I^3, he served as the Group CTO for the Africa and Middle East operations of Dimension Data. He is an accomplished speaker and widely recognized global technology leader who often presents at major conferences around the world.

Research and Development Policies for Ultra-high-presence Video Technology toward 4K/8K Services

Yoshiko Sugaya, Hiroshi Fujii, Atsushi Sato, Hiroaki Matsuda, Shinichiro Chaki, and Hirohito Inagaki

Abstract

Efforts toward achieving 4K/8K broadcast services have been ramping up in countries throughout the world, and in Japan, broadcasters, telecommunications carriers, and equipment manufacturers have teamed up to establish the Next Generation Television & Broadcasting Promotion Forum (NexTV-F) to make 4K/8K communications and broadcasting a reality. As a founding company of NexTV-F, NTT is becoming a world pioneer in the development and promotion of 4K/8K services through a variety of research projects and technical developments. This article introduces these technical developments and the ultra-high-presence services targeted by NTT.

Keywords: 4K/8K, ultra-high-presence services, ultra-high-definition video

1. Introduction

The transition to digital terrestrial broadcasting was completed in Japan in the summer of 2012, and as 2013 got underway, the term 4K/8K suddenly became popular as a reflection of the desire to achieve a level of resolution and picture quality far beyond that of high-definition (HD) images currently in use in terrestrial television (TV) broadcasts. In addition, business possibilities using 4K/8K ultra-high-definition video began to expand amid expectations that Tokyo's successful bid for the 2020 Summer Olympics would increase the demand for high-presence video and accelerate the spread of 4K TV to homes. The NTT laboratories have long been active in the research and development (R&D) of technologies for 4K/8K communications and broadcasting, and today, their efforts toward providing full-scale 4K/8K services are gathering momentum.

In July 2012, NTT Network Innovation Laboratories conducted a joint experiment with NHK on the delivery of Super Hi-Vision 8K video of major sporting events for public viewing [1]. This video was transmitted between London and Japan over a shared Internet protocol (IP) network making use of high-speed error correction technology based on low-density generator matrix (LDGM) code. Meanwhile, NTT Media Intelligence Laboratories, as a long-time promoter of H.265/MPEG-H, also known as HEVC (High Efficiency Video Coding) as an international standard, announced in July 2013 the development of an HEVC-compliant software encoding engine with the world's best video compression performance. Then, in August of the same year, NTT Advanced Technology began domestic and overseas sales of the world's first HEVC software codec development kit, called HEVC-1000-SDK [2]. Studies and trials of 4K video delivery have also been quite active within the NTT Group, which seeks to get an early jump on the commercialization of 4K/8K and the development of peripheral business ventures. For example, NTT WEST held the world's first 4K video Internet-delivery trial to a set-top box (STB) at IMC (Interop Media Convergence) TOKYO 2013 in June 2013. In this trial, HEVC-compressed 4K video content was streamed to an NTT WEST STB over the Internet, thereby demonstrating the feasibility of 4K TV for home use [3]. The delivery of 4K video was also demonstrated at NTT GROUP COLLECTION 2013 held in October 2013.

In addition to the above, NTT Plala has announced plans to hold a 4K trial with the HIKARI TV video service in the fourth quarter of fiscal year (FY) 2013 and to conduct a trial delivery of a 4K video-on-demand (VOD) service in the first quarter of FY2014 ahead of actual IP broadcasts. The company is also collaborating with the Shitamachi Bobsleigh Network Project Promotion Committee on the production of 4K-video content for this project.

2. Overseas trends in 4K/8K broadcasting

Efforts toward achieving 4K/8K broadcasting are not limited to Japan; there has been quite a lot of activity throughout the world since 2012. In France, the 4EVER project got underway in 2012 as a three-year project to advance R&D in the field of HEVC and ultra-high-definition TV (UHDTV). As part of this project, relay broadcasts of the French Open tennis tournament in 4K UHD format were conducted for public viewing using 4K displays in June 2013. In Korea too, 4K trial broadcasts were held in October 2012 by five leading cable broadcasters.

3. Trends toward 4K/8K broadcasting at NexTV-F

In Japan, the Next Generation Television & Broadcasting Promotion Forum (NexTV-F) was founded by NTT, NHK, Sky Perfect JSAT, and Sony in June 2013 with the aim of achieving early deployment of next-generation broadcast services involving 4K/8K broadcasting and smart TVs [4]. A total of 21 Japanese companies consisting of broadcasters, telecommunications carriers, equipment manufacturers, and advertising agencies made up NexTV-F at the time of its founding. In more detail, the objective of NexTV-F is to upgrade broadcast services toward 4K/8K broadcasting through a variety of activities ranging from the study of transmit/receive provisions and specifications and the preparation of broadcast systems, to content production, encoder development, testing, and holding of trial broadcasts. Plans are being made to broadcast various types of sports events as milestones using as a guide the roadmap prepared at the end of May 2013 by the Study Group on Upgrading of Broadcasting Services at the Ministry of Internal Affairs and Communications. For example, NexTV-F is planning to construct an environment for holding trial broadcasts in 2014, the year of the Brazilian World Cup, so that viewers interested in next-generation broadcasting can experience 4K broadcasts. Similarly, in 2016, the year of the Rio de Janeiro Summer Olympics, NexTV-F is planning to construct an environment for holding 8K trial broadcasts in addition to launching 4K broadcast services, and in 2020, the year of the Tokyo Summer Olympics, it is planning to launch 8K broadcast services. As one of the founding companies of NexTV-F, NTT has been promoting studies on specifications for real-time compression encoding in 4K broadcasts and collaborating with Japanese manufacturers in developing hardware-encoding large-scale integrated circuits (LSIs) compliant with HEVC next-generation encoding technology.

4. R&D policies toward ultra-high-presence services

NTT has a future vision of ultra-high-presence services and therefore aims to create an enriching and rewarding user environment by providing a user experience that combines high-definition video, high-presence audio, and diverse human sensations. Various elements will be needed to achieve this, including HD video, free viewpoint video, high-presence audio (ambient sound, free listening point), and communication of the five human senses. Among these, the NTT Group sees the development of services using HD video as the starting point in providing ultra-high-presence services and becoming the world leader in this field (**Fig. 1**).

With a view to 2020 and beyond, NTT laboratories are researching and developing technologies toward the creation of HD-video business in the three areas described below (**Fig. 2**).

(1) Smart TV services combining communications and broadcasting

The market surrounding high-definition video linking 4K broadcasting and Smart TV is expected to expand rapidly both inside and outside Japan. The aim here is to provide HD video services in a lifestyle-friendly format for a wide range of people from mass users to enterprise users.

Our first step will be to simultaneously achieve 4K

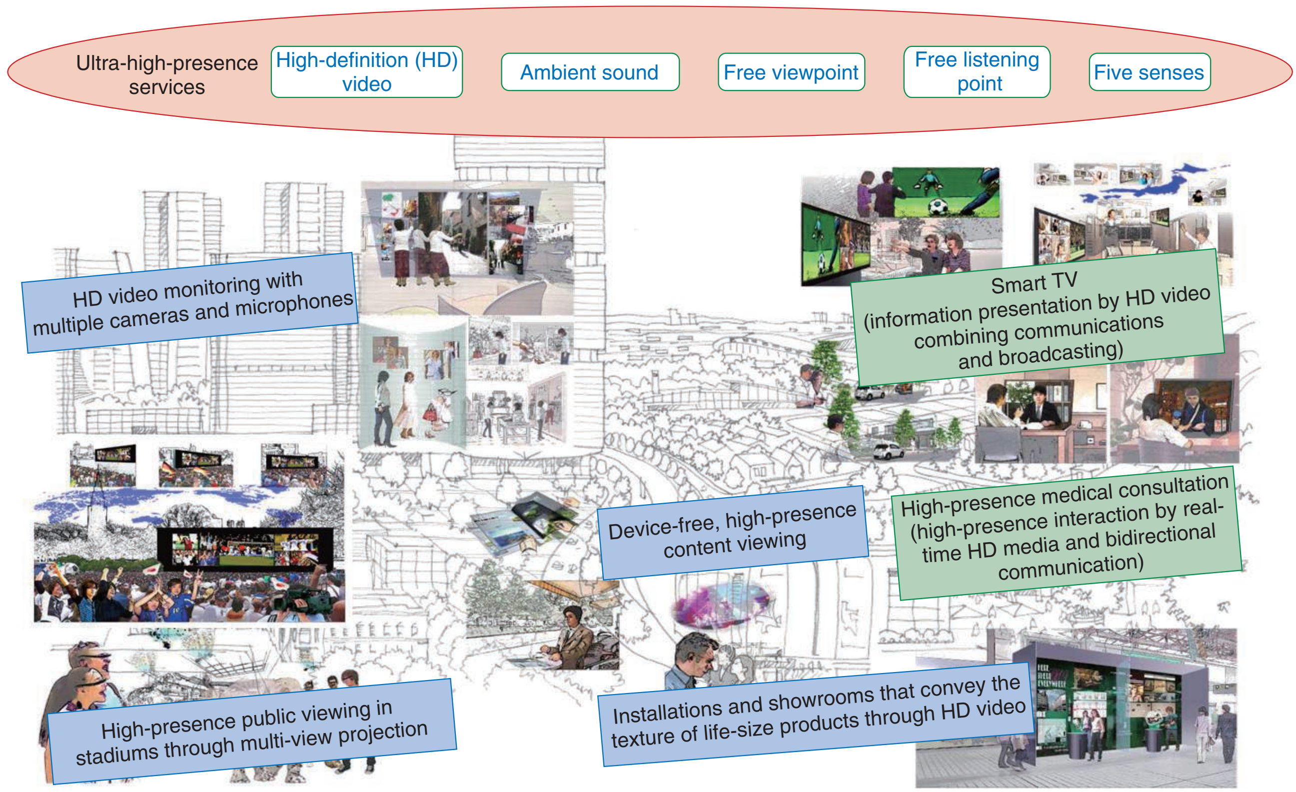


Fig. 1. Future vision of ultra-high-presence services.

delivery and 4K retransmission services over optical fiber. This will be achieved by holding technical trials led by the NTT Group in accordance with the 4K/8K broadcasting roadmap established by NexTV-F as well as by developing the world's first HEVC hardware encoder and developing video transmission technologies. Furthermore, we will conduct studies on the application of 4K technologies to smart TV and the creation of HD smart TV content. In Japan, the smart TV market is expected to reach about 3 million units (2,080,000 TVs and 860,000 STBs)—or about 5% of all TVs in Japan—by 2017 [5], while in the 4K TV world market, the number of units shipped is expected to increase dramatically from 500,000 units in 2013 to 7,250,000 units in 2016 [7]. In such a business environment, we can expect the expansion of high-level, compelling services such as FLET'S TV and HIKARI TV based on 4K HD video to go hand in hand with getting users to enjoy the great advantages that optical circuits can offer.

At the same time, the market for live video delivery services as an enterprise application in Japan is predicted to grow rapidly from 640 million yen in FY2010 to 744.2 billion yen in FY2020 [7], and with the coming of the Tokyo Olympics, the expected growth in live video delivery services for public viewing and other mass-user applications is expected to stimulate the market in Japan for HD video.

(2) Video content transmission services

4K/8K broadcasting must be made to be economically feasible by compressing the bandwidth of video content. In this regard, we are developing an HEVC hardware encoder for achieving low-delay, high-quality video transmission [8]. Using this encoder in video transmission equipment developed by the NTT Group will make it possible to provide HD video transmission services for transmitting and receiving large volumes of HD video data in real time. In this way, we aim to increase the number of network users making use of video relays and other facilities and thereby expand the market.

(3) HD video usage services

Needless to say, we can expect to see various applications using 4K/8K HD video in the medical-care,

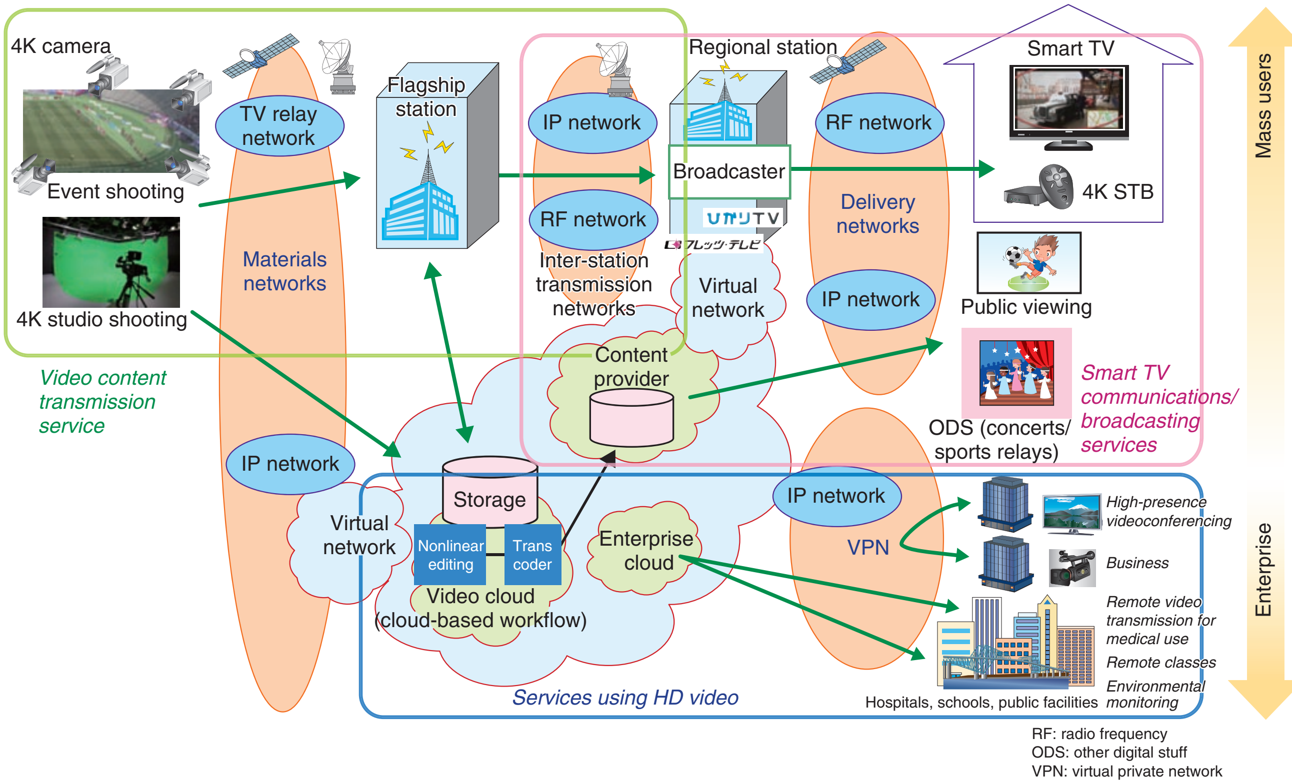


Fig. 2. Concept of HD video services.

education, and content-production industries, and the expansion of the HD solutions business can be anticipated. In particular, it is our desire at NTT to create novel services using HD video such as video transmission for remote medical care, remote classes, and efficient, cloud-based video production. In this way, we hope to improve medical care as well as the social environment and welfare of the people through the use of high-presence video.

5. Issues and technologies surrounding the provision of 4K/8K services

Achieving a low transmission bandwidth by reducing network load and shortening the delay in real-time transmission are two issues that must be addressed if a variety of 4K/8K services are to be provided. There is also a need for technologies that can facilitate the provision of high-level, compelling services that include the transmission of HD video. NTT laboratories are making a concerted effort to resolve these issues by conducting R&D in a variety of fields (**Fig. 3**).

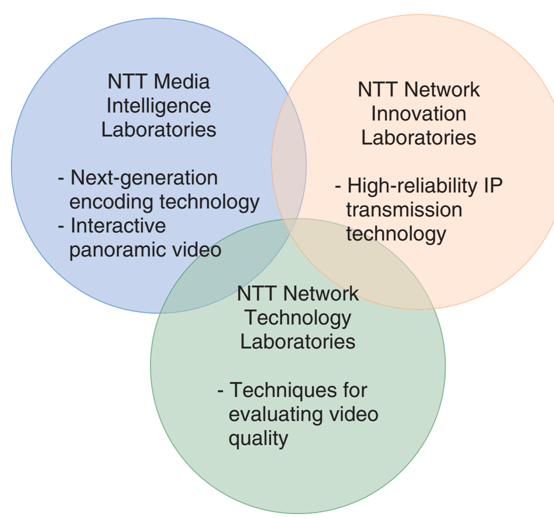


Fig. 3. R&D fields at NTT laboratories toward the provision of 4K/8K services.

5.1 Activities at NTT Media Intelligence Laboratories

NTT Media Intelligence Laboratories is involved in a wide range of 4K/8K activities. These include the standardization of HEVC next-generation encoding technology as a key component in lowering transmission bandwidth by reducing network load, the development of a software encoding engine and encoding LSI, and R&D of high-presence-video viewing technology. The HEVC standard achieves a compression rate that is generally twice that of the existing H.264/MPEG-4 AVC (Advanced Video Coding) (also called H.264) video compression standard. This means that it can keep the transmission bandwidth within 30 Mbit/s even for 4K video and can enable 4K-based communications on the existing NGN (Next-Generation Network).

To provide video delivery services such as 4K VOD (video on demand), it must be possible to set optimal compression parameters depending on the type of content being delivered. This can be achieved by applying a software encoding engine using HEVC [9].

However, high-presence-video transmission services for live video relays, public viewing, and other real-time applications require real-time encoding and transmission, which can be achieved through the use of real-time video transmission equipment using an HEVC LSI.

There will also be a need for innovative technologies such as interactive panoramic video so that high-presence-video services can help make ever-changing individual lifestyles even more enjoyable and interesting [10].

5.2 Activities at NTT Network Innovation Laboratories

NTT Network Innovation Laboratories is engaged in the R&D of high-reliability IP transmission technology to enable the short-term (temporary) use of global, multi-domain networks for transmitting not only 4K video but also large-capacity 8K video content. It aims, in particular, to achieve highly reliable, end-to-end transmission from the IP network (backbone/access circuits) to the reception of IP packets in the user's environment via low-cost network switches and receiving terminals. To this end, NTT Network Innovation Laboratories is participating in discussions at MPEG toward the development of the new MMT (MPEG Media Transport) standard while also proposing and promoting the standardization of LDGM code as a robust error correction technology in the application layer so that users need not rely solely on high-quality leased lines [11].

5.3 Activities at NTT Network Technology Laboratories

NTT Network Technology Laboratories is researching and developing techniques for evaluating video quality. These techniques are needed since the user's quality of experience in relation to video display size, viewing position, and compression rate is thought to differ from that of conventional HD video. Studies are being conducted with the aim of establishing reliable techniques for assessing the quality of video in 4K broadcasting and video transmission services [12].

6. Future activities

A variety of issues must still be addressed in the NTT Group ranging from video production to transmission and viewing given the launch of 4K broadcasting in 2016 and 8K broadcasting in 2020. Solving these issues will, of course, involve the development of key technologies at NTT laboratories, but it will also require R&D activities focused on open innovation, as well as coordination with the NTT business companies and the establishment of alliances with other companies.

If we look at the timeline for future plans, market trends, and business strategies with a view to 2020, we can say that now is the time for the technology born of R&D in 4K/8K high-presence video to break out of its shell, emerge in the form of services, and spread its wings on a global scale. The NTT laboratories will continue to research and develop advanced technologies in a mutually supportive role with the NTT Group as the latter works to become a world leader in the provision of ultra-high-presence services.

References

[1] Press release published by NTT on February 22, 2011 (in Japanese). http://www.ntt.co.jp/news2011/1102/110222a.html
[2] Press release published by NTT on August 8, 2013. http://www.ntt.co.jp/news2013/1308e/130808a.html
[3] Press release published by NTT WEST on June 11, 2013 (in Japanese). http://www.ntt-west.co.jp/news/1306/130611a.html
[4] Handout issued by NexTV-F on June 17, 2013. http://www.nextv-f.jp/en/pdf/press20130617_2.pdf
[5] Information material issued by Nomura Research Institute, Ltd. on November 30, 2012 (in Japanese). http://www.nri.com/jp/event/mediaforum/2012/pdf/forum184.pdf
[6] NPD DisplaySearch: "Quarterly TV Design and Features Report," 2013. http://www.displaysearch.com/cps/rde/xchg/displaysearch/hs.xsl/quarterly_tv_design_features_report.asp
[7] Research summary released by Yano Research Institute Ltd. http://www.yanoresearch.com/press/pdf/775.pdf
[8] T. Oonishi, T. Sano, K. Yokohari, K. So, M. Ikeda, A. Sagata, H. Iwasaki, and A. Shimizu, "HEVC Hardware Encoder Technology," NTT Technical Review, Vol. 12, No. 5, 2014.

https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201405fa3.html
[9] K. Iso, M. Kitahara, R. Tanida, T. Mitasaki, N. Ono, and A. Shimizu, "World's Highest-performance HEVC Software Coding Engine," NTT Technical Review, Vol. 12, No. 5, 2014.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201405fa4.html
[10] Y. Tanaka and D. Ochi, "Interactive Distribution Technologies for 4K Live Video," NTT Technical Review, Vol. 12, No. 5, 2014.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201405fa5.html
[11] T. Nakachi, T. Yamaguchi, Y. Tonomura, and T. Fujii, "Next-generation Media Transport MMT for 4K/8K Video Transmission," NTT Technical Review, Vol. 12, No. 5, 2014.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201405fa6.html
[12] K. Kawashima and J. Okamoto, "QoE Assessment Methodologies for 4K Video Services," NTT Technical Review, Vol. 12, No. 5, 2014.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201405fa7.html

Yoshiko Sugaya
Researcher, Visual Media Project, NTT Media Intelligence Laboratories.
She received the B.E. and M.E. degrees in mechanical engineering from Sophia University, Tokyo, in 2005 and 2007, respectively. She joined NTT Cyber Space Laboratories in 2007 and began conducting research on computer vision and projector-camera systems. She was subsequently engaged in product planning and development of telecommunication devices. Since 2013 she has been working on the development of high-definition visual communication service systems.

Hiroshi Fujii
Senior Research Engineer, Visual Media Project, NTT Media Intelligence Laboratories.
He received the B.E., M.E., and Ph.D. degrees in computer science from Kyoto University in 1989, 1991, and 2000, respectively. He joined NTT Information and Communication Processing Research Laboratory in 1991 and studied video distribution systems including multiple accesses to video data, content protection, video data storage, and digital watermarking. Then he developed digital rights management systems in a development project of the copyright management platform. Since 2006, he has been studying video encoding including SVC, HEVC, and evaluation of subjective video quality. He is a member of the Information Processing Society of Japan.

Atsushi Sato
Senior Research Engineer, Supervisor, Promotion Project 1, NTT Media Intelligence Laboratories.
He received the B.S. degree in mechanical sciences and engineering from Tokyo Institute of Technology in 1987 and the M.E. degree in information processing from Tokyo Institute of Technology in 1989. In 1989, he joined NTT Human Interface Laboratories, where he mainly engaged in R&D of computer vision and medical image processing systems and community ICT service systems. Recently, he has been developing high definition visual communication service systems.

Hiroaki Matsuda
Senior Research Engineer, Supervisor, Promotion Project 1, NTT Media Intelligence Laboratories.
He received the B.S. degree in communications engineering and the M.E. degree in information engineering from Tohoku University in 1987 and 1989, respectively. After joining NTT LSI Interface Laboratories in 1989, he engaged in R&D of VLSI design methodology for video coding and video codec systems, and content sharing platform systems. Recently, he has been working on the development of high definition visual communication service systems.

Shinichiro Chaki
Executive Research Engineer, Promotion Project 1, NTT Media Intelligence Laboratories.
He received the B.E. degree in electric and electrical engineering from Sophia University, Tokyo, in 1986 and the M.E. degree in electrical communication engineering from Waseda University, Tokyo, in 1988. He joined NTT Switching System Laboratories in 1988 and studied broadband NW architecture and switching systems based on ATM (Asynchronous Transfer Mode) technologies. From 1997 he was involved in developing an ATM PON system in NTT Access Network Service Systems Laboratories. He then engaged in developing network architecture and router systems for the FLET's premium network in NTT Service Integration Laboratories and FLET's Next in NTT Network Service Systems Laboratories. At NTT Media Intelligence Laboratories, he was developing systems based on speech, language, and image processing technologies. In March 2014, he moved to NTT Advanced Technology Corporation. He is a member of the Institute of Electrical and Electronics Engineers (IEEE) and the Institute of Electronics, Information and Communication Engineers (IEICE).

Hirohito Inagaki
Vice President, Head of NTT Media Intelligence Laboratories.
He received the B.E. and M.E. degrees in electrical engineering from Keio University, Kanagawa, in 1984 and 1986, respectively. Since joining NTT Electrical Communication Laboratories in 1986, he has been engaged in R&D of natural language processing and its applications. From 1994 through 1997, he was on transfer to NTT Intelligent Technology Co., Ltd., where he developed information security systems and multimedia systems. He moved to NTT Cyber Space Laboratories (now NTT Media Intelligence Laboratories) in 2006. His research interest is R&D of audio, video, and language interfaces and their applications. He is a member of IEICE.

HEVC Hardware Encoder Technology

Takayuki Onishi, Takashi Sano, Kazuya Yokohari, Jia Su, Mitsuo Ikeda, Atsushi Sagata, Hiroe Iwasaki, and Atsushi Shimizu

Abstract

For 4K and 8K video services that are capable of providing high definition pictures with an abundant sense of presence, the latest H.265/MPEG-H video coding standard, also known as HEVC (High Efficiency Video Coding), should allow the encoding of huge quantities of video data to be performed effectively and with high compression efficiency. This article introduces HEVC hardware encoder technology that is capable of encoding 4K video in real time.

Keywords: HEVC, 4K/8K, hardware encoder

1. Introduction

In recent years, advances in device technologies have resulted in the rapid proliferation of devices such as cameras, displays, and tablets that are capable of capturing and displaying 4K video, which has higher resolution than high-definition television (HDTV). With the proliferation of these devices, expectations are growing with regard to next-generation video services for the distribution of HD video via broadcasting and network delivery. In particular, 4K TVs are becoming increasingly popular in ordinary households, with many models now available from major TV manufacturers. It is also becoming possible to experience this technology at 4K live viewings and event broadcasts, and at movie theaters and other venues equipped with 4K projectors.

However, systems that use the H.264/MPEG-4 AVC (Advanced Video Coding) (also known as H.264) international video coding standard—the current mainstream technology—tend to be overwhelmed by the large amount of data (bit rate) needed for 4K video, and thus, there is a strong need for systems that support the latest international video coding standard H.265/MPEG-H, also called HEVC (High Efficiency Video Coding), which offers about twice the coding efficiency of H.264. The development of HEVC systems with high compression efficiency is also an important requirement for making 4K content viewable in every home and for delivering video using a variety of transmission media.

2. The need for real-time encoders

To broadcast and deliver HD content that is best appreciated live, for example, high-profile sporting events and live concerts, rather than pre-recorded content such as movies and dramas, it is essential to have a mechanism for encoding and transmitting video in real time. For transmission applications where video contributions are needed during program production, we require high-performance equipment that not only operates in real time but is also compact and has low latency, low power consumption, and faithful color reproduction. We therefore need a hardware encoder that can stably encode and deliver video while maintaining real-time performance.

3. Recent advances in video encoding hardware technology

At NTT laboratories, we have been researching and

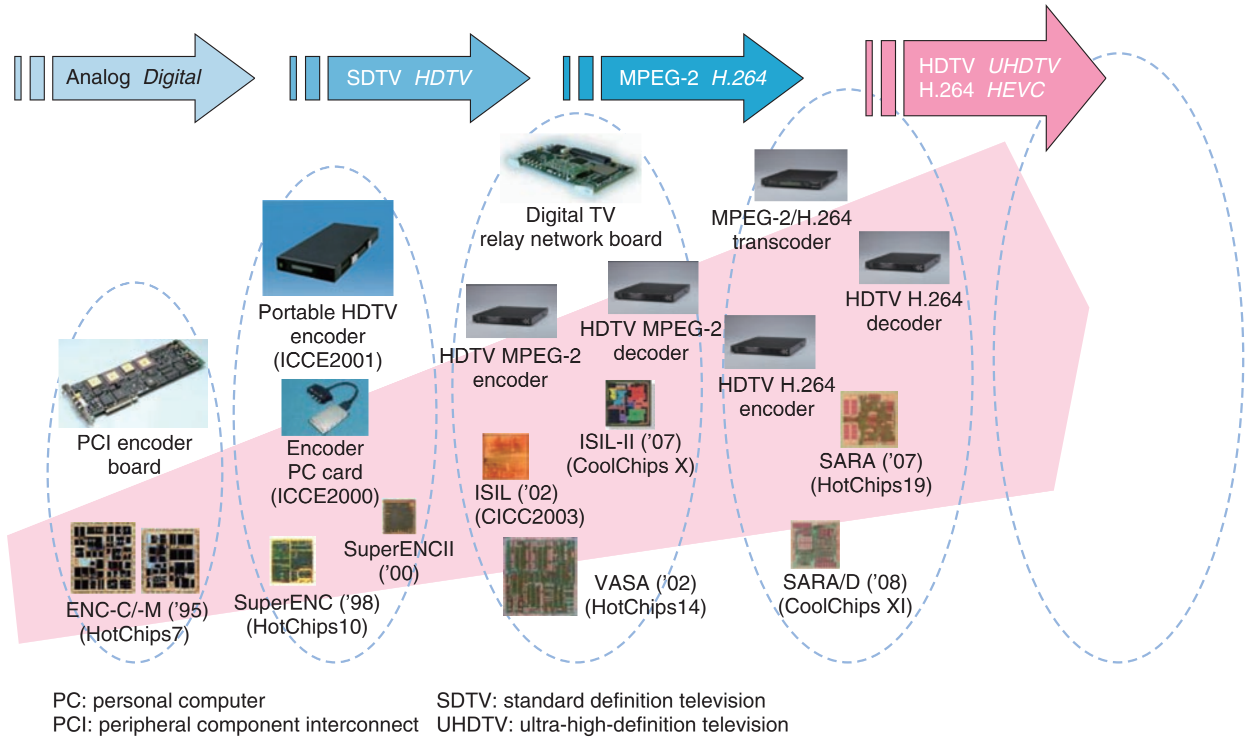


Fig. 1. Evolution of video encoding hardware technology.

developing technologies to realize real-time broadcasting and communication services for high-quality HD video, including the world's first single-chip HDTV MPEG-2 encoder large-scale integrated circuit (LSI) [1] and HDTV-compatible H.264 encoder LSI [2]. NTT's VASA single-chip HDTV MPEG-2 encoder LSI has contributed greatly to the spread and growth of digital TV relay network services transmitting video material, and our SARA H.264 encoder LSI has made a similar contribution to digital TV broadcasting and delivery services, including the HIKARI TV digital broadcasting and on-demand service. The accumulation of these technologies has accelerated the development of HEVC hardware encoder configuration technology for compatibility with next-generation 4K and 8K[*1] video (**Fig. 1**).

4. HEVC: The latest encoding standard

HEVC is the most recent international standard for video coding, the first issue of which (version 1) was published in January 2013 [3]. A second issue (version 2) is currently being drafted. This will include compatibility with high color reproducibility (4:2:2 of 4:4:4 encoding) for high-end professional applications, and hierarchical coding whereby the video quality is changed in a stepwise fashion suited to the available bit rate. The coding performance of HEVC is expected to be capable of achieving comparable picture quality with half the data needed for H.264 (widely used in One-Seg and smartphones) and just a quarter of the data needed for MPEG-2 coding (used for terrestrial digital TV and DVDs (digital versatile discs)). To achieve a practical bandwidth for the broadcasting and delivery of 4K and 8K video, where the amount of video data increases dramatically, it is essential to adopt HEVC with high compression efficiency.

Although the basic framework of HEVC coding is similar to that of the prior standards, it uses a much stronger combination of parameters such as block sizes and block shapes when subdividing and encoding an image, so as to efficiently eliminate redundancies

*1 4K and 8K: Video formats with resolutions of 3840 × 2160 and 7680 × 4320 pixels, respectively (equal to twice and four times the resolution of HDTV in each dimension), so-called because they have approximately 4,000 and 8,000 pixels in each horizontal line. The frame rate (number of frames displayed per second) has also been increased to 60 frames per second, which is twice as fast as HDTV.

Table 1. Comparison of the latest HEVC and H.264 coding standards.

	H.264	HEVC
Picture subdivision units	16 × 16 pixels (Macroblocks)	16 × 16 – 64 × 64 pixels (Coding tree units)
Prediction block size	4 × 4 – 16 × 16 pixels	4 × 4 – 64 × 64 pixels
Intra-prediction directions	8	33
Motion-compensated prediction block shape	Square, rectangular	Square, rectangular (asymmetric motion prediction)
Frequency transform block size	4 × 4, 8 × 8 pixels	4 × 4 – 32 × 32 pixels

contained within the image. A comparison of HEVC and H.264 is presented in **Table 1**. By adapting to image features, HEVC can encode large regions of uniform color (e.g., blue skies) all at once so as to eliminate unnecessary image subdivision overheads, while in regions with complex movements such as crowds of people, it uses detailed predictive coding to achieve a detailed separation of image regions corresponding to each movement. The use of such methods results in high overall coding efficiency.

5. HEVC encoder design issues

Although HEVC achieves highly efficient coding by using a wide combination of block sizes and block shapes, this places a significantly higher processing load on the video encoding processor and can cause severe loading. To improve the encoder performance, it is essential to find a solution that achieves high coding efficiency with fewer combinations, without loss of coding quality. In particular, to achieve high real-time performance in the design of a hardware encoder, a mechanism for efficiently selecting combinations with high coding efficiency must be implemented while balancing the trade-off between picture quality and the scale of the processing circuitry.

6. HEVC hardware design method

The development of a hardware encoder usually begins with a software simulation to evaluate the coding performance prior to the design stage. In practice, however, it is difficult to estimate the circuit scale of an actual hardware implementation, which results in problems such as prolonged simulation runs and the difficulties associated with evaluating the quality of large numbers of video images. We are therefore adopting the following two approaches based on the idea of searching for an optimal architecture by making flexible changes to the hardware configuration in order to improve the coding efficiency while evaluating the picture quality in actual hardware as well as in the simulation.

(1) SystemC development approach

SystemC is a hardware design language that can automatically generate hardware circuits from source code written in C++, a language widely used for software development. Compared with RTL (Register Transfer Level: a conventional hardware design language), the design of typical hardware design aspects such as the arithmetic unit configuration and clock handling is automated so that the designer can concentrate on developing the encoder processing algorithm. When changes are made to the encoding algorithm, a new hardware circuit can easily be generated automatically.

(2) Video codec development platform

We have developed a video codec development platform that can run hardware circuit simulations written in SystemC (**Fig. 2**). This platform can accommodate up to four modules based on rewritable hardware circuits in field programmable gate arrays (FPGAs), allowing hardware circuits generated by SystemC to be written and operated in real time. In addition to setting up input and output terminals for the video and encoded data on the main board, we also added the ability to make flexible changes to the connections between modules by mounting a separate FPGA to be connected between the four base modules.

These two approaches are used to successively implement and improve the hardware while striking a balance between circuit scale and picture quality, and as a result, it has become possible to increase the performance of the HEVC hardware encoder while reducing the time needed for design [4]. An example of a real-time 4K HEVC intra codec that was implemented in the video codec development platform is shown in **Fig. 3**.

This codec performs 4K encoding by splitting the

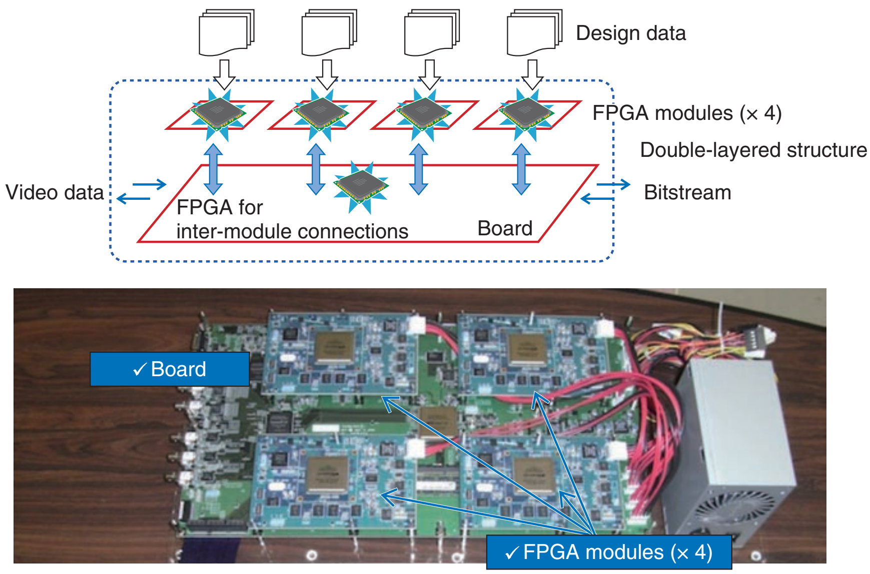


Fig. 2. Video codec development platform.

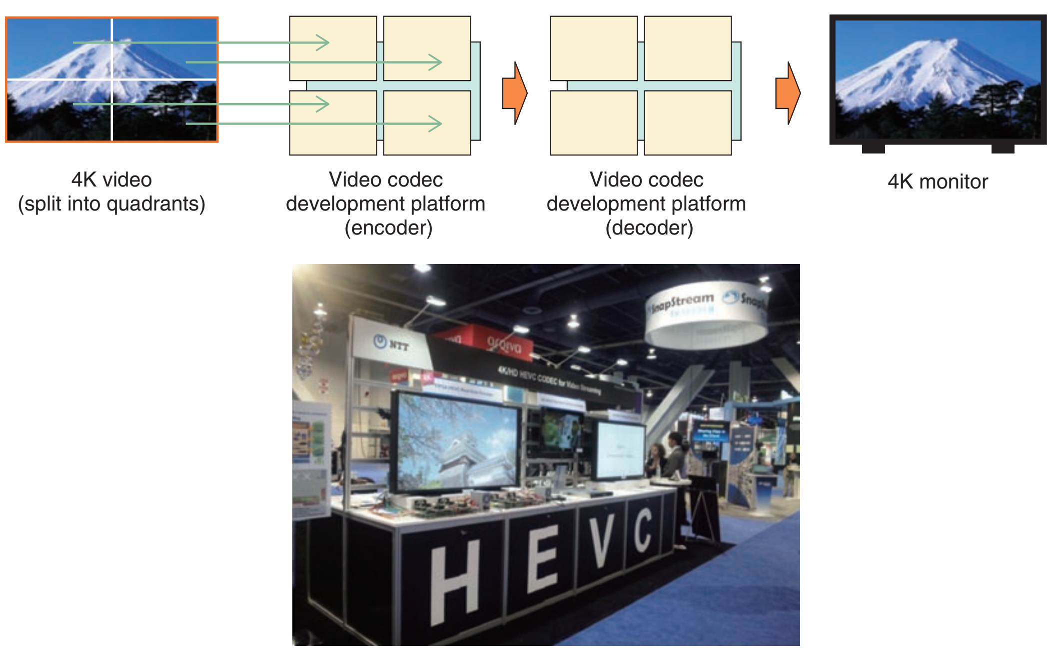


Fig. 3. Operation of 4K HEVC intra codec.

4K video into four quadrants that are separately input to module circuits configured for real-time HEVC intra-coding[*2] and decoding of video sized at HDTV dimensions in each FPGA.

In the future, we aim to continue using the same development method to further improve the hardware architectures for HEVC encoder LSIs.

*2 Intra-coding: A method for performing predictive coding based only on information from within the same video frame. H.264 and HEVC use a method that predicts pixel values extending in a specified direction based on the pixel values around the block to be encoded.

7. Future prospects

High-definition video is likely to be used in a widening range of environments as further improvements are made to display technology and camera performance.

We expect there will be an ever-growing need for real-time encoding technology with good picture quality and superior compression performance, and therefore intend to contribute to the expansion of high-quality broadcasting and video-on-demand services by continuing with our research and development of encoder technology as a key component of 4K and 8K video services.

References

[1] J. Naganuma, H. Iwasaki, K. Nitta, K. Nakamura, T. Yoshitome, M. Ogura, Y. Nakajima, Y. Tashiro, T. Onishi, M. Ikeda, and M. Endo: "VASA: Single-chip MPEG-2 422P@HL Codec LSI with Multi-chip Configuration for Large Scale Processing beyond HDTV Level," Hot Chips 14, Aug. 2002.

[2] K. Nitta, M. Ikeda, H. Iwasaki, T. Onishi, T. Sano, A. Sagata, Y. Nakajima, M. Inamori, T. Yoshitome, H. Matsuda, R. Tanida, A. Shimizu, K. Nakamura, and J. Naganuma: "An H.264/AVC High422 Profile and MPEG-2 422 Profile Encoder LSI for HDTV Broadcasting Infrastructures," 2008 IEEE Symposium on VLSI Circuits, pp.106–107, Honolulu, HI, June 2008.

[3] G. Sullivan, J. Ohm, W. Han, and T. Wiegand: "Overview of the High Efficiency Video Coding (HEVC) Standard," IEEE Transactions on Circuits and Systems for Video Technology, Dec. 2012.

[4] K. Yokohari, K. Nitta, M. Ikeda, and A. Shimizu: "Parallel Design Methodology for Video Codec LSI with High-Level Synthesis and FPGA-Based Platform," 50th Design Automatic Conference (DAC 2013), Austin, USA, June 2013.

Takayuki Onishi

Senior Research Engineer, NTT Media Intelligence Laboratories.

He received the B.E. and M.E. degrees in information and communication engineering from the University of Tokyo in 1997 and 1999, respectively. In 1999, he joined NTT Cyber Space Laboratories (now, NTT Media Intelligence Laboratories), where he engaged in R&D of high quality image coding and transmission. He is currently a senior research engineer in the Visual Media Project of NTT Media Intelligence Laboratories.

Takashi Sano

Research Engineer, NTT Media Intelligence Laboratories.

He received the B.E. and M.E. degrees in electronic engineering from Tokyo University of Science in 2002 and 2004, respectively. In 2004, he joined NTT Cyber Space Laboratories, where he has been engaged in R&D of software and hardware design of high quality image coding. He is currently a research engineer in the Visual Media Project of NTT Media Intelligence Laboratories.

Kazuya Yokohari

Researcher, NTT Media Intelligence Laboratories.

He received the B.E degree in computer science and engineering from the University of Aizu, Fukushima and the M.E. degree in communication engineering from Tohoku University, Miyagi, in 2006 and 2008, respectively. In 2008, he joined NTT Cyber Space Laboratories. He has been engaged in R&D of parallel processing architecture and high efficiency algorithms for video codecs. He is currently a researcher in the Visual Media Project of NTT Media Intelligence Laboratories.

Jia Su

Researcher, NTT Media Intelligence Laboratories.

She received the B.E. and M.E. degrees in telecommunications engineering from the Telecommunications Engineering school of Xidian University, China, in 2006 and 2009, respectively, and the M.E. and Ph.D. degrees from the Graduate School of Information, Production and Systems, Waseda University, in 2008 and 2012, respectively. In 2012, she joined NTT Cyber Space Laboratories. She has been engaged in R&D of parallel processing architecture and high efficiency algorithms for video codecs. She is currently a researcher in the Visual Media Project of NTT Media Intelligence Laboratories.

Mitsuo Ikeda

Senior Research Engineer, Supervisor, NTT Media Intelligence Laboratories.

He received the B.E. and M.E. degrees in mathematics and physics from Kyoto University in 1989 and 1991, respectively. In 1991, he joined NTT. He has been engaged in R&D on video coding and LSI design. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) and the Institute of Image Information and Television Engineers (ITE).

Atsushi Sagata

Senior Research Engineer, Supervisor, NTT Media Intelligence Laboratories.

He received the B.E. degree in electronic engineering from the University of Tokyo in 1994. He joined NTT in 1994 and has been engaged in R&D of video coding systems in NTT Cyber Space Laboratories. His current research interests include digital image processing, real-time hardware/software video coding systems, and video transmission systems. He is a member of IEICE of Japan.

Hiroe Iwasaki

Senior Research Engineer, Supervisor, NTT Media Intelligence Laboratories.

She received the B.S. and Ph.D. degrees in information science from Tsukuba University, Ibaraki, in 1992 and 2006, respectively. She joined NTT Switching System Laboratories in 1992. She has been engaged in R&D of architecture and a real-time operating system for multimedia embedded system LSIs. She moved to NTT Cyber Space Laboratories in 1997. She is a member of IEEE, IEICE of Japan, and the Information Processing Society of Japan (IPSJ).

Atsushi Shimizu

Senior Research Engineer, Supervisor, NTT Media Intelligence Laboratories.

He received the B.E. and M.E. degrees in electronic engineering from Nihon University, Tokyo, in 1990 and 1992, respectively. Since joining NTT in 1992, he has been engaged in video compression algorithm and software development. He is a member of IEICE, the Institute of Image Electronics Engineers of Japan (IIEEJ), and the Institute of Image Information and Television Engineers.

World's Highest-performance HEVC Software Coding Engine

Kazuyuki Iso, Masaki Kitahara, Ryuichi Tanida, Tokinobu Mitasaki, Naoki Ono, and Atsushi Shimizu

Abstract

In recent years, worldwide attention has been focused on the distribution of 4K and 8K high-definition video. Video services for mobile terminals have also been expanding rapidly, and higher compressive power in video coding is expected. In response to this situation, the international standard HEVC (High Efficiency Video Coding, otherwise known as H.265/MPEG-H) was established in April 2013. We introduce here a software coding engine developed by the NTT Media Intelligence Laboratories that conforms to HEVC and features the world's highest coding performance.

Keywords: H.265/HEVC, video coding, video distribution

1. Introduction

Attention has been focused worldwide in recent years on high-speed distribution of 4K and 8K high-definition (HD) video. Transmission of the huge amount of data associated with 4K and 8K video while maintaining sufficient quality requires the use of video coding techniques that have high compression performance. Video services for smartphones have gained in popularity in mobile communication environments in which capacity has expanded due to the increased use of LTE (Long Term Evolution) or other high-speed communication technology. Many video distribution service providers are implementing mobile services, and telecommunications carriers have strong expectations for higher compression of motion picture and video content in order to cope with the rapidly increasing communication traffic. In response to this situation, the NTT Media Intelligence Laboratories developed a software video coding engine that has the world's highest level of video compression and conforms to the Main/Main10 Profile of HEVC (High Efficiency Video Coding). HEVC is said to double the video data compression performance of the previous international standard, H.264/MPEG-4 AVC (Advanced Video Coding), also called H.264. Our software coding engine improves compression performance by a factor of up to 2.5, which enables a reduction in video data of up to 60% compared to H.264 with equivalent image quality.

2. HEVC software coding engine

NTT Media Intelligence Laboratories has developed a technique that takes image features into account to efficiently distribute the coding rate over the image, and in doing so, has achieved the world's highest performance HEVC software coding engine. This technique may eventually be applied to Internet protocol television (IPTV) and other such services. Therefore, we also developed a bit rate control technique that achieves stable video distribution, and a fast technique to compress video in a practical amount of time. To enable viewers to have a pleasant experience when they view video provided by IPTV and similar services, a function for adjusting the amount of data flowing through the network is necessary to achieve efficient video transmission. However, no such function is included in the international standards. We therefore developed a proprietary technique for controlling the coding rate in order to allocate the coding rate efficiently on the basis of statistical data.

Without fast compression techniques, HEVC

Fig. 1. Relation of image properties to subjective image quality.

involves trading high computational load and long processing time for compression performance. The HEVC software coding engine that we developed uses a proprietary fast compression technique to compress full HD video in about five times the playback time; in other words, if the playback time is 1 hour, this technique can compress the full HD video in 5 hours. This factor of five is generally considered a target for a practical processing time when working with actual content, so this software coding engine features both high compression performance and fast processing. With a steady eye on future expansion of the market for video distribution services, we are also targeting the Main 10 Profile for 4K and 8K high definition video. In the following sections, we describe the techniques developed by the NTT Media Intelligence Laboratories in more detail.

3. Local quantization parameter (QP) variation processing for high compression performance

The allocation of a limited coding rate is important for improving image quality. A known characteristic of human vision is that distortion is difficult to detect in regions of video that contain complex patterns. However, there are no complex patterns in the region of a human face in a single image, as shown in the example in **Fig. 1**, and distortion introduced by coding is therefore noticeable in such regions. On the contrary, the background region in the example image contains foliage that has complex patterns due to the many color variations in the tree trunks and leaves, so there is a tendency for coding distortion to go unnoticed. Consequently, techniques that make use of this tendency are often applied to measure the pattern complexity within an image. Then the coding rate is reduced in regions that include complex patterns in order to improve the overall coding efficiency.

We also focused on the relation between regions with complex patterns, peripheral areas, and temporal changes and discovered that the loss of subjective image quality is barely detectable in peripheral areas that have both complex patterns and large temporal changes, even if the coding rate is further reduced. We therefore developed a proprietary local QP variation processing technique in which the coding rate is adjusted for each region in an image. Specifically, the image is analyzed to detect regions that contain complex patterns before the coding process is executed. Controlling the coding rate according to the relation between peripheral areas and the degree of temporal change in regions where the patterns are complex improves the overall video compression performance.

4. Bit rate control for stable video transmission

IPTV and other video distribution services require a control for bit rate stability as well as high compression performance. Bit rate control is a fine control

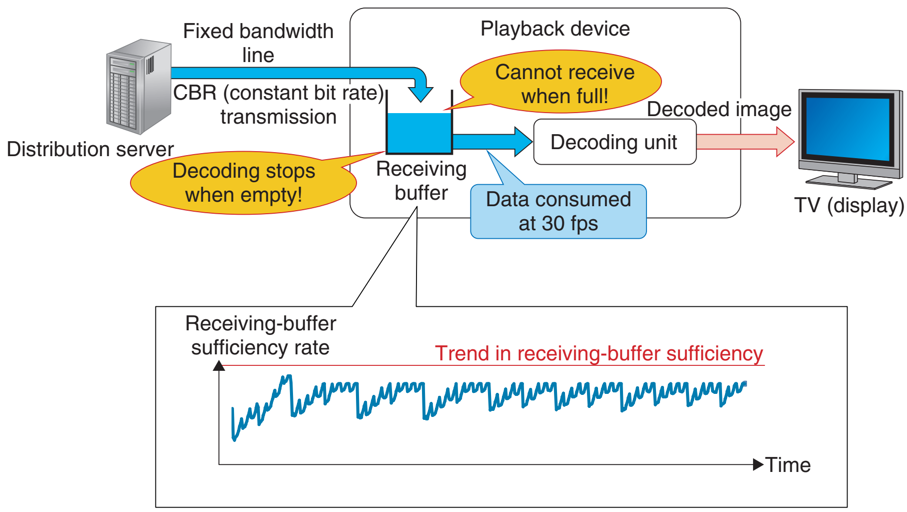


Fig. 2. Coding rate control.

that allocates a bit rate to each video frame and each region of a frame so as to distribute video at a constant bit rate (**Fig. 2**). NTT Media Intelligence Laboratories has developed this technique over a long period of time while developing the H.264 coding engine and other such technologies. Specifically, the coding rate of a previously coded image is divided into a rate for header data and for video data to achieve separate statistical processing. The results of the statistical processing can be used to estimate with high accuracy the coding rate for the subsequent image. HEVC uses a complex hierarchical reference structure that involves random access of more frames than in H.264, so the compression performance is higher. Therefore, performing hierarchical statistical processing can be used to control the coding rate for HEVC as well.

5. Video coding in practical time

In video coding, each image is divided into square blocks. Higher compression performance can be achieved in HEVC than in H.264 by selecting the block size from among four candidates. Additionally, many compression modes are prepared. To select the optimum block size and compression mode for each block, it is necessary to compare all combinations of block sizes and compression modes. However, simply comparing all the combinations requires a huge amount of computation time, so we developed a method that achieves fast processing while maintaining a higher compression rate by analyzing regions of high similarity and regions of local complexity in images.

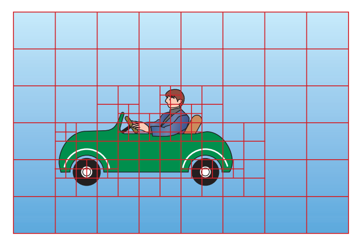


Fig. 3. Division of image into blocks.

A high compression rate can be achieved in HEVC by selecting a small block size for regions with complex video patterns and a large block size for regions with simple patterns (**Fig. 3**). Analysis of the results of block partitioning for different types of video reveals that the reduction in compression rate is slight even when the block sizes are fairly similar between adjacent blocks, because image features are similar for consecutive regions on the screen, even when the image is partitioned. We developed a method in which that information can be used to shorten the computation time by considering the block sizes of adjacent regions that have already been compressed (with compression being done sequentially from the top left of the screen to the bottom right) to reduce the number of block size candidates (**Fig. 4**).

Also, one important compression mode for improving compression performance is the skip mode. The

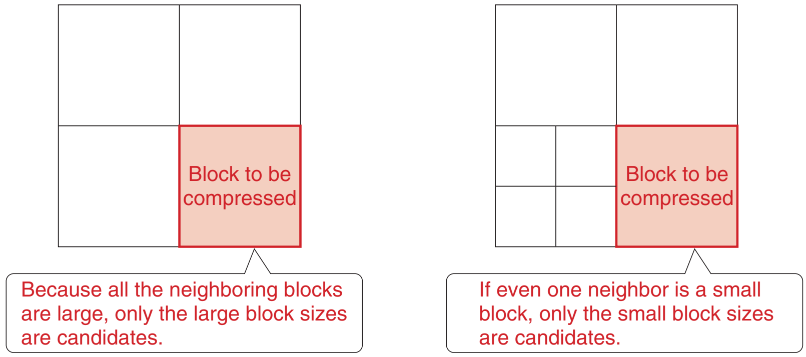


Fig. 4. Selection of candidates for block subdivision.

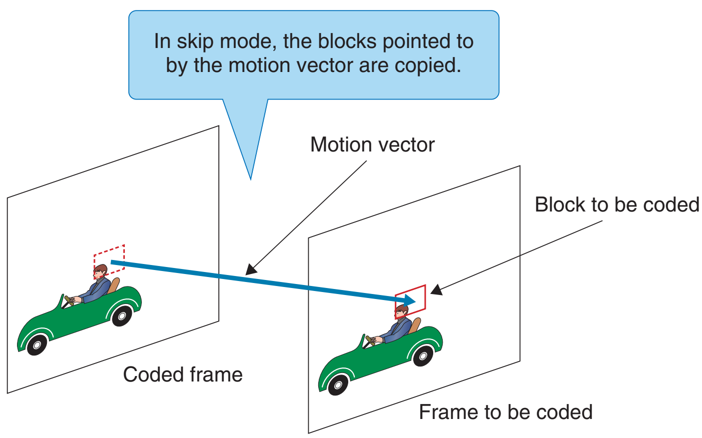


Fig. 5. Skip mode of video compression.

concept of the skip mode is illustrated in **Fig. 5**. In the skip mode, image blocks that are selected according to a motion vector that represents the movement of the subject within the image are coded by copying previous blocks. The skip mode has a wide application range in HEVC, so compression performance can be increased by compressing most of the image using this mode. Therefore, if we can determine that the skip mode is appropriate by estimating the image quality when using the skip mode, we can eliminate the coding processing for trials of other compression modes and thus greatly reduce the computation time. Specifically, we compare the motion vector of the block to be coded with the motion vector of the adjacent block that has already been coded to determine if the two blocks contain the same subject. If they do, we compare the image quality obtained by using the skip mode to the image quality of the adjacent block. If it is possible to achieve the same image quality, the skip mode is judged to be appropriate. That makes it unnecessary to try multiple compression modes and greatly reduces the processing time (**Fig. 6**).

Next, for compression modes other than the skip mode, a linear transform* is applied to the difference between this image and the previous image (predicted difference) to convert it to data that are easily compressed, and then the coding is performed. Many linear transformation methods are possible in HEVC, and comparing all of them would take a huge amount of computation time. We therefore conducted experiments using various types of video to determine the

* Linear transformation: A process applied to the difference data for the block to be compressed and the previous image block to reduce the amount of data involved in video compression. A typical conversion process is the discrete cosine transform.

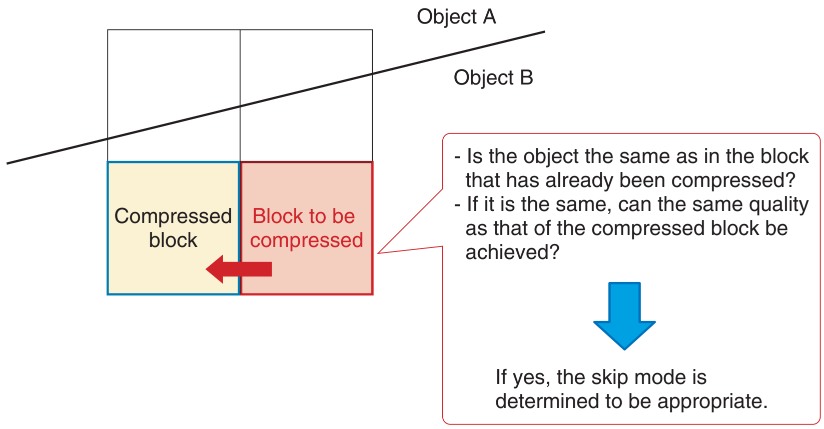


Fig. 6. Skip mode decision.

optimum linear transformation method. The results indicated there was a constant correlation of the image complexity of the regions to be coded and the block size distribution, so we can reduce the computation time by selecting the linear transformation method according to the image complexity and the block partitioning situation.

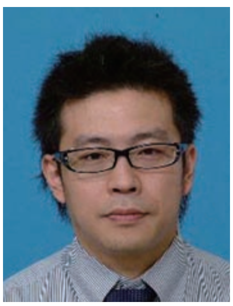
6. Future development

Our software coding engine has been commercialized by NTT Advanced Technologies and is being sold as a codec software development kit that includes the encoder and decoder as a set. The decoding engine for playing a stream generated by this coding engine was developed independently by NTT Advanced Technologies.

We intend to continue developing high image quality video distribution services that require even higher video compression performance. The NTT Media Intelligence Laboratories will push forward with research and development to contribute to the further expansion of the 4K and 8K video distribution market.

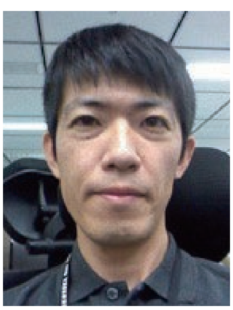
Kazuyuki Iso
Senior Research Engineer, Visual Media Coding Group, Visual Media Project, NTT Media Intelligence Laboratories.

He received the B.E. degree in electronics from Gunma National College of Technology in 1998, and the M.S. degree in knowledge science from Japan Advanced Institute of Science and Technology, Ishikawa, in 2000. He joined NTT in 2000 and engaged in research on visual communication systems. He is a member of the Institute of Electrical and Electronics Engineers (IEEE), the Virtual Reality Society of Japan (VRSJ), and the Information Processing Society of Japan (IPSJ).



Masaki Kitahara
Senior Research Engineer, Visual Media Coding Group, Visual Media Project, NTT Media Intelligence Laboratories.

He received the B.E. and M.E. degrees in industrial and management engineering from Waseda University, Tokyo, in 1999 and 2001, respectively. After joining NTT in 2001, he engaged in the development of compression and rendering algorithms for 3D video. He moved to NTT Advanced Technologies (NTT-AT) and managed various software development projects. In 2009, he returned to NTT and has been researching and developing HEVC encoding technologies. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).



Ryuichi Tanida
Research Engineer, Visual Media Coding Group, Visual Media Project, NTT Media Intelligence Laboratories.

He received the B.E. and M.E. degrees in electrical engineering from Kyoto University in 2001 and 2003, respectively. He joined NTT to study video coding for developing a real-time H.264 software encoder application. He moved to NTT-AT in 2009 to manage various software development projects related to video compression, distribution, and TV phones. He has been at NTT Media Intelligence Laboratories since 2012, as a research engineer in the visual media coding group. His current research focuses on the development of HEVC software encoders. He is a member of IEICE.

Tokinobu Mitasaki
Research Engineer, Visual Media Coding Group, Visual Media Project, NTT Media Intelligence Laboratories.

He received the B.E. and M.E. degrees in electrical engineering from Hokkaido University in 2001 and 2003, respectively. After joining NTT in 2003, he engaged in the development of holographic memory and video compression. He moved to NTT Communications in 2008 to develop Digital TV relay service. He has been at NTT Media Intelligence Laboratories since 2012, as a research engineer in the visual media coding group. His current research focuses on the development of HEVC software encoders. He is a member of IEICE.

Naoki Ono
Senior Research Engineer, Visual Media Coding Group, Visual Media Project, NTT Media Intelligence Laboratories.

He received the B.E. and M.E. degrees in information engineering from Niigata University in 1989 and 1991 respectively. He joined NTT in 1991 and has been engaged in video compression algorithm and software development. He is a member of IEICE and the Institute of Image Electronics Engineers of Japan (IIEEJ).

Atsushi Shimizu
Senior Research Engineer, Supervisor, Visual Media Coding Group, Visual Media Project, NTT Media Intelligence Laboratories.

He received the B.E. and M.E. degrees in electronic engineering from Nihon University, Tokyo, in 1990 and 1992, respectively. Since joining NTT in 1992, he has been engaged in video compression algorithm and software development. He is a member of IEICE, IIEEJ, and the Institute of Image Information and Television Engineers.

Interactive Distribution Technologies for 4K Live Video

Yasuaki Tanaka and Daisuke Ochi

Abstract

Video technology is becoming more sophisticated, and the amount of available content is increasing substantially. Accordingly, individual preferences for video content are becoming more diverse. We have perceived a need for a means of personalized viewing that enables viewers to select their preferred subjects and scenes within the content rather than just passively selecting video content created by third parties, as in the conventional style. This article introduces a system that partitions 4K video into tiles, compresses it, and enables only parts selected by the viewer within the video to be distributed live at the desired size and with high quality.

Keywords: interactive viewing, MVC, video distribution

1. Introduction

The selection of video content available to individual viewers has been increasing recently because of enhancements in video distribution services and the popularity of video sharing sites. Additionally, with set-top boxes and terminals such as smartphones and tablets, viewers are getting used to watching whatever program or video they like at any time. Moreover, the video content produced most recently includes camera work and scene switching that are intended to make the content accessible to the maximum number of viewers. Thus, there are cases in which not all subjects or scenes in a video are shown, even though some viewers might be interested in them. Consequently, it is not necessarily possible to satisfy all viewers.

NTT Media Intelligence Laboratories has been working to resolve this sort of issue and has been conducting research and development (R&D) on interactive viewing technology that enables individual viewers to dynamically select their preferred subject or scene within the content. Thus, viewers can personalize their view of the content to their preferences. This technology enables viewers to change the viewing size or area within the image using pinch or flick operations on a smartphone or tablet (**Fig. 1**). High-definition (HD) video sources such as 4K and 8K, which are higher definition than conventional HD, are expected to spread quickly in the future. However, on devices with a small viewing area, especially mobile terminals, it will not be possible to provide adequate resolution for viewing, so a viewing style that allows viewers to select the area and field of view themselves will be useful. However, video with 4K class image quality or greater requires an extremely large amount of data, so it is important to find a distribution method that will not overburden the communications bandwidth or terminal resources.

To this end, we have created a compression and distribution technology that can efficiently distribute only selected areas of a video, in order to achieve interactive viewing that enables users to view the specific areas they want to see, for video sources with resolutions of 4K and higher. We have also built an interactive distribution system for 4K live video that can perform all processing on 4K video in real time, from capture through compression and distribution.

2. Technologies for interactive viewing

2.1 Video compression and distribution technology

The system described here uses H.264/AVC

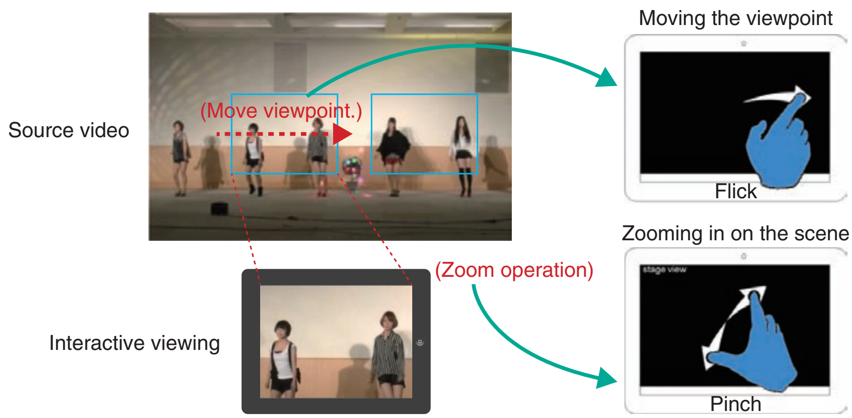


Fig. 1. 4K interactive viewing.

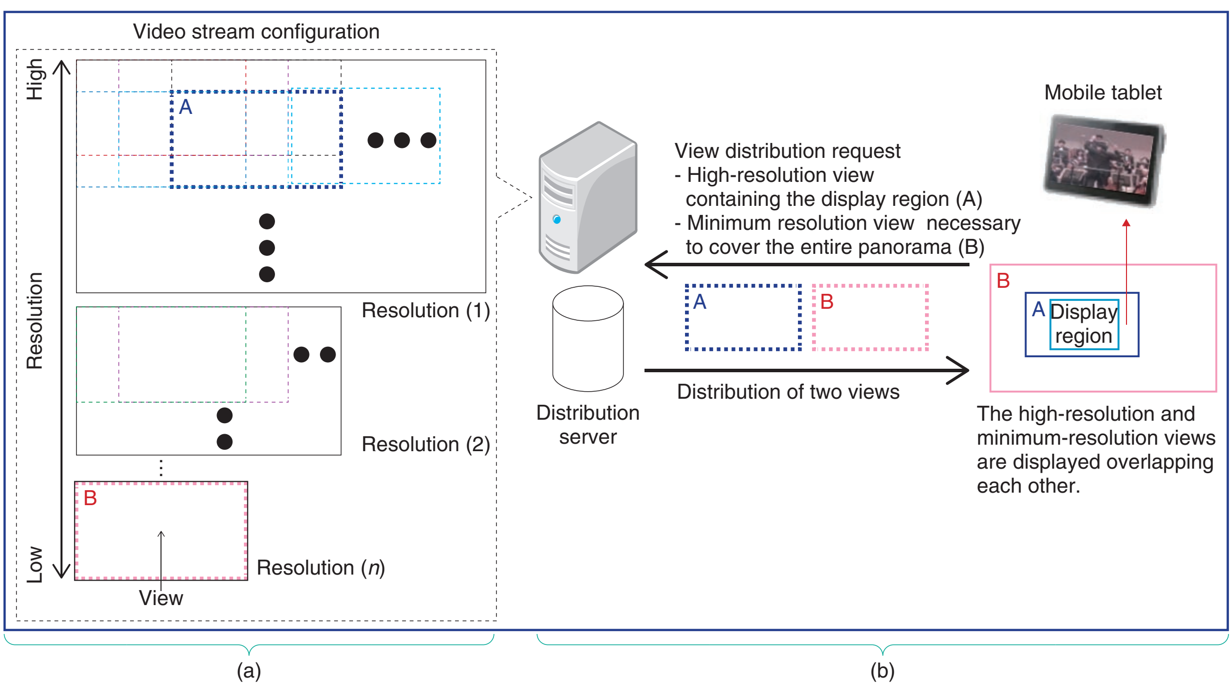


Fig. 2. 4K interactive distribution technology.

(Advanced Video Coding) Annex H multiview video coding (MVC) in order to efficiently distribute the areas selected by viewers. Originally, this method was standardized to handle multiple video streams (e.g., the right- and left-eye images for three-dimensional (3D) video), compressing them together while maintaining synchronization between them. In this case, the standard is used to compress the multiple videos together after partitioning the video source into small tile regions (**Fig. 2(a)**). Specifically, the video source is down-sampled to generate multiple videos at lower resolutions. These multiple videos together with the source video are partitioned into multiple tiles in the compression system. Next, all tile images are encoded in a synchronized compression that is coded into a video stream using MVC. Then, the tiles at each resolution are converted to multi-bit-rate data by coding at multiple compression rates. The

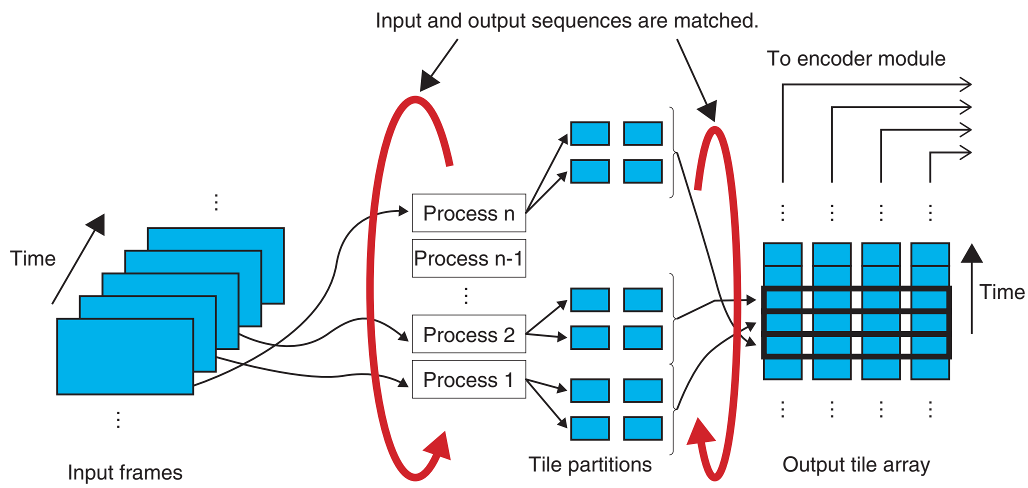


Fig. 3. Multi-thread processing of tile partitions.

compression-coded video stream is stored on the distribution server, and tile images with the appropriate position and resolution are distributed according to requests from individual viewers.

A distribution scheme is used in which the tiles containing the viewing area selected by the viewer are distributed with high resolution and at a high code rate, and tiles covering the range outside of the screen are distributed at low resolution and at a lower code rate (**Fig. 2(b)**). This is advantageous in that even if the viewer changes the viewing area randomly, the low-resolution tile can be viewed until the high-resolution tile for that area has been requested and delivered. Thus, no areas are dropped from the image, viewers can change the viewing area seamlessly, and network bandwidth is used more efficiently than simply distributing the entire high-resolution video stream.

2.2 High-speed compression technologies for live distribution

To provide live interactive viewing, all tiles generated from the video source must be compressed in real time, which requires more resources and processing time than for generating ordinary streams such as H.264/MPEG-4 AVC. Consequently, the speed of the compression process needed to be increased. The following three techniques were used to increase speed in the current system.

(1) Pipelining of module input/output (I/O) processing

The compression system in the current proposal requires several modules in sequence for resolution conversion, tile partitioning, encoding, and multiplexing. If the processes in each module are executed sequentially, the total processing time is the total of the processing times for each module, which results in large delays. To reduce the delay, data processing and data I/O for each module were done asynchronously, reducing the cumulative delay. Processing is pipelined within each module into three stages: input, processing, and output, and each stage is processed on a different thread. Data are transferred between threads using data buffers (FIFO[*1] queues) within the application.

(2) Multi-threaded tile generation

Resolution conversion and tile partitioning of the input video are done in units of video frames. There are no dependencies between these two processes, so an increase in speed was achieved by processing data frames in parallel. This can result in changes in frame order unless the I/O sequence is controlled. Therefore, we used a configuration in which the parallelization processing stage ran on a separate thread, and each thread had a separate FIFO queue. At the input stage, data are placed in each FIFO queue in a simple round-robin[*2] sequence, and at the output queues, data are removed in the same sequence as they were input on the input side (**Fig. 3**). Inputting and outputting data in the same sequence prevents any changes in the order.

*1 FIFO (First-In, First-Out): A method of storing data in a location and extracting it in the same order as it was stored.

*2 Round robin: A method of allotting a finite resource by taking turns in order.

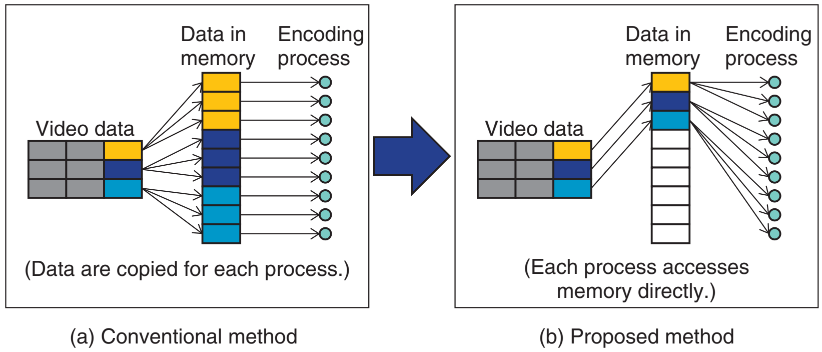


Fig. 4. Shared memory for the encoding process.

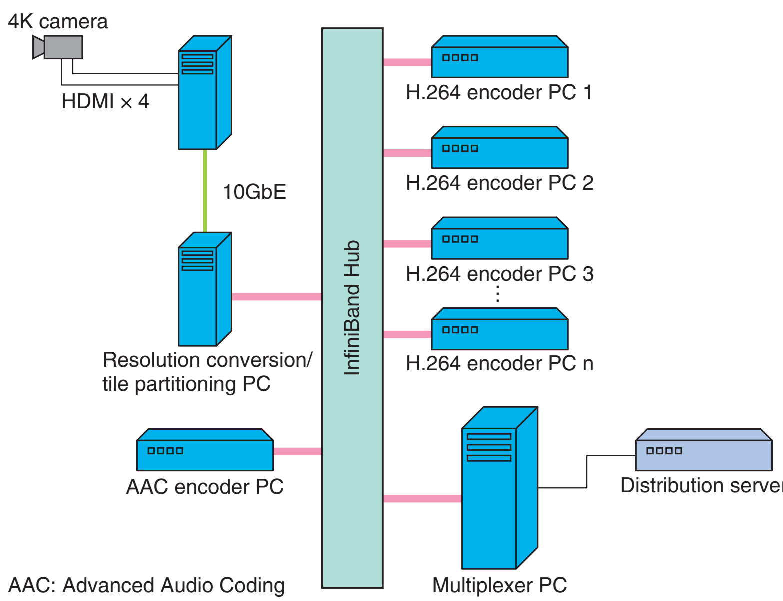


Fig. 5. System device configuration for live distribution tests.

(3) Handling multiple bit rates with shared memory

When this system encodes a tile at a given resolution and partition position at multiple bit rates, the data are copied to shared memory, and each encoding process accesses the same memory. This allows the tile data to be copied to shared memory only once, reducing the amount of copying done for inter-process communication (**Fig. 4**).

3. Interactive distribution system for 4K live video

The above technology was implemented in a system for interactive distribution of 4K live video, and live experiments were conducted to distribute video from the NTT Musashino R&D Center research facility, where the video was first captured, to the NTT Yokosuka R&D Center, where it was compressed and distributed, and back to the NTT Musashino R&D Center for playback.

An overview of the equipment configuration is shown in **Fig. 5**. Data transmission between the personal computers (PCs) was carried out using an InfiniBand Fourteen Data Rate (FDR) (4X) with theoretical throughput of 56 Gbit/s. There were 22 encoder PCs, and a commercial 4K camera was used for capturing and producing a video source with 4 HDMI (High-Definition Multimedia Interface) lines

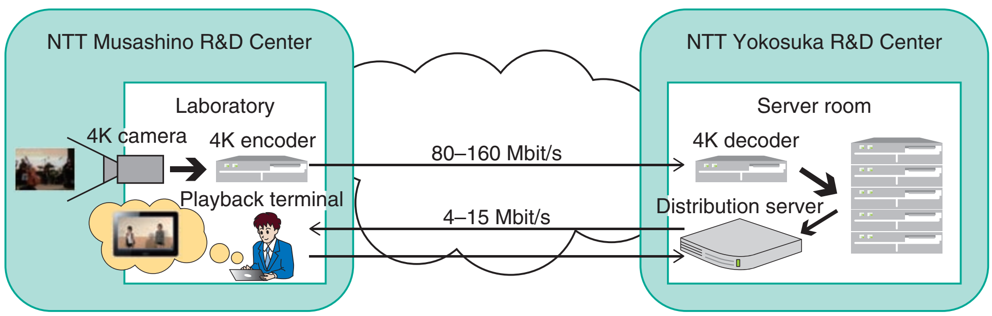


Fig. 6. Live distribution test configuration.

at 30 fps. This was compressed with MPEG-2 encoders generating total data at a rate ranging from 80 to 160 Mbit/s, which was transmitted to the remote location via a network using a network adapter. After this video was decoded, it was input to the encoder PC for interactive viewing. The video was compression-coded on the encoder PC, then sent in real time to the distribution server. We confirmed that viewers were able to view the video interactively over the network by connecting to the distribution server at approximately 4 to 15 Mbit/s, with a delay of approximately 7 s (**Fig. 6**).

4. Future prospects

These experiments achieved live interactive viewing, enabling users to view the areas of their choice at high quality from a 4K video source. This distribution and viewing technology is not limited to 4K video; it could also be applied to 8K video, wide panoramas, or whole-sky imagery. We are also conducting R&D in an effort to develop free-viewpoint video distribution, which allows a greater degree of freedom when changing viewpoints. This includes activities toward international standardization of 3DV/FTV (3D video/free-viewpoint television)[*3].

*3 3DV/FTV: International standards for 3D video coding methods by making use of multi-viewpoint video and depth maps under study at JCT-3V, a Joint Collaborative Team on 3D video that includes the international standardization organizations MPEG (Moving Picture Experts Group) and ITU-T (International Telecommunication Union, Telecommunication Sector). Multiple standards are also under study depending on the base compression method used, including 3D-AVC and 3D-HEVC (High Efficiency Video Coding).

Yasuaki Tanaka
Research Engineer, Visual Media Project, NTT Media Intelligence Laboratories.
He received the B.E. and M.E. degrees in electrical engineering from Osaka University in 1986 and 1988, respectively. He joined NTT in 1988 and studied optical disk systems. He is currently studying live video distribution techniques and 3D video communication systems.

Daisuke Ochi
Research Engineer, Visual Media Project, NTT Media Intelligence Laboratories.
He received the B.E. degree from the Department of Electrical Engineering and Computer Science and the M.E. degree in information engineering from Kyushu University, Fukuoka, in 2001 and 2003, respectively. His research interests include computer-human interaction, human interfaces, and multimedia information systems.

Next-generation Media Transport MMT for 4K/8K Video Transmission

Takayuki Nakachi, Takahiro Yamaguchi, Yoshihide Tonomura, and Tatsuya Fujii

Abstract

MPEG Media Transport (MMT) for heterogeneous environments is being developed as part of the ISO/IEC (International Organization for Standardization/International Electrotechnical Commission) 23008 standard. In this article, we overview the MMT standard and explain in detail the MMT LDGM (low-density generator matrix) FEC (forward error correction) codes proposed by NTT Network Innovation Laboratories. We also explain remote collaboration for content creation as a use case of the MMT standard.

Keywords: MPEG Media Transport, Low-density Generator Matrix Codes, ISO/IEC

1. Introduction

It has been about 20 years since the widely used MPEG-2 TS (MPEG-2 Transport Stream) standard*1 was developed by ISO/IEC (International Organization for Standardization/International Electrotechnical Commission) MPEG (Moving Picture Experts Group). Since then, media content delivery environments have changed. While video signals have been diversified, 4K/8K video systems have been developed. Moreover, there are a wide variety of fixed and mobile networks and client terminals displaying multi-format signals. The MPEG Media Transport (MMT) standard [1] is being developed as part 1 of ISO/IEC 23008 for heterogeneous environments. MMT specifies technologies for the delivery of coded media data for multimedia services over concatenated heterogeneous packet-based network segments including bidirectional IP networks and unidirectional digital broadcasting networks. NTT Network Innovation Laboratories has developed 4K transmission technologies for digital cinema, ODS (other digital stuff)*2, super telepresence, and other applications. We have contributed to the standardization of MMT in order to diffuse 4K transmission core technologies such as forward error correction (FEC) codes and layered signal processing.

2. MMT overview

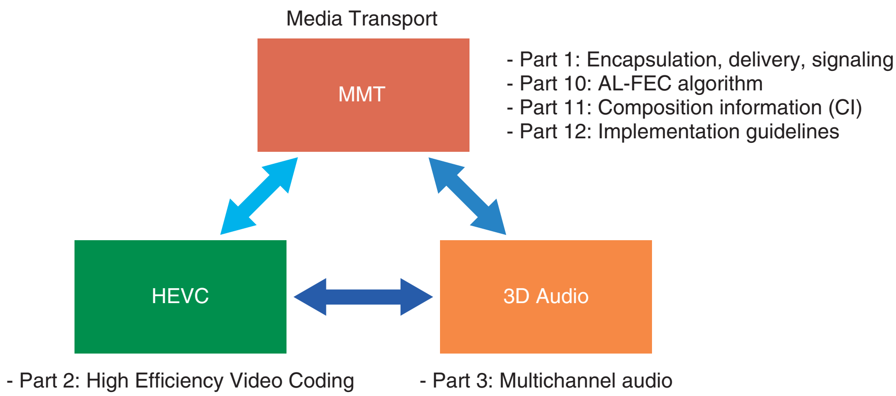
MPEG-2 TS provides efficient mechanisms for multiplexing multiple audio-visual data streams into one delivery stream. Audio-visual data streams are packetized into small fixed-size packets and interleaved to form a single stream. This design principle means that MPEG-2 TS is effective for streaming multimedia content to a large number of users.

In recent years, it has become clear that the MPEG standard is facing several technical challenges due to the emerging changes in multimedia service environments. One particular example is that the pre-multiplexed stream of 188-byte fixed size MPEG-2 TS packets is not quite suitable for IP (Internet Protocol)-based delivery of emerging 4K/8K video services due to the small and fixed packet size and the rigid packetization and multiplexing rules. In addition, it might be difficult for MPEG-2 TS to deliver multilayer coding data of scalable video coding (SVC) or multi-view coding (MVC) via multi-delivery channels. Error resiliency is also insufficient.

To address these technical weaknesses of existing

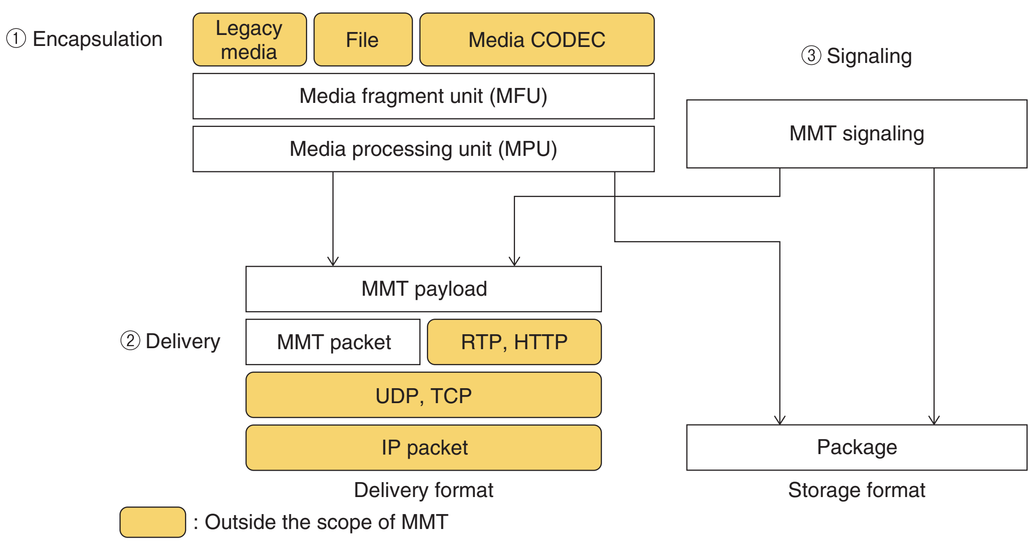
*1 MPEG-2 TS: A standard format for transmission developed by ISO/IEC MPEG. It is used for the current digital terrestrial broadcasting.

*2 ODS: A live-streaming application for theaters.



AL-FEC: application layer forward error correction

Fig. 1. Overview of MPEG-H.



HTTP: Hypertext Transfer Protocol
RTP: Real-time Transport Protocol
UDP: User Datagram Protocol

Fig. 2. MMT protocol stack.

standards and to support the wider needs of network-friendly transport of multimedia over heterogeneous network environments including next-generation broadcasting system, MPEG has been developing transport and synchronization technologies for a new international standard, namely MMT, as part of the ISO/IEC 23008 High Efficiency Coding and Media Delivery in Heterogeneous Environments (MPEG-H) standard suite. An overview of the MPEG-H standard is shown in **Fig. 1**. The MMT standard consists of Parts 1, 10, 11, and 12.

The protocol stack of MMT, which is specified in MPEG-H Part 1 [2] is shown in **Fig. 2**. The white boxes indicate the areas in the scope of the MMT specifications. MMT adopts technologies from three major functional areas. 1) Encapsulation: coded audio and video signals are encapsulated into media fragment units (MFUs)/media processing units (MPUs). 2) Delivery: an MMT payload is constructed by aggregating MPUs or fragmenting one MFU. The

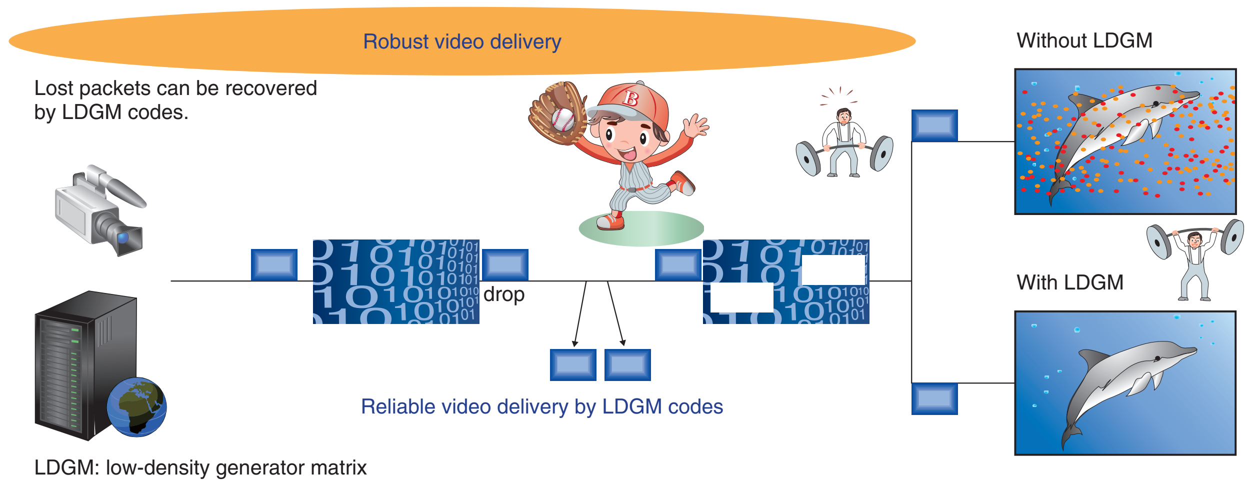


Fig. 3. Reliable video transmission.

size of an MMT payload needs to be appropriate for delivery. An MMT packet is carried on IP-based protocols such as the User Datagram Protocol (UDP) or Transmission Control Protocol (TCP). It is a variable-length packet appropriate for delivery in IP packets. Each packet contains one MMT payload, and MMT packets containing different types of data and signaling messages can be transferred in one IP data flow.
3) Signaling: MMT signaling messages provide information for media consumption and delivery to the MMT client.

A client terminal identifies MPUs constituting the content and their presentation time by processing the signaling message. The presentation time is described on the basis of Coordinated Universal Time (UTC)*3. Therefore, the terminal can consume MPUs in a synchronized manner even if they are delivered on different channels from different senders. MMT defines application layer forward error correction (AL-FEC) codes in order to recover lost packets, as shown in **Fig. 3**. MPEG-H Part 1 specifies an AL-FEC framework, and MPEG-H Part 10 [3] specifies AL-FEC algorithms. Furthermore, MPEG-H part 11 defines composition information (CI), which identifies the scene to be displayed using both spatial and temporal relationships among content. These functions mentioned above are the novel features of MMT, and are not shared by MPEG-2 TS.

3. Powerful FEC codes

MPEG-H Part 10 defines several AL-FEC algorithms, including Reed-Solomon (RS) codes*4 and low-density generator matrix (LDGM) codes proposed by NTT Network Innovation Laboratories. Each AL-FEC code has advantages and disadvantages in terms of error recovery performance and computation complexity. RS codes are based on algebraic structures. They give an optimal solution in terms of MDS (minimum distance separation) criteria. However, an RS code makes it difficult to increase the block length because the decoding process features polynomial-time decoding. It cannot reach Shannon's capacity*5, which gives the theoretical limit. This can result in low coding efficiency and excessive computation overhead. Unlike RS codes, LDGM codes can handle block sizes of over several thousand packets because of their very low computation complexity. In particular, LDGM codes are suitable for the transmission of huge data sets such as 4K/8K video (**Fig. 4**).

LDGM codes are one type of linear code; the party check matrix H*6 consists of a lower triangular matrix (LDGM structure) that contains mostly 0's and only

*3 UTC: UTC is the time standard commonly used across the world. It is used to synchronize time across Internet networks. In the MMT standard, it can be used as a timestamp.

*4 Reed-Solomon (RS) codes: RS codes are error-correcting codes in which redundant information is added to data so that it can be reliably recovered despite errors in transmission. Decoding is time-consuming when the block length is increased.

*5 Shannon's capacity: Error-correcting codes can alleviate some errors, but cannot alleviate all the errors introduced by the channel in a digital communications system. Shannon's capacity theorem states that error-free transmission is possible as long as the transmitter does not exceed the channel's capacity.

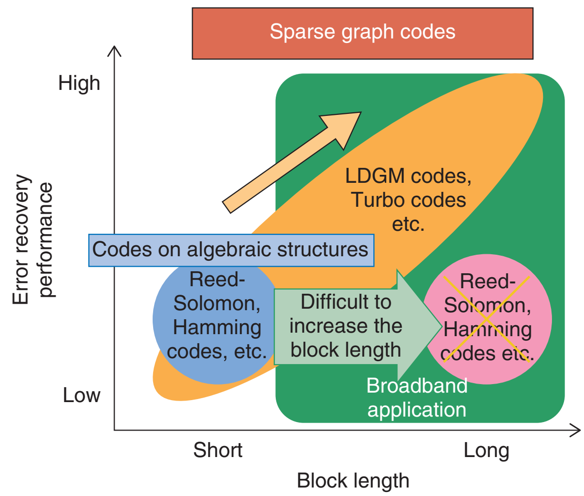


Fig. 4. FEC codes.

a small number of 1's. An example of LDGM encoding and decoding procedures is shown in **Fig. 5**. Using a sparse parity generator matrix makes it possible to process large code blocks of over a thousand IP packets as a single code block, which offers robustness to packet erasure error in networks. Our proposed LDGM codes provide good error recovery performance while keeping the computational complexity low because the sub-optimal parity generator matrix is used when applied to the message passing algorithm (MPA). Furthermore, the proposed LDGM codes can use irregular matrices that provide good error recovery performance for both MPA and MLD (maximum likelihood decoding).

Furthermore, the specified LDGM codes can support the following two schemes. One is a sub-packet division and interleaving method [4]. Generally, LDGM codes provide superior error recovery performance for large code-block sizes. However, the error recovery performance for short code-blocks is inferior to that of the RS codes. The proposed sub-packet division and interleaving LDGM scheme solves this problem. This scheme increases the number of symbols in one block size and increases the error recovery performance.

The other is a Layer-Aware LDGM (LA-LDGM) scheme [5]. The structure of conventional LDGM codes does not support partial decoding. When the conventional LDGM codes are used for scalable video data such as JPEG2000 created by JPEG (Joint Photographic Experts Group) or SVC code streams, performance is low and scalability is lost. The LA-LDGM scheme maintains the corresponding relationship of each layer. The resulting structure supports partial decoding. Furthermore, the LA-LDGM codes create highly efficient parity data by considering the relationships of each layer. Therefore, LA-LDGM codes raise the probability of recovering lost packets.

4. Use cases of MMT standard

MMT has various functions. Several representative MMT functions are shown in **Fig. 6**. Reliable 4K/8K video transmission via public IP networks is one of the typical features of MMT due to its high error recovery capability. Synchronization*7 between the content on a large screen and the content on a second display can also be realized by MMT. Users can enjoy public viewing while displaying alternative camera views on their second display. Different content selected by a user can be simultaneously transported via different networks. Hybrid delivery of content in next-generation Super Hi-Vision broadband systems can also be achieved. Applications such as widgets can be used to present information together with television (TV) programs. In particular, the large and extremely high-resolution monitors needed for displaying Super Hi-Vision content would be able to present various kinds of information obtained from broadband networks together with TV programs from broadcasting channels.

5. Remote collaboration for content creation [6]

High quality video productions such as those involving films and TV programs made in Hollywood are based on the division of labor. The production companies for video, visual effects (VFX), audio, and other elements of a production all create their content in different locations. Post-production is the final stage in filmmaking, in which the raw materials are edited together to form the completed film. In conventional program production, as shown in **Fig. 7(a)**, these raw materials are gathered physically from the separately located production sites. An overview of

*6 Parity check matrix H: The parity check matrix defines the relationships between the encoding symbols (source symbols and repair symbols), which are used by the decoder to reconstruct the original k source symbols if some of them are missing.

*7 Synchronization: Several media content files that have been delivered via various networks can be synchronized by using UTC based timestamps. However, the method of implementing the synchronization is outside the scope of the MMT standard.

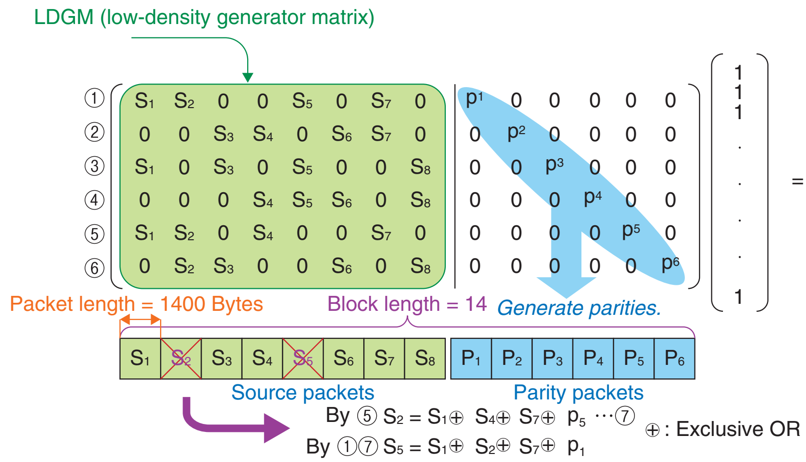


Fig. 5. LDGM codes.

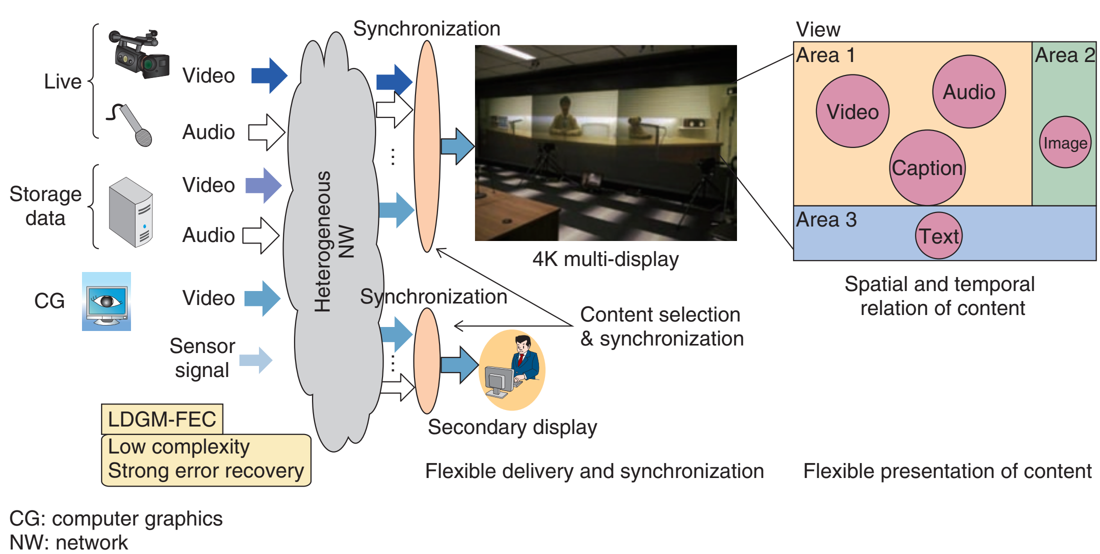
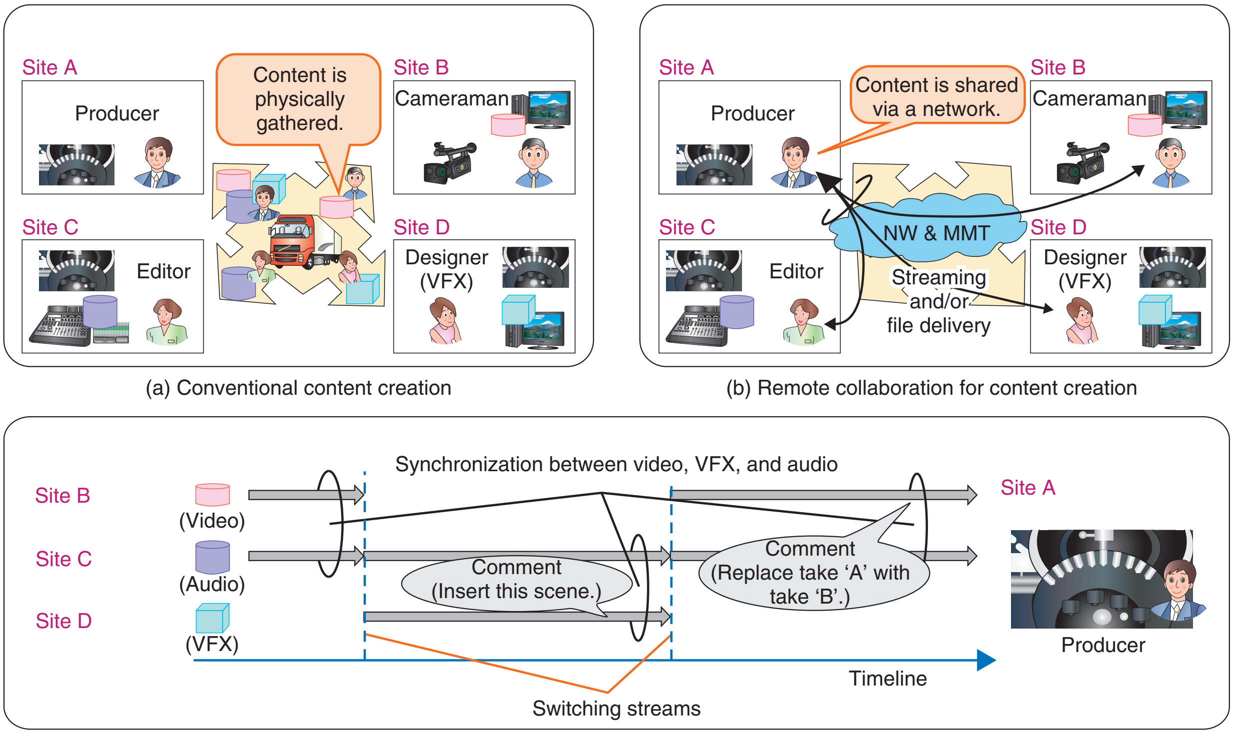


Fig. 6. MMT functions.

remote collaboration for content creation and editing actions on a timeline is shown in **Fig. 7(b)** and **(c)**. By using the MMT standard, multiple forms of content such as video, VFX, and audio are shared via the network, and the selected content can then be synchronized based on the producer's request. All the staff members at each location perform their tasks and share their comments simultaneously. A remote collaboration system that uses MMT enhances the speed of decision-making. As a result, we can increase the efficiency and productivity of content creation. This is an example use case of the MMT standard. Various other services are expected to emerge with the full implementation of MMT standard technologies.

6. Future direction

NTT Network Innovation Laboratories will continue to promote innovative research and development of reliable and sophisticated transport technologies such as stable video transmission on shared networks consisting of multi-domain networks and virtual network switching to achieve dynamically configurable remote collaboration.



(a) Conventional content creation

(b) Remote collaboration for content creation

(c) Editing on timeline

Fig. 7. Remote collaboration for content creation.

References

[1] T. Nakachi, "An Emerging MMT Standard for a Next Generation Media Platform," The 26th Communication Systems Workshop (CSWS), Nov. 2013.

[2] ISO/IEC 23008-1 FDIS, Information technology - High Efficiency Coding and Media Delivery in Heterogeneous Environments - Part 1: MPEG Media Transport (MMT).

[3] ISO/IEC 23008-10 DIS, Information Technology - High Efficiency Coding and Media Delivery in Heterogeneous Environments - Part 10: MPEG Media Transport Forward Error Correction Codes (MMT FEC Codes).

[4] T. Tonomura, D. Shirai, K. Kitamura, T. Nakachi, T. Fujii, and H. Kiya, "Construction Method and Theoretical Analysis of Packet-level Low-density Generator Matrix Codes to Allow Backward Compatibility for Video Streaming," IEICE Trans. Fundamentals (Japanese edition), vol. J93-A, no. 3, pp. 212–215, March 2010.

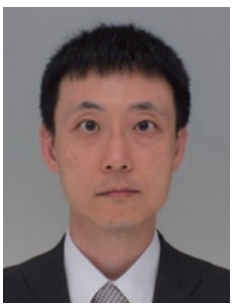
[5] Y. Tonomura, D. Shirai, T. Nakachi, T. Fujii, and H. Kiya, "Layered Low-density Generator Matrix Codes for Super High Definition Scalable Video Coding System," IEICE Trans. on Fundamentals, vol. E92-A, no.3, pp. 798–807, March 2009.

[6] T. Nakachi, Y. Tonomura, and T. Fujii, "A Conceptual Foundation of NSCW Transport Design Using an MMT Standard," IEEE ICSPCS2013, Dec. 2013.

Takayuki Nakachi

Senior Research Engineer, Media Processing Systems Research Group, NTT Network Innovation Laboratories.

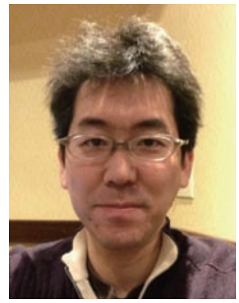
He received a Ph.D. degree in electrical engineering from Keio University, Tokyo, in 1997. Since joining NTT in 1997, he has been engaged in researching super high definition (SHD) image coding, especially in the area of lossless and near-lossless coding. In recent years, he has been researching scalable image/video coding in order to distribute SHD image content. From 2006 to 2007, he was a visiting scientist at Stanford University, CA, USA. He is currently Senior Research Engineer of the Media Processing Systems Research Group in NTT Network Innovation Laboratories. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) and IEEE.



Takahiro Yamaguchi

Senior Research Engineer, Supervisor, NTT Network Innovation Laboratories.

He received his B.E., M.E., and Ph.D degrees in electronic engineering from the University of Electro-Communications, Tokyo, Japan, in 1991, 1993, and 1998, respectively. He joined NTT Optical Network Systems Laboratories in 1998 and has been researching SHD image distribution systems. He is a member of IEICE and the Information and Television Engineers (ITE) of Japan.



Yoshihide Tonomura

Research Engineer, NTT Network Innovation Laboratories.

He received his B.S. and M.S. degrees in electronics engineering from Nagaoka University of Technology, Niigata, in 2002 and 2004, respectively, and the Ph.D. degree from Tokyo Metropolitan University in 2010. He joined NTT Network Innovation Laboratories in 2004. His research is focused on image processing theories and applications. From 2011 to 2012 he was a visiting researcher in the Media Lab of Massachusetts Institute of Technology, Boston, MA, USA. He is a member of IEICE.

Tatsuya Fujii

Group Leader, Media Processing Systems Research Group, Director of Digital Cinema Project, NTT Network Innovation Laboratories.

He received his B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Tokyo in 1986, 1988, and 1991, respectively. He joined NTT in 1991. He has been researching parallel image processing and SHD image communication networks. In 1996, he was a visiting researcher at Washington University in St. Louis, MO, USA. He is currently a group leader of the media processing systems research group and a director of the Digital Cinema Project in NTT Network Innovation Laboratories. He is a member of IEICE, ITE of Japan, and IEEE.

QoE Assessment Methodologies for 4K Video Services

Kimiko Kawashima and Jun Okamoto

Abstract

The technological developments of 4K/8K video enable users to watch video content with high quality and a high *sense of presence*. To provide attractive 4K video content, it is important to design and manage services based on the viewers' quality of experience (QoE). This article introduces QoE assessment methodologies and standardization activities.

Keywords: 4K video, QoE, subjective quality assessment

1. Introduction

The 4K and 8K video specifications have been attracting attention in recent years for their use in next-generation video services, and various electronics manufacturers are already selling 4K televisions (TVs). Also, the Next Generation Television & Broadcasting Promotion Forum (NexTV-F) established by the Ministry of Internal Affairs and Communications has begun studying standards and specifications for 4K video transmission and reception. This forum's objective is to create an environment in which any user can enjoy an immersive high quality and high sense of presence video experience, which gives users a feeling of *being there*, beginning with experimental 4K broadcasting of the opening of the World Cup to be held in Brazil in 2014 and 8K broadcasting of the opening of the 2016 Rio de Janeiro Olympics and the 2020 Tokyo Olympics.

To provide attractive 4K video, it is important to design and manage the video service based on the user quality of experience (QoE). A methodology for assessing the QoE of 4K video service is therefore essential. We describe here efforts concerning QoE assessment of 4K video services.

2. 4K video QoE

The resolution of 4K video is four times that of conventional full high-definition (HD) images, so 4K video can be viewed with very high image quality. Additionally, the bit depth of 4K video is 8 bits or more, and the frame rate is two to four times higher than that of conventional full HDTV. Consequently, users can experience a higher sense of presence, sense of depth, and sense of immersiveness such as that experienced by users watching 3D movies, than with conventional HDTV [1]. We therefore need to evaluate the sense of presence, sense of depth, and sense of immersiveness as QoE factors in addition to the video quality as has been done in the past. In this article, we narrow our focus to a method of assessing 4K video quality.

3. 4K video subjective assessment methodology

Subjective quality assessment methods are the most fundamental way of assessing the QoE of video communication services. Subjective quality assessments involve visual psychological tests where participants are subjected to a video stimulus and then assess its quality based on their own subjective judgment. Special knowledge and evaluation equipment are needed to plan and implement the tests, including the selection of a video presentation method, a suitable way of measuring the assessments, and a means of adjusting the viewing conditions. Therefore, the International Telecommunication Union (ITU) has published recommendations relating to subjective quality assessments in order to regulate the viewing environments

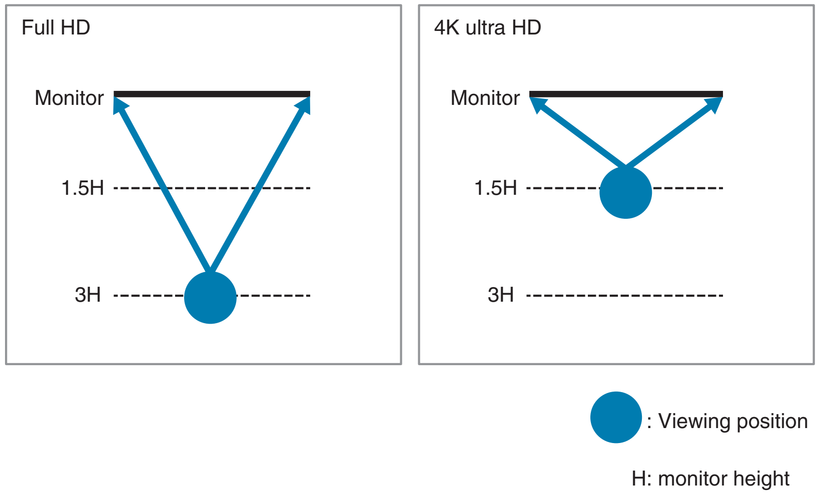


Fig. 1. Viewing distance.

and viewing criteria with the aim of obtaining repeatable quality assessments.

Some organizations are working to develop methods to assess 4K video quality. The Video Quality Experts Group (VQEG) has established the Ultra-HD (4K) project, which has begun studies on 4K video quality assessment. This project is targeting subjective 4K video quality assessment methods and objective quality assessment methods that can estimate the results of subjective assessment of QoE from the physical characteristics of the 4K video signal or other such means. The 4K subjective quality assessment method basically follows the same process as conventional full HD subjective quality assessment methods.

Here, we introduce one of the conditions in subjective quality assessment tests that are standardized in conventional full HD subjective quality assessment methods. For example, the environment (viewing conditions) in which the video quality is assessed is set forth in ITU-R (ITU Radiocommunication sector) recommendations including BT.710 [2] and BT.1129 [3]. These recommendations state that the assessment tests must be carried out in a special room (quality assessment booth) where it is possible to create the specified environmental conditions. In addition, these recommendations define criteria for the viewing distance, luminance of the monitor, and number of participants. The quality experienced by participants (subjective quality) can vary widely, even when they are watching the same video. To reduce the variation in the assessment results, a single video sample must be scored by a large number of participants. ITU-R recommendation BT.500 [4] defines that the participants should consist of at least 15 non-experts (people who are not routinely involved in work relating to the quality of TV pictures and who have limited experience in assessment tests).

The subjective assessment method of 4K video quality follows the subjective assessment method of full HD video quality, so we discuss the possibility of applying those standards to the subjective assessments of 4K video.

3.1 Viewing distance for 4K video assessment

The viewing distance for full HD video is specified in ITU-R BT.710 to be 3H, where H is the monitor height (**Fig. 1**). This viewing distance is the same as the distance at which a person with a visual acuity of 1.0 on the Japanese scale cannot perceive the scanning lines on the screen [5]. The high resolution of 4K video, however, makes it possible to view from a closer distance without perceiving the scan lines. Specifically, because 4K video has a horizontal pixel count of 3840, and the viewing resolution is 60 pixels per degree of screen angle, the screen angle is about 64 degrees. For a viewing distance of 1.5H, that is about the same as the viewing angle (61°). Actually, 4K video can provide a strong sense of detail, and the image quality is maintained even when the viewer moves as close as 1.5H. We therefore assume a viewing distance of 1.5H for the subjective assessment experiments for 4K video; this distance is closer to the monitor than in subjective assessment of

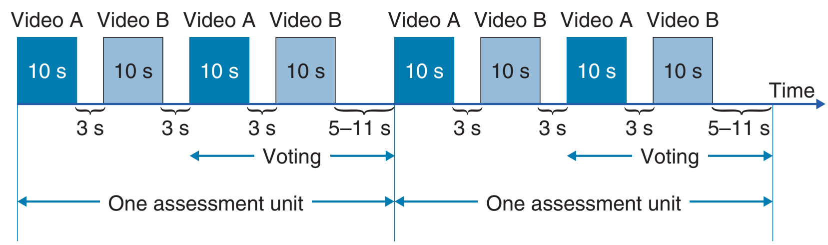


Fig. 2. Flow of DSCQS method.

conventional full HD video (Fig. 1). Moreover, most 4K video quality assessment studies that have been done have used a viewing distance of 1.5H.

3.2 Assessment metric of 4K video quality

The Double Stimulus Continuous Quality Scale (DSCQS) method set forth in ITU-R recommendation BT.500 is widely used in assessing the quality of systems and transmission paths used for television broadcasts. This method is particularly effective in cases where it is not possible to present the full range of quality conditions, and it is useful for simultaneously assessing the difference in quality between a reference video and an assessment video, and for assessing the absolute quality of the assessment video.

The DSCQS method, as shown in **Fig. 2**, involves using a pair of videos comprising the reference video and an assessment video that has been subjected to some sort of processing such as video coding. These videos are presented twice, and the assessment is performed the second time the videos are presented. The videos are presented in random order, and the participants are not told which is the reference video. The participants score the videos on a continuous quality assessment scale based on five categories, as shown in **Fig. 3**. The assessment scale is normalized to the range 0–100 (maximum value: 100, minimum value: 0), and the difference in video assessment values for the reference video and assessment video in each pair is calculated. These video quality differential values are averaged across all the participants to yield a DSCQS value. Because the DSCQS value is calculated from the differences in video quality, a smaller value indicates higher quality (closer to the reference video), and a larger value indicates lower quality.

With 4K video, the quality of the encoded video may be close to that of the reference video, depending on the encoding rate. Consequently, there is concern that non-experts will not be able to make a stable assessment of the difference in quality between a reference video and an assessment video (**Fig. 4**). Therefore, we conducted two subjective assessment tests using high-quality assessment videos that were close in quality to the 4K reference videos and using assessment videos that ranged in quality from low to high. The 36 participants (18 males, 18 females) ranged in age from 20 to 29 years old and had visual acuity of 1.0 on the Japanese scale. All participants joined both of the subjective assessment tests. We compared the dispersion around the average DSCQS values at the 95% confidence intervals for 4K video viewed at a distance of 1.5H with that for full HD video quality at a viewing distance of 3H (**Fig. 5**). The graph in Fig. 5 indicates that the dispersion in DSCQS values is about the same for 4K video and full HD video. We can thus show that non-experts can sufficiently assess the difference in quality between

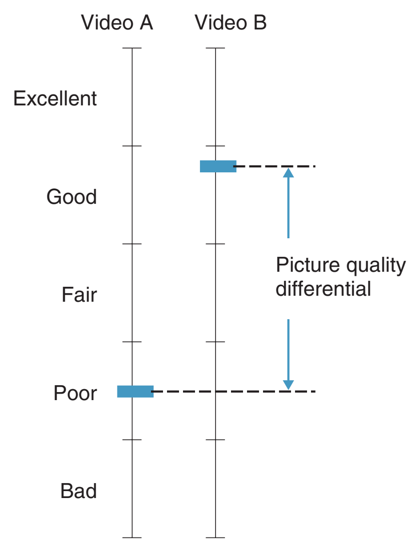


Fig. 3. DSCQS assessment categories.

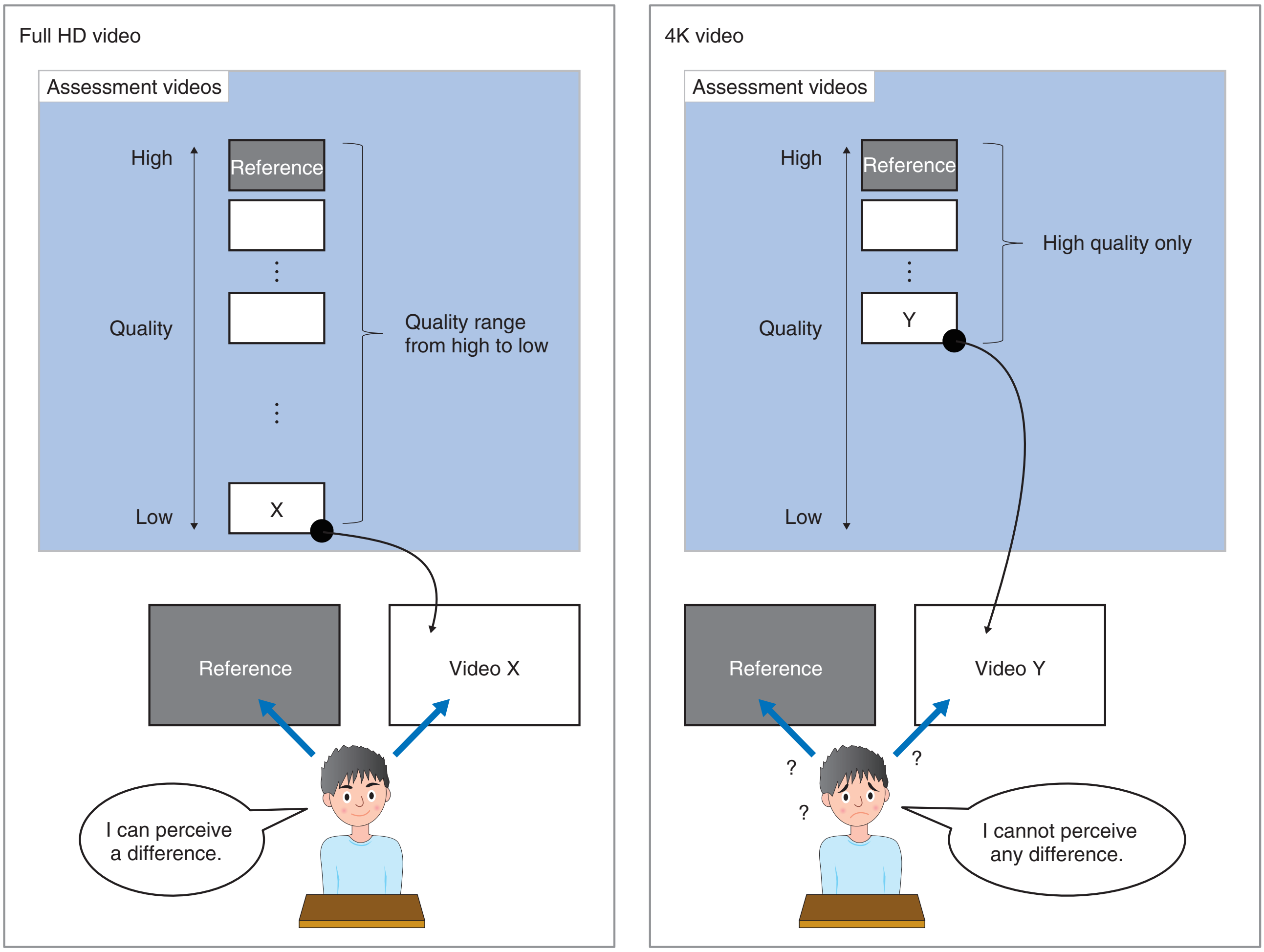


Fig. 4. Concerns about using the DSCQS method with 4K assessment videos.

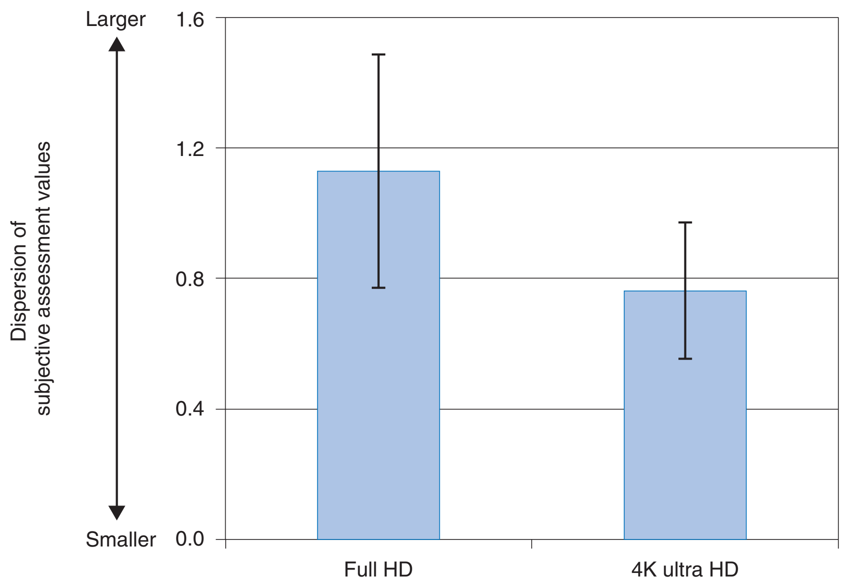
the reference video and 4K assessment videos that were close in quality to the reference video. It is therefore possible to use DSCQS as a metric for 4K video quality in the same way as has been done in the past.

4. Conclusion

We have described here a subjective technique for quality assessment of 4K video. By assessing the quality of 4K video encoded at an assumed bit-rate in 4K video services, we will be able to set the encoding rate to provide 4K video services that satisfies user expectations.

Many attractive video services are possible with 4K video services, for example, remote communication services and interactive video distribution services that allow users to select and view what they want to see. Therefore, to ensure a level of quality that satisfies users, it is important to evaluate the video quality described here as well as other sensory aspects of user experience such as the feeling of engaging in face-to-face communication, the sense of presence that seems to draw users into the proposed 4K video space, the sense of immersion, and an overall positive feeling. Techniques for evaluating such sensory experiences are particularly needed for 8K video services in the future, because viewing on larger monitors is expected.

In the future, we will develop techniques for evaluating a sense of presence and other sensory experiences that are not assessed in the conventional subjective assessment techniques, and we will continue to work on techniques to provide high quality and high sense of presence 4K and 8K video services.



Participants: 36
Viewing distance: 1.5H (4K), 3H (full HD)
Number of assessed video scenes: 4
Number of assessed video conditions: 4 (reference + three versions of encoding conditions)
Error bars: 95% confidence intervals

Fig. 5. Dispersion of subjective assessment values of full HD and 4K video.

References

[1] M. Sugawara, "Advances in Super Hi-vision in 2012—Public viewings at the London Olympics, and standardization at the ITU-R," New Breeze, Winter, pp. 4–9, 2013.

[2] ITU-R Recommendation BT.710, "Subjective Assessment Methods for Image Quality in High-definition Television," Nov. 1998.

[3] ITU-R Recommendation BT.1129, "Subjective Assessment of Standard Definition Digital Television (SDTV) Systems," Feb. 1998.

[4] ITU-R Recommendation BT.500, "Methodology for the Subjective Assessment of the Quality of Television Pictures," Jan. 2012.

[5] N. Narita, M. Kanazawa, and F. Okano, "Optimum Screen Size and Viewing Distance for Viewing Ultra High-Definition and Wide-Screen Images," J. Inst. Image Inform. TV Eng., Vol. 55, No. 5, pp. 773–780, 2001 (in Japanese).

Kimiko Kawashima

Researcher, Service Assessment Group, Communication Traffic & Service Quality Project, NTT Network Technology Laboratories.

She received the B.E. and M.E. degrees in engineering from Keio University, Kanagawa, in 2008 and 2010, respectively. She joined NTT in 2010 and has been engaged in researching the quality assessment of visual communication services. She is currently working on quality assessment of 4K video services. She is a member of the Institute of Information and Communication Engineers (IEICE).

Jun Okamoto

Senior Research Engineer, Supervisor, Service Assessment Group, Communication Traffic & Service Quality Project, NTT Network Technology Laboratories.

He received the B.S. and M.S. degrees in electrical engineering from the Science University of Tokyo in 1994 and 1996, respectively. He joined NTT in 1996 and has been engaged in the quality assessment of visual communication services. He is currently working on developing subjective and objective methods to assess audio and visual communication services. He has been contributing to VQEG (Video Quality Experts Group) since 2004. He received the Telecommunications Advancement Foundation Award in 2009 and the International Telecommunication Union Encouragement Award in Japan in 2010.

Silicon-germanium-silica Monolithic Photonic Integration Platform for Telecommunications Device Applications

Koji Yamada, Tai Tsuchizawa, Hidetaka Nishi, Rai Kou, Tatsurou Hiraki, Kotaro Takeda, Hiroshi Fukuda, Mitsuo Usui, Kota Okazaki, Yasuhiko Ishikawa, Kazumi Wada, and Tsuyoshi Yamamoto

Abstract

We have developed a silicon (Si)-germanium (Ge)-silica monolithic photonic integration platform for telecommunications applications of highly integrated silicon-based photonic devices. On this platform, high-performance silica-based passive devices and compact, high-speed Si/Ge-based active devices are monolithically integrated. We have used this platform to develop a photonic receiver for wavelength-division multiplexing systems, in which a silica-based 16-channel arrayed waveguide grating wavelength filter and Ge photodiodes are monolithically integrated. This Si-Ge-silica photonic integration platform also enables direct integration of electronic circuits.

Keywords: silicon photonics, integrated photonic circuit, photonics electronics convergence

1. Introduction

Telecommunications network systems are now facing an explosive traffic increase with a growth rate of 30% per year [1]. However, such large and rapid growth is not sustainable from economic and environmental viewpoints because within ten years, the energy consumption and capital/operating expenditure for network systems would increase by ten times or more. One of the breakthrough technologies for dealing with this information explosion is high-density photonic integration with ultra-small photonic devices. The smaller the device is, the less power is required. Moreover, integrated photonic functions with redundancy enable flexible and energy-efficient network operation [2]. Silicon photonics, one of the most promising technologies for such high-density photonic integration, enables monolithic integration of photonic devices made of silicon (Si) and germanium (Ge) on a Si wafer. In particular, Si photonic wire waveguides with a micrometer-scale bending radius and a sub-micrometer-scale core enable ultra-high density photonic device integration [3].

The potential of silicon photonics technology in data transmission systems is shown in **Fig. 1**. The compactness of the silicon photonic devices could lead to a 100-fold improvement in both energy efficiency and integration density, compared to those in conventional technology. Although the improvement could be implemented in a wide area of applications ranging from on-chip interconnects to long-distance telecommunications, the hurdles to telecommunications applications are still very high. In telecommunications applications, where long-distance optical

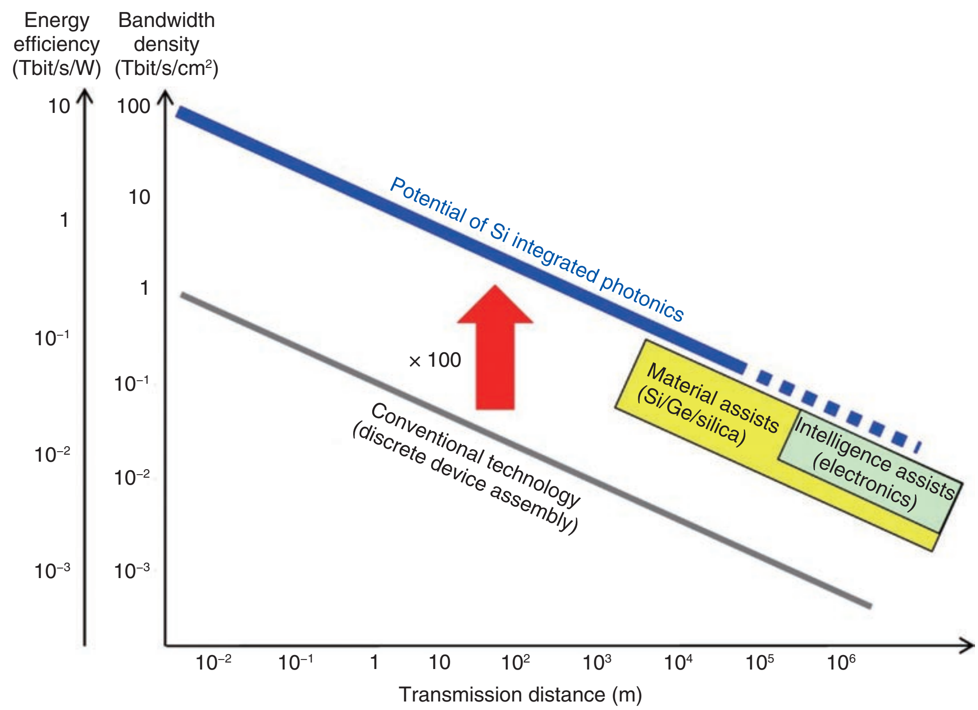


Fig. 1. Potential of silicon photonics in data transmission applications.

transmission is essential, devices should have low optical losses, low interchannel crosstalk, and insensitivity to polarization. For example, in dense wavelength-division multiplexing (WDM) transmission systems, interchannel crosstalk should be less than –25 dB. The present Si photonic platform, however, cannot provide such high-performance photonic devices. In particular, the performance of passive photonic devices, for example, wavelength filters, is a critical issue. Several wavelength filters based on Si photonic wire waveguides have been reported [4], [5], but none of them meet the requirements for telecommunications applications. The effective refractive index of a Si photonic wire waveguide is extremely sensitive to the core geometry, and therefore, the geometric tolerance required is far below that of the errors in the present fabrication technology. For example, a 1-nm geometric error significantly increases the interchannel crosstalk of a wavelength filter to –18 dB, which is not acceptable for telecommunications applications [6].

To eliminate these obstacles to practical telecommunications applications, some assisting technologies should be implemented in Si photonics. These technologies are known as *intelligence assists* and *material assists*, as shown in Fig. 1. Intelligence assists are aimed at compensating for performance deficiencies by using digital electronics, which are now being developed in digital coherent transmission systems. Material assists are designed to improve performance by introducing the right materials for the right functions. One of the most promising material-assist technologies is Si-Ge-silica photonic integration, which can potentially relax the geometric tolerance in high-performance passive devices. We recently developed such a Si-Ge-silica photonic integration technology, by which high-performance silica-based passive devices and Si/Ge-based compact active/dynamic devices can be monolithically integrated. In this article, we report the details of the Si-Ge-silica monolithic integration platform and its application to receivers for telecommunications systems.

2. Si-Ge-silica monolithic photonic integration platform

2.1 Waveguide system

A schematic image of the Si-Ge-silica monolithic integration platform is shown in **Fig. 2**. This platform consists of Si-waveguide-based dynamic/active devices such as Si modulators and Ge detectors,

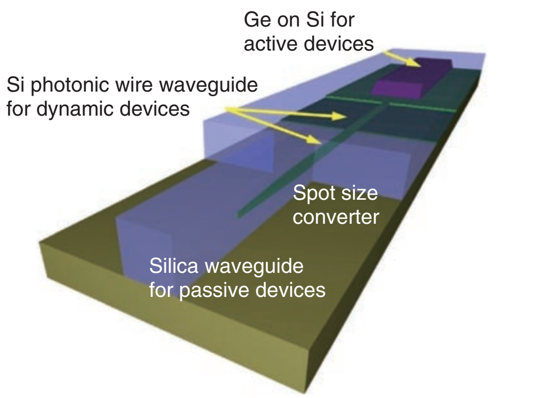


Fig. 2. Schematic of Si-Ge-silica monolithic integration platform.

silica-waveguide-based high-performance passive devices such as wavelength filters, and inverse-taper spot size converters (SSCs), which connect the Si and silica waveguides to each other [7]. These devices are monolithically fabricated on a silicon-on-insulator (SOI) wafer. Integration of Si- and Ge-based devices is possible using conventional Si photonics technology. However, additional silica-based devices cannot be integrated using conventional technology because conventional silica waveguide fabrication generally involves processes with temperatures exceeding 1000ºC for silica film deposition [8]. Such high-temperature processes can destroy the Si/Ge electronic structures for modulators and detectors. Our monolithic integration process features low-temperature silica film deposition using the electron-cyclotron-resonance plasma-enhanced chemical vapor deposition method (ECR-PECVD) [9]. ECR-PECVD involves the deposition of silica-based materials such as silicon oxinitride (SiO_xN_y) and silicon-rich silica (SiO_x). A gas mixture of oxygen (O_2), nitrogen (N_2), and silane (SiH_4) is used for the SiO_xN_y, and O_2 and SiH_4 are used for the SiO_x deposition. The O_2 and N_2 gases are introduced into the plasma chamber, and the SiH_4 gas is introduced into the deposition chamber. ECR plasma generated in the plasma chamber is transported to the deposition chamber by a divergent magnetic field and irradiated to a wafer in the deposition chamber. The energy of the ions irradiated to the wafer is about 10–20 eV [10]. This moderate energy induces a reaction on the wafer surface so that high-quality films are formed at low temperature. The ECR-PECVD system does not add bias to the substrate, so the wafer temperature during the film deposition can be kept below 200ºC even without wafer cooling. The refractive index of the deposited films is controlled by adjusting the flow rate of O_2 and N_2 for SiO_xN_y films and the flow rate of O_2 for SiO_x films while maintaining the flow of SiH_4 at a fixed rate. The refractive index and deposition rate as a function of gas flow rate are shown in **Fig. 3**. Index controllable ranges are 1.47–1.72 for SiO_x and 1.47–1.95 for SiO_xN_y. The deposition rate is over 100 nm/min. The atomic percentages were measured by Rutherford backscattering spectrometry, and we confirmed that the composition of the thermal-oxide film (reference) was precisely measured to be $SiO_{2.0}$ with this technique. The compositions of the SiO_x and SiO_xN_y cores with a refractive index contrast Δ (delta) of 2.9%, and of the clad films were $SiO_{1.7}$, $SiO_{1.8}N_{0.2}$, and $SiO_{2.0}$, respectively. These results indicate that the refractive indices of the silica-based films can be successfully controlled by the atomic ratio of Si/O (N). Here, for the WDM devices, the SiO_x core was used for most of the passive devices because it has lower optical absorption in the C-band than the SiO_xN_y core.

Transmission characteristics of monolithically integrated Si and SiO_x waveguides with a delta of 2.9% are shown in **Fig. 4**. The core size of the Si and SiO_x waveguides are 200 nm × 400 nm and 3 µm × 3 µm, respectively. In the sample preparation, we fabricated the Si waveguide core using electron beam lithography and low-pressure-plasma reactive ion etching before constructing the SiO_x waveguides. The propagation losses of the Si and SiO_x waveguides on the Si/Ge/silica monolithic platform were less than 3 dB/cm and 0.9 dB/cm, respectively. Without Ge, these losses can be further reduced to 1 dB/cm and 0.2 dB/cm, respectively. The polarization-dependent propagation loss of the SiO_x waveguides is less than 0.1 dB/cm [11].

The SiO_x waveguide and Si waveguide can be connected by an SSC using an inverse adiabatic Si waveguide taper. The taper tip of the SSCs is 200 nm × 80 nm, which provides a reflection ratio in the taper region of less than –40 dB [12]. The typical coupling loss per SSC is 0.35 dB for TE (transverse electric)-like modes and 0.31 dB for TM (transverse-magnetic)-like modes, respectively [13]. These coupling losses are low enough to meet the severe telecommunications-grade requirements of the platform.

The transmission spectrum of a Si photonic wire waveguide with an SSC is shown in **Fig. 5**. Although there is absorption due to oxygen-hydrogen (O-H)

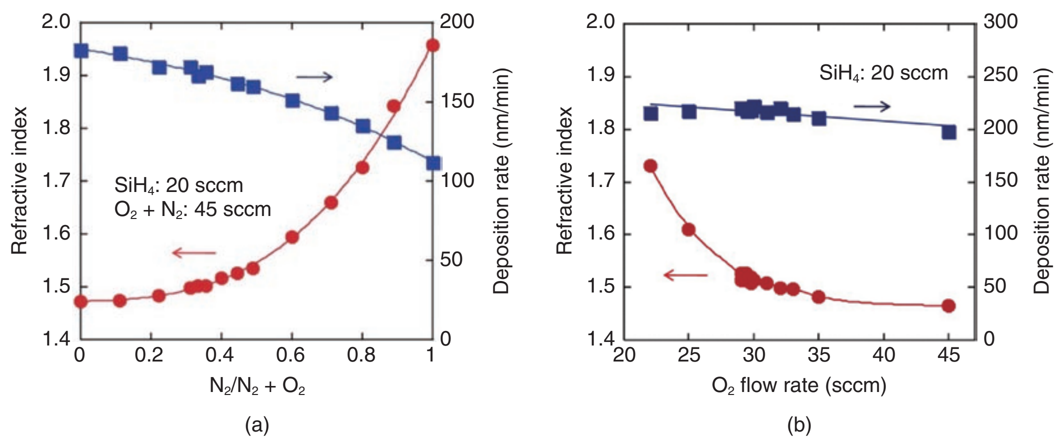
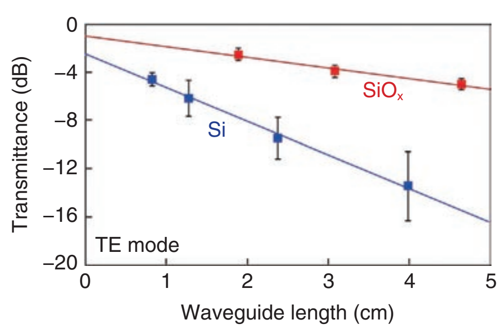


Fig. 3. Refractive index and deposition rate as a function of gas flow rate: (a) SiO_xN_y and (b) SiO_x.



TE: transverse electric

Fig. 4. Transmission characteristics of Si and SiO_x waveguides.

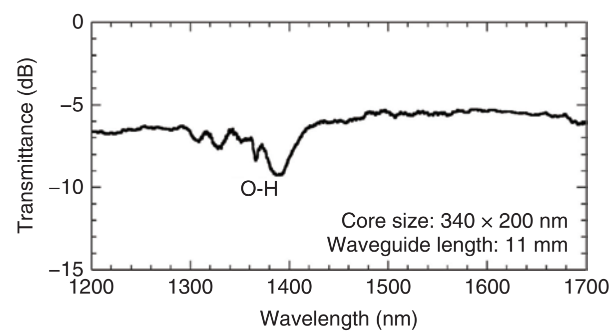


Fig. 5. Transmission spectrum of Si photonic wire waveguide with SSC.

residue in the SiO_x film, the applicable bandwidth of the SSC is very wide, ranging from 1400 to 1700 nm for 1-dB loss. The O-H absorption can easily be eliminated by thermal annealing, and the applicable bandwidth can be further expanded.

2.2 Stand-alone devices

We used the SiO_x waveguides to develope a 12-channel arrayed waveguide grating (AWG) wavelength filter with 1.6-nm channel separation. The structure and filtering spectra of the device are shown in **Fig. 6**. The fiber-to-fiber insertion loss is 5 dB and interchannel crosstalk is less than –25 dB. Thanks to a large waveguide core and moderate index contrast, the filtering performance is better than that in Si-based AWGs. The insertion loss and crosstalk would be further improved by applying sophisticated designs developed for conventional silica waveguide devices. Although small polarization dependence due to film stress remains in the SiO_x AWG, we can compensate for it by introducing a multilayer core technology using SiO_x and silicon nitride films [11].

We also constructed Ge photodetectors on our Si-Ge-silica photonic platform. The structure of a Ge-based photodiode (Ge PD) with a Si waveguide input is shown in **Fig. 7(a)** [13]. The fabrication process for the Ge PD on the Si/Ge/silica platform is much more complicated than that of modulation devices. First, Si waveguide cores and tapers for SSCs are fabricated on the SOI wafer, and boron implantation for p-type electrodes is performed. Next, selective growth of Ge is carried out by using the UHV-CVD (ultrahigh vacuum chemical vapor deposition) method [14], and a thin Si film is deposited on the Ge surface to prevent the Ge film from being damaged during subsequent processes. After that, phosphine is implanted into the top of the Ge PD, followed by the formation of

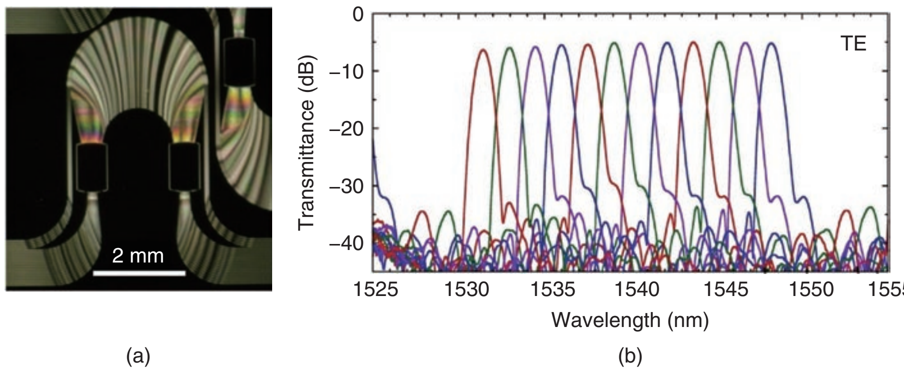


Fig. 6. (a) Structure and (b) filtering spectra of SiO_x AWG.

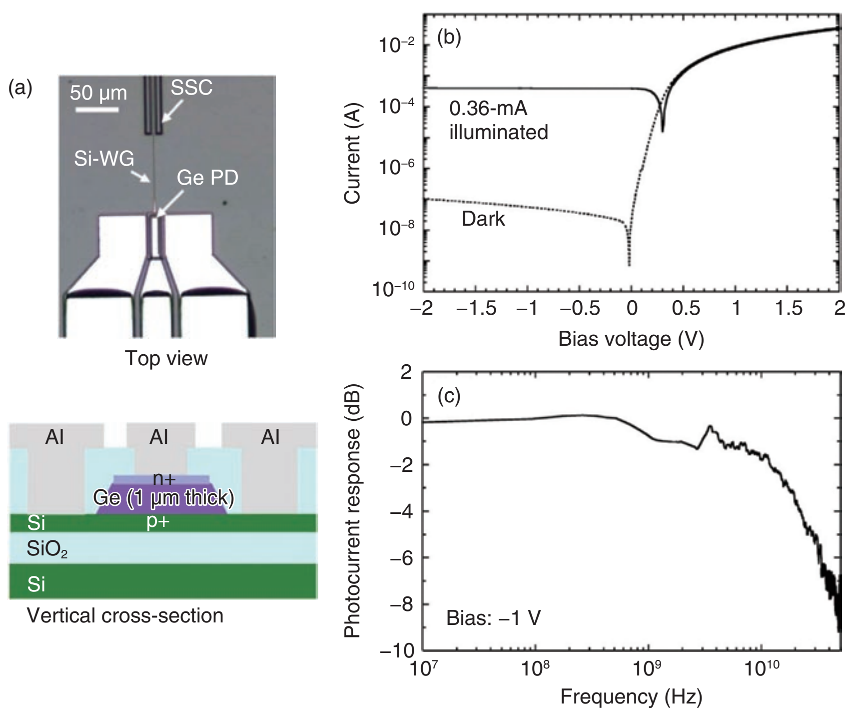


Fig. 7. (a) Structure, (b) current-voltage characteristics, and (c) frequency response of Ge PD.

Ti/TiN/Al electrodes. Then, SiO_x film is deposited by using ECR-PECVD, which is then etched to form cores of SiO_x waveguides. Finally, an over-cladding SiO_2 film is deposited, and contact-holes are formed. The current-voltage curves of a Ge PD in the dark and in an illuminated state are shown in **Fig. 7(b)**. Input light power was 0.36 mW at the SSC before the Ge PD. Dark current and photocurrent were 56 nA and 0.4 mA at a –1-V bias. The responsivity of the stand-alone Ge PD was estimated to be about 1.1 A/W at the entrance of the SSC. The series resistance of the Ge PD was estimated to be about 140 Ω. The low parasitic resistance makes it possible to reduce RC (resistance capacitance) delay and improves the operation speed of the PD. The frequency response of the Ge PD is shown in **Fig. 7(c)**. The 3-dB cutoff frequency is about 20 GHz at a –1-V bias, which would be fast enough for application to 25-Gbit/s data transmission. The relationship between photocurrent and wavelength is shown in **Fig. 8** [13]. The Ge PDs show

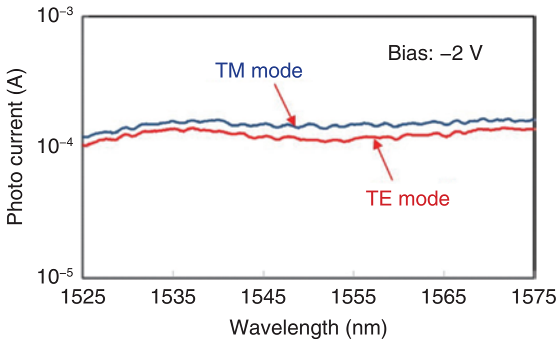


Fig. 8. Dependence on wavelength of photocurrent of Ge PD with TE and TM mode inputs.

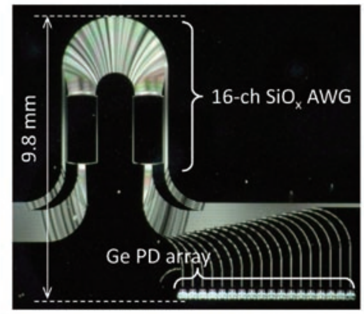


Fig. 9. Structure of AWG-PD.

a flat wavelength dependence in the C-band, and the polarization dependence of the power amplitude is less than 1.0 dB.

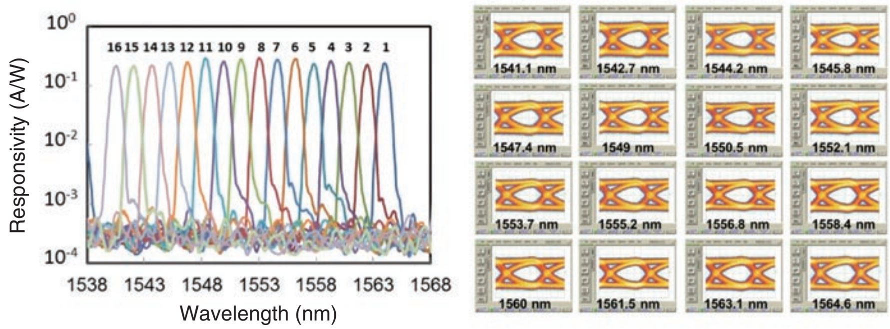
For modulation devices, we developed a polarization-independent Si-based variable optical attenuator with a p-i-n carrier injection structure [15] and Si-based p-n depletion-type optical modulators with a 10-Gbit/s modulation speed.

3. Photonic integration for WDM receivers

For WDM receiver applications, we integrated Ge PDs and an AWG on our Si-Ge-silica photonic platform. The structure of the device is shown in **Fig. 9** [13]. The basic structure of the AWG is the same as that of the stand-alone device, except for the number of channels and the polarization treatment. We adopted a 16-ch AWG in this integration to use in a feasibility study of a 400-Gbit/s WDM receiver system. A slight polarization-dependent wavelength shift (PDλ) of approximately 2.0 nm remained in this AWG because we used a single-layer SiOx core to ascertain the feasibility of the integration process. The Ge PD was the same as that described in the previous section. At each output of the AWG, Ge PDs were connected via SSCs. Measured demultiplexing (DEMUX) spectra for all 16 channels of the Ge PDs are shown in **Fig. 10(a)** [13]. The spectra show a channel spacing of 1.6 nm as designed, and interchannel crosstalk is around –22 dB or less. The insertion loss of the AWG at the central wavelength is about 5.1 dB, which includes waveguide propagation loss of 1.8 dB and diffraction loss of 3.3 dB. The fiber-to-waveguide coupling loss is 0.48 dB/facet. The fiber-to-PD responsivity, which is normalized at the input power before the AWG, is 0.29 A/W at the central wavelength. The monolithic integration of the high-performance SiO$_x$ AWG with Ge PDs contributes to reducing crosstalk and increasing responsivity compared to a Si AWG with Ge PDs [16] and a SiN AWG with Ge PDs [17].

Next, we input non-return-to-zero (NRZ) pseudo-random bit sequence (PRBS) data with a pattern length of 2^{31}-1 into the AWG. Measured eye diagrams of DEMUX photocurrent with 22-Gbit/s signal inputs for all channels are shown in **Fig. 10(b)** [13]. We confirmed that all 16 channels show clear eye openings. The AWG-PD chip achieved a total system bandwidth of 352 Gbit/s. Here, the bit rate was limited to 22 Gbit/s by the measurement equipment. Since the Ge PDs can work at 25 Gbit/s as mentioned in the previous section, the total system bandwidth could potentially reach 400 Gbit/s.

We also examined long-distance transmission in order to confirm the feasibility of the device for practical network applications. Continuous-wave infrared light from a tunable laser diode was modulated by a lithium-niobate Mach-Zehnder modulator. The modulation format was simple intensity modulation with 12.5-Gbit/s NRZ PRBS data. The word length here was also 2^{31}-1. The signal bandwidth was restricted to 12.5 Gbit/s by our experimental equipment. The modulated optical signal was transmitted through standard single-mode fibers with dispersion of 17 ps/km/nm at distances of 20, 40, and 60 km. After the polarization was prepared, the signal was input into the AWG-PD device by butt coupling of a high-NA (numerical aperture) fiber. The experiment was performed at room temperature without temperature control. The electrical signals from Ge PDs were directly guided to an oscilloscope or an error detector for bit-error rate (BER) measurement without transimpedance amplifiers (TIAs). We used a commercial



(a) (b)

Fig. 10. (a) DEMUX spectra of the AWG-PD chip and (b) eye diagrams of DEMUX photocurrent with 22-Gbit/s signal for all channels.

clock-data recovery (CDR) module to emulate a real receiver. The input light was tapped by a 10-dB coupler and guided into the CDR module.

The measured BER curves obtained from channel 12 for various transmission distances, and a received eye diagram after 40-km transmission (inset), are shown in **Fig. 11** [13]. The channel numbers are shown in Fig. 10(a). We confirmed error-free signal transmission up to 40 km. The receiver sensitivity measured at the AWG input was –6.8 dBm for the BER of 10^{-9} after 40-km transmission. After 60-km transmission, the BER increased considerably because of waveform distortion due to the dispersion and nonlinearity of the fiber. We estimated the sensitivity of the Ge PD on channel 12 to be –13.6 dBm for the 40-km transmission, which takes into account the transmission loss of the AWG. Although the theoretical sensitivity of the AWG-PD device was –32 dBm, which was estimated from shot noise and Johnson noise [18], the measured sensitivity was restricted by the measurement system in that there were no TIAs on this chip. We believe the sensitivity will be improved by on-chip integration of TIAs.

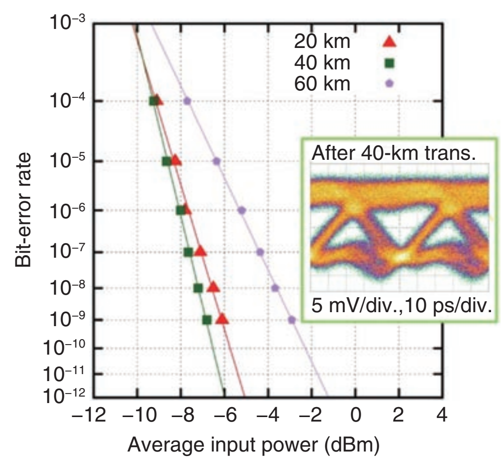
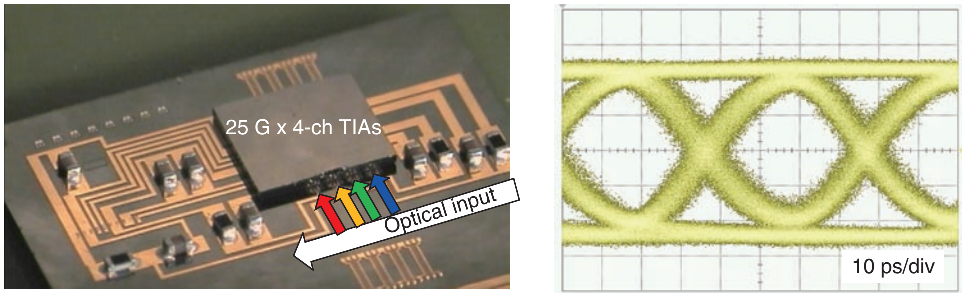


Fig. 11. BER curves of channel 12 of the AWG-PD for various transmission distances and eye diagram after 40-km transmission (inset).

4. Integration of electronics on photonic chip

As mentioned in the previous section, it is very important to use TIAs to improve the receiver sensitivity. Generally, electrical wiring between a TIA and PD should be as short as possible, typically a few hundred micrometers or less for 25-Gbit/s signals. Conventionally, wire bonding technology is used to connect a TIA to a PD. In the wire bonding connection, however, it is difficult to shorten the wiring, and the frequency response is therefore easily degraded by parasitic impedances. The wiring should also be shortened for other high-speed electronics for photonic interfaces, for example, modulator drivers. In other words, the electronic components should be located as close as possible to the photonic ones. From this point of view, the electronics should be located on photonic chips. For this photonics-electronics convergence, electronics back-end processes should be performed on photonic chips, upon which the electronic wiring should be constructed and electronic chips mounted. The Si-based photonic platform is suitable for this kind of photonics-electronics convergence because it is very robust for electronics



(a) (b)

Fig. 12. (a) Multi-channel TIA integration on a photonic chip with 4-ch optical DEMUX and Ge PDs and (b) typical eye diagram of the TIA.

back-end processes.

An example of photonics-electronics convergence for a WDM receiver system is shown in **Fig. 12(a)**. Four-channel Si-based ring resonator wavelength filters and Ge PDs are monolithically integrated on a photonic chip. This photonic chip contains electronic wiring that includes transmission lines, capacitors, and resistors, as well as a four-channel TIA mounted using flip-chip bonding technology. The frequency response of the wiring between the TIA and PD is flat throughout a 40-GHz range. Thus, we obtained a very clear eye pattern of the TIA output for 25-Gbit/s signals, as shown in **Fig. 12(b)**.

5. Summary

We reported a Si-Ge-silica monolithic photonic integration platform for telecommunications applications. This platform enables monolithic integration of Si/Ge-based compact modulators/detectors and silica-based high-performance wavelength filters. The monolithic integration process features low-temperature silica film deposition by ECR PECVD so as not to damage the Si/Ge active devices. Since the right materials are applied for the right functions, the Si-Ge-silica platform can provide well-balanced performance for telecommunications applications, whose specifications are too strict for the conventional silicon-based photonic platform.

We have developed various integrated photonic devices for broadband WDM applications using this platform. An integrated AWG-PD device works very well as a multichannel WDM receiver with a 400-Gbit/s transmission capacity. High-speed electronics can also be integrated on this Si-Ge-silica photonic platform.

Thus, with the Si-Ge-silica monolithic photonic integration platform, we can construct a highly functional WDM photonic module with an electronic interface on one chip, which can be applied in compact, energy-efficient, and cost-effective photonic network systems.

References

[1] Cisco Global Cloud Index: Forecast and Methodology, 2012–2017. http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns1175/Cloud_Index_White_Paper.html

[2] H. Kimura, N. Iiyama, Y. Sakai, and K. Kumozaki, "A WDM-based Future Optical Access Network and Support Technologies for Adapting the User Demands' Diversity," IEICE Trans. Commun., E93-B, pp. 246–254, 2010.

[3] Y. Urino, Y. Noguchi, M. Noguchi, M. Imai, M. Yamagishi, S. Saitou, N. Hirayama, M. Takahashi, H. Takahashi, E. Saito, T. Shimizu, M. Okano, N. Hatori, M. Ishizaka, T. Yamamoto, T. Baba, T. Akagawa, S. Akiyama, T. Usuki, D. Okamoto, M. Miura, J. Fujikata, D. Shimura, H. Okayama, H. Yaegashi, T. Tsuchizawa, K. Yamada, M. Mori, T. Horikawa, T. Nakamura, and Y. Arakawa, "Demonstration of 12.5-Gbps Optical Interconnects Integrated with Lasers, Optical Splitters, Optical Modulators and Photodetectors on a Single Silicon Substrate," Optics Express, Vol. 20, No. 26, pp. B257–B263, 2012.

[4] W. Bogaerts, P. Dumon, D. V. Thourhout, D. Taillaert, P. Jaenen, J. Wouters, S. Beckx, V. Wiaux, and R. G. Baets, "Compact Wavelength-selective Functions in Silicon-on-Insulator Photonic Wires," IEEE J. Select. Topics Quant. Electron., Vol. 12, No. 6, pp. 1394–1401, 2006.

[5] Q. Fang, T-Y. Liow, J. F. Song, K. W. Ang, M. B. Yu, G. Q. Lo, and D-L. Kwong, "WDM Multi-channel Silicon Photonic Receiver with 320 Gbps Data Transmission Capability," Optics Express, Vol. 18, No. 5, pp. 5106–5113, 2010.

[6] K. Okamoto, "Progress and Technical Challenge for Planar Waveguide Devices: Silica and Silicon Waveguides," Laser & Photonics Review, Vol. 6, No. 1, pp. 14–23, 2012.

[7] T. Tsuchizawa, K. Yamada, H. Fukuda, T. Watanabe, J. Takahashi, M. Takahashi, T. Shoji, E. Tamechika, S. Itabashi, and H. Morita, "Microphotonics Devices Based on Silicon Microfabrication Technology," IEEE J. Select. Topics Quant. Electron., Vol. 11, No. 1, pp. 232–240, 2005.

[8] M. Kawachi, "Silica Waveguides on Silicon and Their Applications to Integrated-optic Components," Opt. Quantum Electron., Vol. 22, No. 5, pp. 391–416, 1990.

[9] T. Tsuchizawa, "Guided Light in Silicon-based Materials," Handbook of Silicon Photonics, Ed. by L. Vivien and L. Pavesi, CRC Press, 2013.
[10] S. Matsuo and M. Kiuchi, "Low Temperature Chemical Vapor Deposition Method Utilizing an Electron Cyclotron Resonance Plasma," Jpn. J. Appl. Phys., Vol. 22, pp. L210–L212, 1983.
[11] H. Nishi, T. Tsuchizawa, H. Shinojima, T. Watanabe, S. Itabashi, R. Kou, H. Fukuda, and K. Yamada, "Low Polarization-dependent Silica Waveguide Monolithically Integrated on SOI Photonic Platform," IEEE Journal of Lightwave Technology, Vol. 31, pp. 1821–1827, 2013.
[12] T. Tsuchizawa, K. Yamada, T. Watanabe, S. Park, H. Nishi, R. Kou, H. Shinojima, and S. Itabashi, "Monolithic Integration of Silicon-, Germanium- and Silica-based Optical Devices for Telecommunications Applications," IEEE Journal of Selected Topics in Quantum Electronics, Vol. 17, No. 3, pp. 516–525, 2011.
[13] T. Hiraki, H. Nishi, T. Tsuchizawa, R. Kou, H. Fukuda, K. Takeda, Y. Ishikawa, K. Wada, and K. Yamada, "Si-Ge-Silica Monolithic Integration Platform and Its Application to a 22-Gb/s × 16-ch WDM Receiver," IEEE Photonics Journal, Vol. 5, No. 4, 4500407, 2013.
[14] H-C. Luan, D. R. Lim, K. K. Lee, K. M. Chen, J. G. Sandland, K. Wada, and L. C. Kimerling, "High-quality Ge Epilayers on Si with Low Threading-dislocation Densities," Applied Physics Letters, Vol. 75, No. 19, pp. 2909–2911, 1999.
[15] H. Nishi, T. Tsuchizawa, T. Watanabe, H. Shinojima, K. Yamada, and S. Itabashi, "Compact and Polarization-independent Variable Optical Attenuator Based on a Silicon Wire Waveguide with a Carrier Injection Structure," Japanese Journal of Applied Physics, Vol. 49, 04DG20, 2010.
[16] J. M. Fedeli, L. Virot, J. M. Hartmann, P. Grosse, W. Bogaerts, and L. Vivien, "16 Channel Receiver with 20 GHz Ge Photodiodes," Proc. of IEEE International Conference on Group IV Photonics, Vol. 1, No. 1, pp. 156–158, San Diego, CA, USA, 2012.
[17] L. C. Chen, C. R. Doerr, L. Buhl, Y. Baeyens, and R. A. Aroca, "Monolithically Integrated 40-Wavelength Demultiplexer and Photodetector Array on Silicon," IEEE Photonics Technology Letters, Vol. 23, No. 13, pp. 869–871, 2011.
[18] G. P. Agrawal, "Fiber-Optic Communication Sytems (Wiley Series in Microwave and Optical Engineering)," 2nd Edition, Wiley-Interscience, NewYork.

Koji Yamada
Senior Research Engineer, Supervisor, Distinguished Technical Member of the NTT NanoPhotonics Center and NTT Microsystem Integration Laboratories.

He received the B.E., M.E., and Ph.D. degrees in nuclear engineering from Kyushu University, Fukuoka, in 1986, 1988, and 2003, respectively. He joined NTT in 1988 and began studying electron beam dynamics in synchrotron light sources. He has been engaged in studies on silicon photonic platforms since 2000, and he has also been a leader of the photonic wiring and waveguide research team of the PECST project supported by the Japan Society for the Promotion of Science since 2010. He is a member of the Japan Society of Applied Physics (JSAP), the Atomic Energy Society of Japan, the Particle Accelerator Society of Japan, the Institute of Electronics, Information and Communication Engineers (IEICE), and IEEE.

Tai Tsuchizawa
Senior Research Engineer, NTT Microsystem Integration Laboratories.

He received the B.S. and M.S. degrees in physics from Sophia University, Tokyo, in 1984 and 1986, respectively, and the Ph.D. degree from The University of Tokyo in 1990. He has been involved in studies on plasma processes for microfabrication. His current research interest includes fabrication technologies for silicon-based optoelectronics devices. He is a member of JSAP.

Hidetaka Nishi
Research Engineer, NTT Microsystem Integration Laboratories and NTT Nanophotonics Center.

He received the B.S. and M.S. degrees in mechanical science and engineering from Tokyo Institute of Technology in 2005 and 2007, respectively. Since joining NTT Microsystem Integration Laboratories in 2007, he has been engaged in research on integrated silicon photonic devices. He is a member of JSAP and the Optical Society of America.

Rai Kou
NTT Microsystem Integration Laboratories and NTT Nanophotonics Center.

He received the B.E. and M.E. degrees in applied physics from Waseda University, Tokyo, in 2007 and 2009, respectively. From 2007 to 2009, he was a Visiting Research Student at the National Institute for Materials Science (NIMS), Tsukuba, Japan, where he was engaged in research on nonlinear optics. He has been at NTT Microsystem Integration Laboratories since 2009 and NTT Nanophotonics Center since 2012 (concurrent posts). His current research interests include integrated photonic circuits with silicon and its related materials. He is a member of JSAP, IEICE, and IEEE.

Tatsuro Hiraki

NTT Microsystem Integration Laboratories and NTT Nanophotonics Center.

He received the B.E. and M.E. degrees in mechanical engineering from Tohoku University, Miyagi, in 2009 and 2011, respectively. He has been at NTT Microsystem Integration Laboratories and NTT Nanophotonics Center since 2012. His current research interests include integrated silicon photonic devices. He is a member of JSAP.

Kotaro Takeda

NTT Microsystem Integration Laboratories and NTT Nanophotonics Center.

He received the B.E., M.E., and Ph.D. degrees in materials science from Tokyo Institute of Technology in 2007, 2009, and 2012, respectively. He has been at NTT Microsystem Integration Laboratories and NTT Nanophotonics Center since 2012. His current research interests include integrated silicon photonic devices. He is a member of JSAP.

Hiroshi Fukuda

Senior Research Engineer, NTT Microsystem Integration Laboratories.

He received the B.E. and M.E. degrees in nuclear engineering from Tohoku University, Miyagi, in 1993 and 1995, respectively. In 1995, he joined NTT Microsystem Integration Laboratories, where he has been engaged in research on microphotonics devices. He is a member of JSAP, IEICE, and IEEE.

Mitsuo Usui

Senior Research Engineer, NTT Microsystem Integration Laboratories.

He received the B.E. and M.E. degrees in electronic engineering from Tokyo Denki University in 1988 and 1990, respectively. He joined NTT in 1990 and has been engaged in researching packaging technology for electronic and optoelectronic components. He is a member of JSAP, IEICE, and the Japan Institute of Electronics Packaging (JIEP). He received the Young Engineer Award from IEICE in 1997, and the Best Paper Award and the Inose Award from IEICE in 2000.

Kota Okazaki

NTT Microsystem Integration Laboratories.

He received the B.S. and M.S. degrees in quantum radiation sciences and the Ph.D. degree in electrical and electronic science from Kyushu University, Fukuoka, in 2008, 2010, and 2013, respectively. In 2013, he joined NTT Microsystem Integration Laboratories, where he has been researching plasma process microfabrication for microphotonics devices with silicon and its related materials. He is a member of JSAP.

Yasuhiko Ishikawa

Associate Professor, Department of Materials Engineering, The University of Tokyo.

He received the B.S., M.S., and Ph.D. degrees from Hokkaido University in 1993, 1995, and 1998, respectively. During 2001–2002, he was a Visiting Scientist at the Materials Processing Center, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. His current interests are Si-based photonic devices and their applications to communications and biosensing systems. He is a member of JSAP and the Materials Research Society.

Kazumi Wada

Professor, Department of Materials Engineering, The University of Tokyo.

He received the Ph.D. degree in instrumentation engineering from Keio University, Kanagawa, in 1982. He has been a Professor in the Department of Materials Engineering at The University of Tokyo since 2004. In 1998, he was a principle research scientist at the Materials Processing Center of MIT and conducted research on Ge photodetectors integrated on a Si chip. He was with NTT LSI Laboratories for over 20 years (since 1975), where he was engaged in research on defects in silicon and silicon-based photonic integrated circuits. He served as chair of Group IV Photonics 2007, sponsored by IEEE, and was on the Board of Directors of the Materials Research Society (MRS), and on the Board of the Japan Society of Applied Physics. He runs and was a co-proposer of the Japanese Programs on Si photonics, including the FIRST program, referred to as Photonics-Electronics Convergence System Technology.

Tsuyoshi Yamamoto

Senior Research Engineer, Supervisor, Group Leader, NTT Microsystem Integration Laboratories.

He received the B.E. degree in electrical engineering from Kansai University, Osaka, in 1991. In 1991, he joined NTT Communication Switching Laboratories, where he was engaged in R&D of several optical interconnection systems using free-space optics. Recently, he has been involved in R&D of photonic-integrated-circuit devices based on silicon photonics technology. During 1998–1999, he was a Visiting Research Engineer at the Department of Electrical and Computer Engineering, McGill University, QC, Canada. During 2007–2011, he was the Director of the Optical Node System Department, Optical Communications Systems Group at NTT Electronics Corp. He received the Microoptics Conference Best Paper Awards in 2003 and 2009.

Advanced Technologies for Wide-area Ethernet Networks

Tomonori Shibuya, Takahiro Madokoro, Yuki Takagi, and Yoshiaki Sone

Abstract

This article introduces a partial-bandwidth-guaranteed service system and a service-area expansion system that we have developed in order to improve the performance of wide-area Ethernet networks. The developed systems will enable new added services and provide a scheme for cost-efficient network expansion.

Keywords: wide-area Ethernet service, partial-bandwidth-guaranteed service system, service-area expansion system

1. Wide-area Ethernet

Wide-area Ethernet services have been commercially available since 2008 and have been providing high-speed circuit services at a reasonable cost compared to conventional leased-line services. NTT Access Network Service Systems Laboratories is doing continuous research and development (R&D) in areas such as L2 (Layer 2) networking technology and operating technology, with the aim of further advancing wide-area Ethernet networks [1].

Two requirements must be met in order to improve the existing NGN (Next Generation Network) Ethernet networks.

- Offering various services that supplement data-center/cloud needs
- Sharply reducing network capital expenditure

This article introduces a *partial-bandwidth-guaranteed service system* and a *service-area expansion system* that address these requirements by improving the access systems of wide-area Ethernet networks (**Fig. 1**).

2. Partial-bandwidth-guaranteed service system

We have developed a partial-bandwidth-guaranteed service system that satisfies various customer needs (**Fig. 2**). Conventionally, wide-area Ethernet networks offer only bandwidth-guaranteed services. Therefore, there is unutilized bandwidth when the total contracted bandwidth is less than the physical bandwidth capacity of the network, or when the actual working bandwidth for forwarding user traffic is less than the contracted bandwidth. To efficiently utilize this unutilized bandwidth, we have developed a new system that achieves the transfer of user traffic as bandwidth-guaranteed traffic, as long as the user traffic does not exceed the predetermined guaranteed bandwidth. In addition, it can transfer the user traffic flexibly with the unutilized bandwidth even if user traffic exceeds the predetermined guaranteed-bandwidth limit. The system has the following three key functions.

(1) Bandwidth monitoring function

This function supervises the flow of user traffic for every frame in order to determine whether to transmit traffic as guaranteed traffic or best-effort traffic. An optical subscriber unit (OSU) transfers the user frames that arrive under the designated committed information rate (CIR). User frames that exceed the CIR but not the peak information rate (PIR) are transferred by the OSU as best-effort traffic. If user frames exceed the PIR, the OSU discards them.

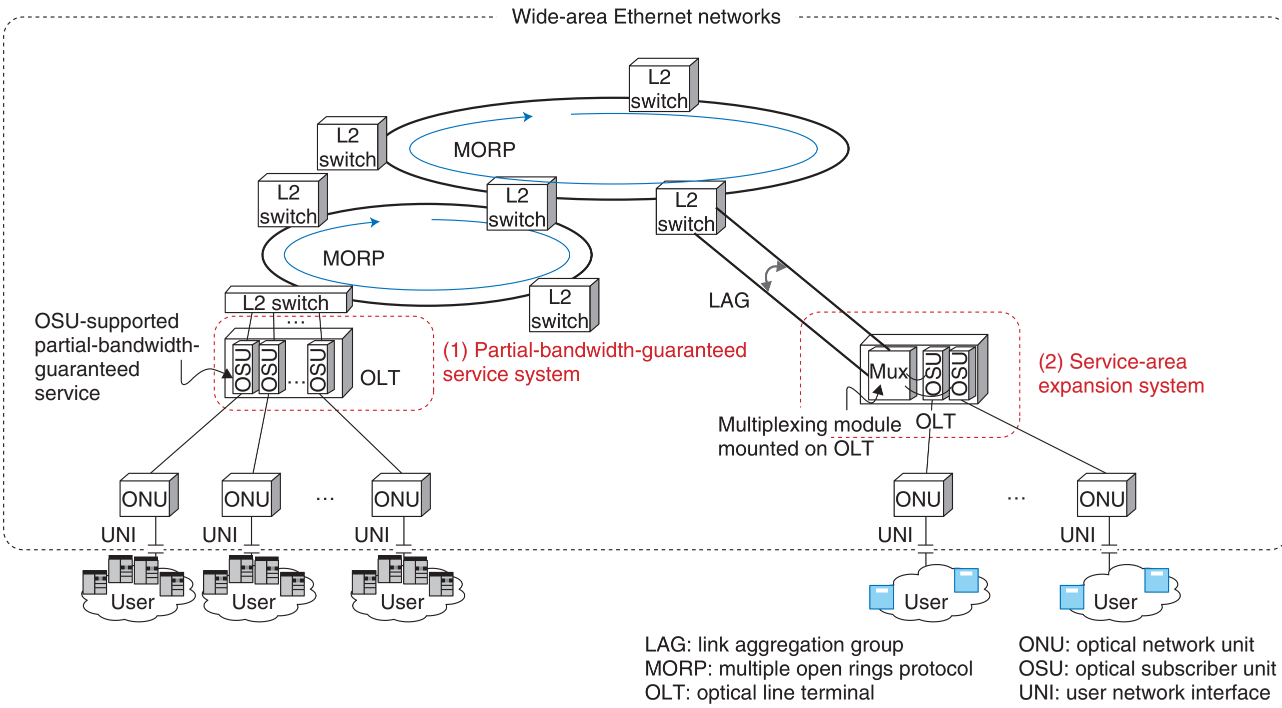


Fig. 1. Advanced technologies for wide-area Ethernet networks.

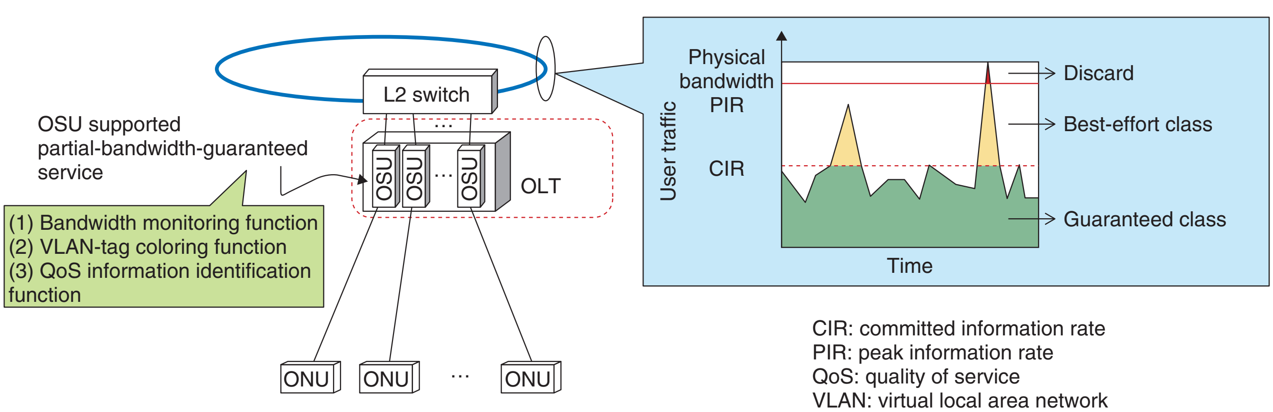


Fig. 2. Partial-bandwidth-guaranteed service system.

(2) VLAN tag coloring function

This function adds colored virtual local area network (VLAN) tags to user frames so that L2 switches can determine whether to use guaranteed class or best-effort class. The L2 switches refer to a class-identifier in the tag to execute the frame quality of service (QoS) control according to the congestion state.

(3) QoS information identification function

In the developed system, an access system determines the QoS class of the frame by checking the various designated priority fields in the user frames, which are arbitrarily set up by the user's local policy, so that core switches can conduct QoS control according to the core network policy. Specifically, the OSU in the access system determines the color of VLAN tags according to the value of designated

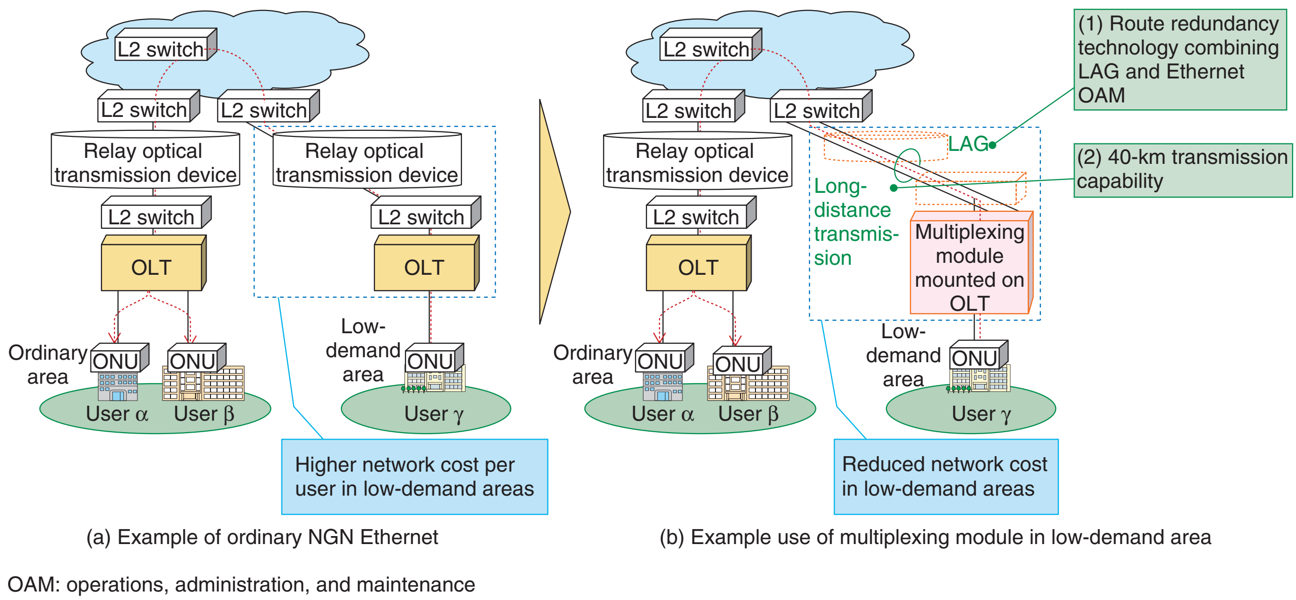


Fig. 3. Service-area expansion system.

fields such as ToS (type of service)/TC (traffic class) for IP (Internet protocol) headers and CoS (class of service)/VID (VLAN identifier) for VLAN tags [2].

A wide-area Ethernet network that applies these functions can transfer even bursty traffic flexibly and efficiently according to the available bandwidth capacity in the network while keeping the conventional guaranteed transfer function.

3. Service-area expansion system

To provide wide-area Ethernet service, it is necessary to establish an access network using several kinds of network equipment, for example, a media converter (MC), L2 multiplexing switches, and long-distance transmission systems. Conventionally, these multiple types of equipment have been applied to all service areas. However, in low-demand areas, the equipment cost per user becomes larger than in other areas (**Fig. 3(a)**). We have been working to reduce this cost and have developed a new multiplexing module for MCs that has a route redundancy capability and long-distance transmission capability. This module features a conventional multiplexing switching function that is integrated with a long-distance transmission function. This module adds VLAN tags to received user frames and multiplexes multiple access lines into one of the redundant trunk lines. This module employs the following two key technologies.

(1) Route redundancy technology
The system's reliability is enhanced through the use of a novel route redundancy technology we have developed that combines LAG (link aggregation group) and Ethernet OAM (operation, administration, and maintenance) [3] ((1) in **Fig. 3(b)**). The active route is switched if the system detects a link down alarm in the physical layer or a frame loss in the data-link layer.

(2) Long-distance transmission technology
This system makes use of a long-distance transmission module (ZX-SFP (small form-factor pluggable)) that has 40-km transmission capability over single-mode optical fiber ((2) in Fig. 3(b)). In addition, the system supports 1000BASE-SX (up to 550 m) and 1000BASE-LX (up to 5 km), which are standardized in IEEE 802.3, so that the operator can select the optimum SFP module depending on the distance between the L2 switch and the OLT.

4. Summary

We have developed a partial-bandwidth-guaranteed service system and a service-area expansion system that promise to improve wide-area Ethernet services.

These systems enable new added services and provide a scheme for cost-efficient network expansion. We believe these systems will open a path to further expanding the use of wide-area Ethernet services.

References

[1] S. Narikawa, M. Maruyoshi, and K. Okura, "Carrier Ethernet Technologies for the NGN Era," NTT Technical Review, Vol. 7, No. 11, 2009.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr200911sf6.html

[2] IEEE 802.1ac.
http://www.ieee802.org/1/pages/802.1ac.html

[3] "Special Feature: NGN Communication Styles and Support Technologies," NTT Technical Review, Vol. 5, No. 6, 2007.
https://www.ntt-review.jp/archive/2007/200706.html

Tomonori Shibuya

Senior Research Engineer, First Promotion Project, NTT Access Network Service Systems Laboratories.

He received the B.E. and M.E. degrees in electronics and communication engineering from Meiji University, Tokyo, in 1998 and 2000, respectively, the MBA degree from Tokyo Metropolitan University in 2009, the M.Env. Sci. degree from Tohoku University, Miyagi, in 2011, and the M.A. degree from Kochi University of Technology in 2014. He joined NTT Access Network Service Systems Laboratories in 2000. He worked at NTT EAST R&D Center during 2004–2008. Since 2008, he has been engaged in R&D of next-generation Ethernet technology.

Takahiro Madokoro

Engineer, First Promotion Project, NTT Access Network Service Systems Laboratories.

He received the B.E. and M.E. degrees in electrical engineering and computer science from Nagoya University, Aichi, in 2005 and 2007, respectively. After joining NTT WEST in 2007, he worked on service development for the high-speed optical fiber Internet connection service FLET'S HIKARI Next. Since 2012, he has been engaged in R&D of next-generation Ethernet technology.

Yuki Takagi

Engineer, First Promotion Project, NTT Access Network Service Systems Laboratories.

He received the B.E. degree in mechanical engineering from Osaka University in 2009. In 2011, he joined NTT Access Network Service Systems Laboratories, where he has been engaged in R&D of next-generation Ethernet technology.

Yoshiaki Sone

Research Engineer, First Promotion Project, NTT Access Network Service Systems Laboratories.

He received the B.E. and M.E. degrees in electronics engineering from Tohoku University, Miyagi, in 2001 and 2003, respectively. In 2003, he joined NTT Network Innovation Laboratories and engaged in research on network engineering technology for optical transport networks. He is currently engaged in the development of systems for wide-area Ethernet services in NTT Access Network Service Systems Laboratories. He is a member of the Institute of Electronics, Information and Communication Engineers.

Invisible Watermarks Connecting Video and Information: Mobile Video Watermarking Technology

Ken Tsutsuguchi, Shingo Ando, Atsushi Katayama, Hidenori Tanaka, Susumu Yamamoto, and Yukinobu Taniguchi

Abstract

Mobile video watermarking technology is a digital watermarking technology that can detect invisible information embedded in external videos with both high speed and accuracy, simply by directing the camera in the mobile device towards the video. In this article, we give an overview of the technology and describe two example use case applications in existing broadcast TV services and a new opportunity using video synchronized augmented reality.

Keywords: digital watermark, second screen services, video synchronized AR

1. Second screen service using video watermarks

With the explosive growth of mobile devices such as smartphones and tablets in recent years, services that present information related to television (TV) programs, commercials, events, and digital signage that can be retrieved from videos are becoming more and more well known.

In particular, expectations for a *second screen service* or *multi-screen service* have been growing, and extensive trials are being carried out. This type of service typically treats a TV and a mobile device as a first and second screen, respectively, with service being provided by linking information from both screens.

In our trials, applications installed on smartphones automatically recognized the sounds or images of a TV program or commercial being broadcast. That information can then be applied to direct users towards social networking services and retailers, or to give reward points and coupons to users depending on the service.

The difference between these services and existing technology is that broadcasters can be actively involved in the consumer's experience and can provide a new and engaging style of watching TV with mobile phone in hand.

These services use automatic content recognition (ACR) technology*1.

An example of a second screen service using video watermarking technology is shown in **Fig. 1**.

In the remainder of this article we explain typical methods and application examples of ACR technology and introduce an overview and use cases of the mobile video watermarking technology that has been researched and developed by NTT Media Intelligence Laboratories.

2. ACR technology

ACR technology can be classified by method into four categories as follows:

*1 ACR Technology: we define this term as technology that automatically recognizes and identifies published visual media (e.g., TV programs and commercials).

Fig. 1. Second screen services using digital watermarking technology.

(1) Fingerprints: a recognition and identification technique that matches a query signal with features extracted from content (called fingerprints) that are registered to a database in advance. The robust media search technology [1] developed by NTT Communication Science Laboratories is an example of this method.
(2) Digital watermarking: a technique by which information such as an identification (ID) tag is embedded in content by making otherwise negligible changes that are subject to recognition and identification at a later point. Our mobile video watermarking technology is an example of this.
(3) Visible codes: a technique that presents the ID in the form of a QR (quick response) code or bar code that can be clearly perceived in the content.
(4) Metadata: a technique that imparts information in conjunction with the content in a way that a computer can understand. The hybrid cast technique falls into this category.

There are advantages and disadvantages to each method. For example, the fingerprint technique is advantageous in that it does not modify the original content, but content features must be registered in advance for successful recognition.

The digital watermarking technique can distinguish the same content in different settings (such as TV channels) by embedding different information in each setting. However, it requires preprocessing the original content prior to broadcasting.

These methods have diverse functionality, and consequently, the most effective service may be provided by combining several techniques simultaneously.

3. Mobile video watermarking technology

The mobile video watermarking technology developed by NTT Media Intelligence Laboratories is a digital watermarking technology for video content that is capable of high-speed detection by cell phones and smartphones. The primary focus of our research has been to develop inter-media synchronization tools for mobile devices. Simply by directing the camera of a mobile device installed with the detection application to video content with embedded watermarks, the user can rapidly obtain the watermark ID and can access related information.

This technology uses a high-speed quadrilateral tracking side trace algorithm [2] to detect the video area from camera images (resulting in the embedded digital watermark partially consisting of a thin frame that surrounds the image area). Then, it detects the watermark at any time point using the single-frequency plane spread spectrum technique [3], which is resistant to variance in synchronization time.

An overview of the features and technology of the mobile video watermarking technology is shown in **Fig. 2**.

In our current implementation, we can embed a 29-bit-wide ID number as watermark information. This enables us to categorize and identify content in approximately 500 million ways. In addition, this provides a reliable quantitative guarantee that the ID false positive rate is below a reasonable probability.

The nature of this technique means that slight changes are inevitably made to the original content. However, the degradation of image quality is so negligible that it does not affect content viewing. In addition, it has a very high watermark detection speed. Successful identification typically occurs within one second after the camera is directed at the video, so we can provide a high level of convenience to the user.

4. Demonstration experiment using real-world TV broadcast

In order to provide second screen services based on commercials and TV programs, we first must verify the digital watermark detection performance in a real world broadcast environment. In addition, this allows us to verify the ease of usability of our tool by the

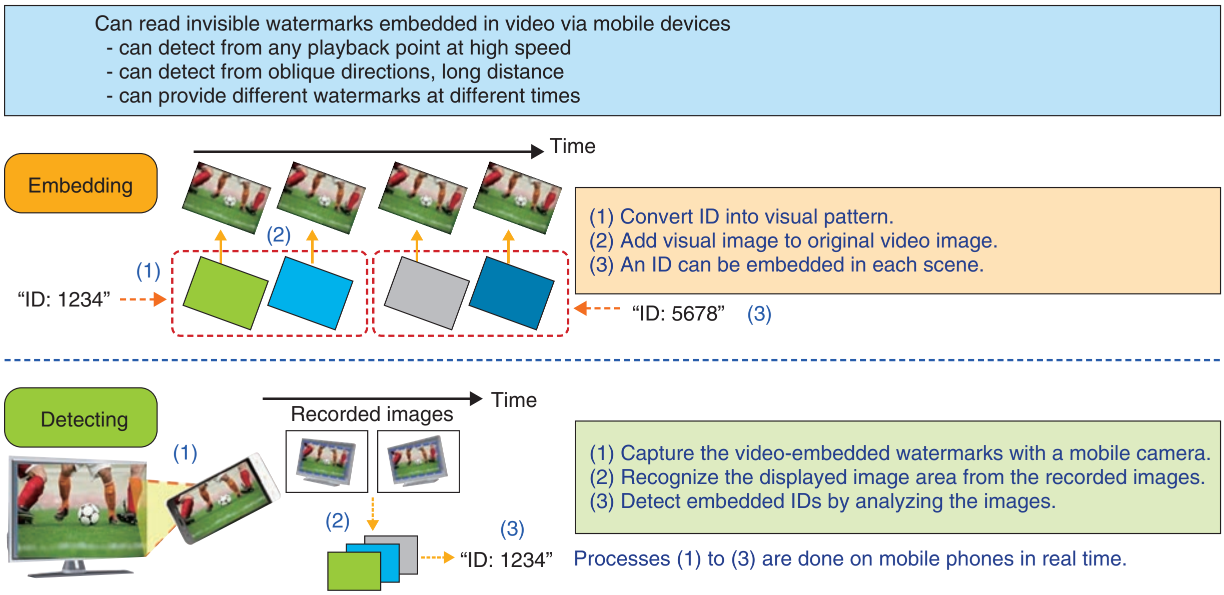


Fig. 2. Mobile video watermarking technology.

customer. We conducted a demonstration experiment using a real world television broadcasting network in December 2012 that showed the feasibility of our technology as an ACR technique.

The subject of our experiment was the general audience of the Mie Television Broadcasting Co., Ltd. We broadcast watermark-embedded video in a TV shopping program produced and delivered by Oak Lawn Marketing, Inc. By detecting embedded watermarks using a smartphone application, viewers of the TV program were able to go directly to the appropriate product purchase site. To our knowledge, this was the first experiment of its type using digital watermarking for video in a real TV broadcasting environment in Japan.

During the experiment, we released an Android application for detecting watermarks on the Google Play[*2] marketplace, so that anyone with a compatible device could participate in the experiment. By sending operation logs of the application to the server, we were able to aggregate and analyze the detection success rate and the time required for detection.

The flow of this experiment is shown in **Fig. 3**. We verified through this experiment that for TV program videos provided via terrestrial digital broadcasting, watermark detection was possible with a success rate in excess of 90% with commercial smartphones, even from a distance equivalent to four to six times the height of the TV screen[*3]. Thus, we have confirmed that our technology can detect the watermark in a variety of receivers and viewing environments in general households without significant problems.

In addition, we were able to detect the watermark from recorded videos and test videos uploaded to YouTube, as well as the initial broadcast source with no problem.

On average, it took 1.7 seconds to detect watermarks using our technology. Previous studies indicate that users will wait between 2 and 8 seconds; we therefore confirmed that our technology meets this rapid identification requirement. In addition, we confirmed that general users were able to operate our application easily.

We also interviewed several users about their experience with the application and the provided service. The interviews concluded that users were highly satisfied in terms of the convenience and efficiency of the service, specifically, the ability to purchase products immediately without having to place a call. In conclusion, this test has shown that our mobile video watermarking technology serves as an effective contact point for new customer acquisition.

*2 We have considered only Android smartphones.
*3 The recommended distance at which to watch widescreen TV is said to be about three times the height of the TV screen.

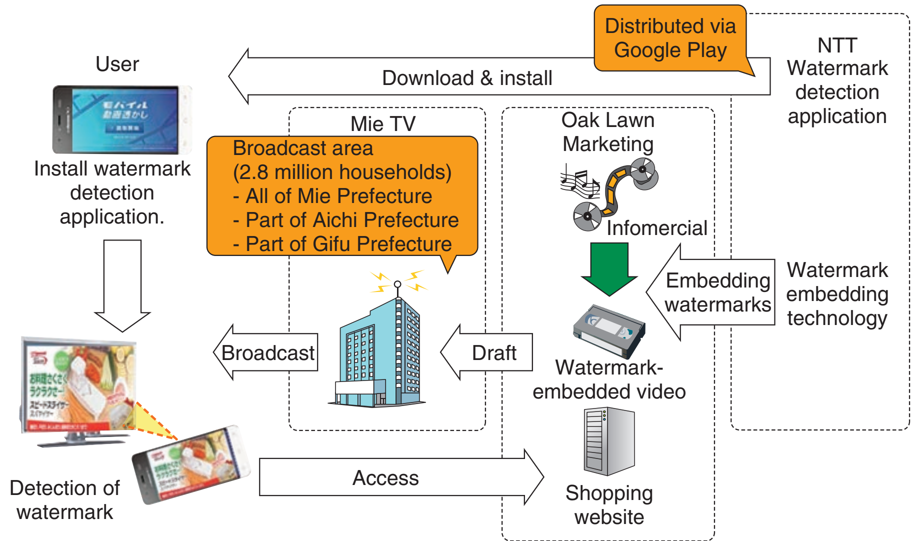


Fig. 3. Flow of demonstration experiment.

5. Video synchronized augmented reality technology: Visual SyncAR

Augmented reality (AR) is a technology that creates an augmented world by superimposing additional information (e.g., virtual computer graphics, text, etc.) onto images and video captured from the real world. The use of high-performance camera-equipped mobile devices such as smartphones has significantly increased in recent years, and because of this, AR has often been used in attractions such as event advertising and product promotion. Frequently, existing applications function by identifying two-dimensional markers or still images, then superimposing three-dimensional (3D) computer graphics (CG) onto the original image.

Traditional AR technology has not augmented temporally relevant information to reality. Although some animated videos have been augmented, they have been independent of still images or markers to be identified.

Visual SyncAR (pronounced "visual sinker"), developed by NTT Media Intelligence Laboratories, is characterized by the ability to superimpose 3D CG animations that are synchronized with the timing of the video captured by the smartphone camera. This technology creates a new content representation that is synchronized between the video and CG animation, in both time and space.

The basis of Visual SyncAR is the mobile video watermarking technology described above. By embedding timestamps into the video's digital watermarks in advance and detecting these watermarks at a reliably high speed, it is possible to synchronize the video and the CG animation. Also, estimating the direction and position of the camera is possible by detecting the region of the captured images that contain the watermarked video. This enables us to superimpose CG content that is spatially synchronized in the same manner as traditional AR [6]. The mechanism of Visual SyncAR is shown in **Fig. 4**.

Visual SyncAR can continuously display the CG content in the appropriate screen position even if the camera position changes. Moreover, even if the playback position is changed by fast forwarding or rewinding the video, it can quickly adapt and display AR content that is synchronized with the video's new playback position.

Visual SyncAR enables unprecedented video expression for mobile devices that jumps out beyond the *fourth wall*[*4].

A demonstration of Visual SyncAR is shown in **Fig. 5**. On the tablet screen, we can see that the stairs have come out beyond the edge of the tablet, and a CG character is dancing in sync with the dancers in the original video. The actual demonstration video can be viewed on the Diginfo TV website [7].

Visual SyncAR is not limited to only new video expressions as in our example; it can be used in a

*4 Fourth wall: This concept describes the boundary between the audience and the content they are observing.

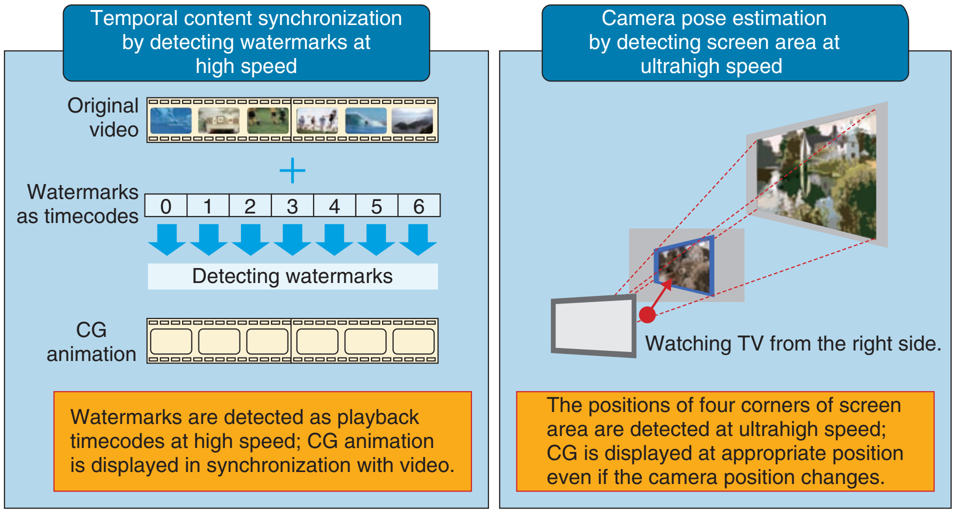


Fig. 4. Mechanism of Visual SyncAR.

Fig. 5. Viewing video content using Visual SyncAR.

Fig. 6. Trial at Kumamon Square using Visual SyncAR.

variety of applications. For example:

a) It can be paired with guide videos of public facilities; sign language CG animation for hearing impaired persons or translations for foreigners can be displayed in real time.
b) Supplemental information can be superimposed on educational videos.
c) Stored navigation information can be displayed by simply holding a smartphone to the floor guide on a digital sign.

These ideas are by no means exhaustive, but they showcase the potential and the versatility of Visual SyncAR.

Currently, a collaborative trial using Visual SyncAR is ongoing in the tourist hub known as Kumamon Square with NTT WEST*5, in the Smart HIKARI-Town Kumamoto project [8]. An image of this trial is shown in **Fig. 6**.

6. Future development

We have described an overview of the Mobile Video Watermarking technology developed by NTT Media Intelligence Laboratories and introduced various use cases related to second screen services, culminating in Visual SyncAR. NTT IT has started marketing this technology under the trade name MagicFinder [9]. In the future, we will continue our research and development aiming at highly detailed service creation.

*5 This trial is planned to run until July 2014.

References

[1] T. Kawanishi, R. Mukai, K. Hiramatsu, T. Kurozumi, H. Nagano, and K. Kashino, "Media Content Identification Technology and Its Applications," Transactions of the Japan Society for Industrial and Applied Mathematics, Vol. 21, No. 4, pp. 49–52, 2011 (in Japanese).

[2] A. Katayama, T. Nakamura, M. Yamamuro, and N. Sonehara, "A New High-speed Corner Detection Method: Side Trace Algorithm (STA) for i-appli to Detect Watermarks," IEICE Trans. Inf. & Syst. (Japanese Edition), Vol. J88-D-II, No. 6, pp. 1035–1046, 2005 (in Japanese).

[3] S. Yamamoto, T. Nakamura, A. Katayama, and T. Yasuno, "Temporal Desynchronization-resistant Video Watermarking Using Single Frequency Plane Spread Spectrum," IEICE Trans. Inf. & Syst. (Japanese Edition), Vol. J90-D, No. 7, pp. 1755–1764, 2007 (in Japanese).

[4] NTT news release (in Japanese).
http://www.ntt.co.jp/news2012/1212/121206a.html

[5] S. Ando, S. Yamamoto, K. Tsutsuguchi, A. Katayama, and Y. Taniguchi, "Field Test of Mobile Video Watermarking Technology over Terrestrial Digital Broadcasting," ITE Technical Report, Vol. 37, No. 38, pp. 57–62, 2013 (in Japanese).

[6] S. Yamamoto, S. Ando, K. Tsutsuguchi, A. Katayama, and Y. Taniguchi, "Visual/Video Synchronizing AR based on Mobile Video Watermark: Visual SyncAR," Proc. of the Media Computing Conference, IIEEJ (the Institute of Image Electronics Engineers of Japan), June 2013 (in Japanese).

[7] Diginfo TV.
http://www.diginfo.tv/v/13-0014-r-en.php

[8] Smart Hikari Town Kumamoto Project (in Japanese).
http://www.hikarikumamoto.jp/service2.html

[9] NTT IT MagicFinder (in Japanese).
http://www.ntt-it.co.jp/product/MagicFinder/top.html

Ken Tsutsuguchi

Senior Research Engineer, Visual Media Project, NTT Media Intelligence Laboratories.

He received the B.S. degree in science and M.E. degree in engineering in 1989 and 1991, respectively and the Ph.D. degree in informatics in 2001 from Kyoto University. Since joining NTT in 1991, he has been engaged in researching computer graphics, computer animation, computer vision, and video/image processing. He is now working on R&D of video watermarking technology. He is a member of the Information Processing Society of Japan (IPSJ), the Institute of Electronics, Information and Communication Engineers (IEICE), the Institute of Image Electronics Engineers of Japan, the Institute of Image Information and Television Engineers (ITE), the IEEE Computer Society, and the Association for Computing Machinery (ACM).

Shingo Ando

Research Engineer, Visual Media Project, NTT Media Intelligence Laboratories.

He received the B.E. degree in electrical engineering and the Ph.D. degree in engineering from Keio University, Kanagawa, in 1998 and 2003, respectively. He joined NTT in 2003. He has been engaged in research and practical application development in the fields of image processing, pattern recognition, and digital watermarking. He is a member of IEICE and ITE.

Atsushi Katayama

Senior Research Engineer, Visual Media Project, NTT Media Intelligence Laboratories.

He received the B.E. and M.E. degrees in mechanical engineering from Kyushu University, Fukuoka. He received the Ph.D. degree in information science from National Institute of Informatics, Tokyo. He has been engaged in researching infrared laser sensing, digital watermarking, and FPGA watermarking design. He received the EMM (the Technical Committee on Enriched MultiMedia) Technical Committee Award in 2013. He is a member of the Japan Society of Mechanical Engineers and the Robotics Society of Japan.

Hidenori Tanaka

Researcher, Visual Media Project, NTT Media Intelligence Laboratories.

He received the B.E., M.E., and Ph.D. degrees in science and technology from Keio University, Kanagawa, in 2004, 2006, and 2011, respectively. Since joining NTT Cyber Space Laboratories in 2006, he has been engaged in research on computer vision and pattern recognition. He moved to NTT Strategic Business Development Division in 2010 and was involved in starting up an e-learning company. He moved to NTT Media Intelligence Laboratories in 2013. His research interests include computer vision, pattern recognition, and digital watermarking.

Susumu Yamamoto

Senior Research Engineer, Visual Media Project, NTT Media Intelligence Laboratories.

He received the B.E. degree in electrical engineering and the M.E. degree in computer science from Keio University, Kanagawa, in 1997 and 1999, respectively. Since joining NTT in 1999, he has been conducting research on content distribution systems, digital watermarking, and AR applications. He is a member of IPSJ.

Yukinobu Taniguchi

Research Engineer, Supervisor and Group Leader of Image Media Processing Group, Visual Media Project, NTT Media Intelligence Laboratories.

He received the B.E., M.E., and Dr.Eng degrees in mathematical engineering from the University of Tokyo in 1990, 1992, and 2002, respectively. He joined NTT in 1992. His research interests include image/video processing and multimedia applications. He is a member of IPSJ, IEICE, and ACM.

NTT Group's Involvement in ITS, and the Direction of Standardization

Tetsufumi Shoji, Norifumi Yasue, Seiji Kuno, and Kazunari Moriuchi

Abstract

In 2013, the ITS (Intelligent Transport Systems) World Congress event returned to Japan for the first time in nine years. NTT participated in this event, and we based our exhibit on the key concept that all mobile objects should function as network terminals that are continuously connected to the cloud. This article introduces the contents of NTT's exhibit, discusses the information and communication technology needed to realize our vision of "2020: Life on the Move," and reports on the direction in which the future efforts of the NTT Group should be steered.

Keywords: ITS, ICT, cloud

1. Introduction

The Intelligent Transport Systems (ITS) initiative aims to use information and communication technology (ICT) to achieve increased efficiency and comfort in the transport field. This initiative is now entering a new stage with the nationwide expansion of technologies such as VICS (Vehicle Information and Communication System) and ETC (electronic toll collection), and the spreading use of car navigation systems. Amid growing interest in technologies such as smartphones, car navigation systems with communication functions, so-called *connected cars* (vehicles equipped with their own communication functions), and self-driving vehicles, the roles of telecommunications providers are becoming increasingly prominent and important.

2. NTT Group's view of the ITS concept

On the basis of our vision, "2020: Life on the Move," we formulated the concept that all mobile objects should function as network terminals that are continuously connected to the cloud (**Fig. 1**). This concept focuses on the applications on the network side to which smartphones and connected cars are connected, and is aimed at the total provision of services including processing large quantities of data on a cloud computing platform, and managing large volumes of traffic. In particular, it could be said that the management of data traffic, which will continue to increase in the future, is precisely a technical field in which NTT has exhibited considerable expertise as a telecommunications carrier. Based on this idea, the NTT Group's presentations were themed around the

Fig. 1. The concept of all mobile objects serving as network terminals that are continuously connected to the cloud.

slogan “YOU GO with ICT. We Can Support You” to indicate the fusion of ICT and mobile living while bearing in mind the ongoing selection of Next Value Partners, even in the traffic sector, through ICT.

3. Exhibiting at the ITS World Congress

The ITS World Congress is a unique international event held in a different country each year by a consortium of ITS organizations representing three regions of the world — Europe (ERTICO), North America (ITS America), and Asia-Pacific (ITS Japan). Its purpose is to create business and find solutions to the issues that face information societies as ITS becomes more widespread. To achieve this, it supports the exchange of information from a wide range of viewpoints on not only technologies but also government policies and market trends. At the 2013 congress, which was held in Tokyo, events such as symposiums, presentations, and showcases were held based on the theme “Open ITS to the Next,” and the NTT Group presented technologies and services aimed at making “life on the move” safe, secure, comfortable, and optimal. Five companies in the NTT Group (NTT DOCOMO, NTT DATA, NTT COMWARE, NTT Geospace, and Nippon Car Solutions) exhibited services and technologies supporting current mobile systems, NTT’s laboratories (five laboratories of three laboratory groups) exhibited technologies supporting future mobile systems, and NTT Research and Development Planning Department exhibited a concept video of the ITS of the future (**Photo 1**).

Photo 1. The NTT Group booth.

One of the technologies that attracted the attention of visitors to the exhibition was the millimeter-wave camera that is capable of finding obstacles on the road even under conditions of poor visibility such as fog or snow (**Fig. 2**). This was exhibited as the result of a joint study by NTT Photonics Laboratories and Nissan Research Center. At NTT Photonics Laboratories, we have been conducting research into technologies that handle radio waves at even higher frequencies than those conventionally referred to as millimeter waves. Millimeter waves in the 140-GHz

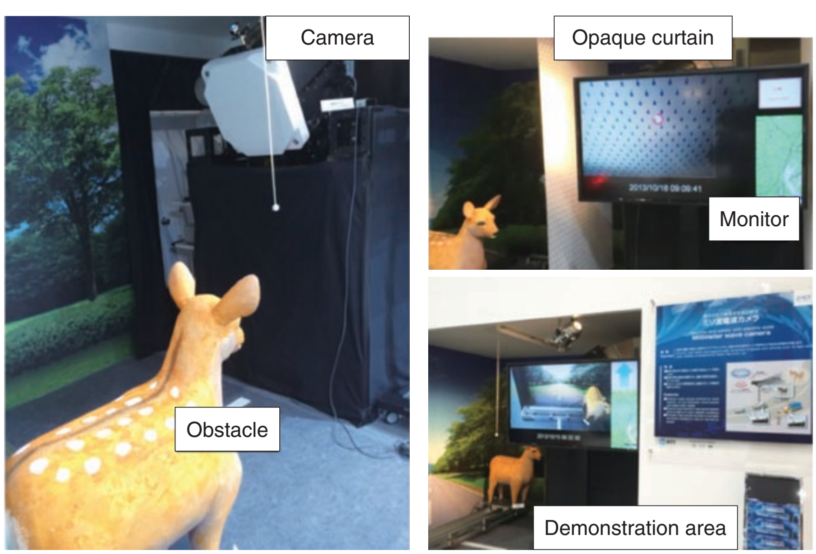


Fig. 2. Demonstration of a millimeter-wave camera.

band are particularly good at penetrating through fog, snow, and the like, and it is even possible to detect millimeter waves emitted naturally by people and obstacles in poor weather conditions where visibility is impaired. At the ITS World Congress, we demonstrated a system capable of tracking and sensing an obstacle represented by a model of a deer moving on rails. This demonstration system was still able to function when visible light had been blocked out by an opaque screen representing the effects of snow. The system received an enthusiastic reception from visitors to the exhibition, several of whom asked us to develop the system for practical applications. This millimeter-wave technology originated from pioneering research into devices for wireless communications. It demonstrates that the NTT Group's technical expertise cultivated over many years in the core areas of ICT can contribute to innovation in ITS, which departs somewhat from NTT Group's usual field of expertise.

4. ITS standardization system

In the ITS field, the International Organization for Standardization/Technical Committee 204 (ISO/TC204) plays a central role in international standardization, and from Japan, the Japanese Industrial Standards Committee (JISC) is participating with the approval of Japan's cabinet [1], [2]. In Japan, an ITS standards committee has been set up in offices at the Society of Automotive Engineers of Japan (JSAE) (**Fig. 3**), and is working on issues to be covered by ISO/TC204. Underneath the ITS standards committee there is a technical committee, and below that are subcommittees that deliberate on the actual standards proposals and liaise with international WGs (Working Groups). These subcommittees are promoted by organizations such as JSAE, the Universal Traffic Management Systems (UTMS) Society, the Japan Electronics and Information Technology Industries Association (JEITA), and the Highway Industry Development Organization (HIDO), which serve as secretariats. The discussions at these subcommittees contribute to various standardization efforts such as the specifications of map data and vehicle-to-vehicle/road-to-vehicle communication, cruise control system interfaces, and traffic management systems.

5. Future direction of the NTT Group

While progress is being made in standardization by manufacturers and industry groups, the theme of the Tokyo ITS World Congress was "Open ITS to the Next," highlighting the issue that the systems of each manufacturer do not necessarily possess the interoperability needed to retain customers. We are approaching standardization in an open, fair, and adaptable way so as to provide users with services that are capable of working together. We believe this approach will go a long way towards increasing Japan's international competitiveness (**Fig. 4**). By cultivating and expanding cloud services and new services using big data to create a bigger pie to share across the entire industry, we believe that it will become more important to request the circulation and utilization of data beyond the frameworks of individual businesses. While the use of big data is attracting interest, it is also rumored that businesses are unwittingly running the risk of harming their reputations through handling of personal data. The NTT Group can mitigate this risk because it operates independently of the systems currently in use by the automotive industry, and it has the experience and technology to be able to handle large amounts of data reliably and securely. As a standard-bearer for the ICT industry, it is essential that we implement our concept of 2020: Life on the Move safely, securely, comfortably, and optimally.

ITS standards committee organization

ITS standards committee*1
- ITS standardization strategy development
- Discussion of draft standards

Secretariat: Society of Automotive Engineers of Japan

Technical committee*2
- Checks the progress of work at subcommittees etc.
- Exchanges information

Liaison
ITS ICT systems forum
Secretariat: Association of Radio Industries and Businesses

Subcommittees and business teams
- Discusses draft standards
- Liaises with international WGs

Subcommittees and business teams	Secretariat
System functional configuration subcommittee (WG1)	Japan Automobile Research Institute
ITS database technology subcommittee (WG3)	Japan Digital Road Map Association
Vehicle/cargo automatic recognition subcommittee (WG4)	UTMS Association
Automatic toll collection subcommittee (WG5)	Highway Industry Development Organization
Commercial freight vehicle fleet management subcommittee (WG7)	Highway Industry Development Organization
Public transportation subcommittee (WG8)	Japan Institute of Country-ology and Engineering
Traffic management subcommittee (WG9)	UTMS Association
Travel information subcommittee (WG10)	UTMS Association
Navigation/route guidance subcommittee (WG11)*3	Society of Automotive Engineers of Japan
Cruise control subcommittee (WG14)	Society of Automotive Engineers of Japan
Narrowband communication subcommittee (WG15)*4	Japan Electronics and Information Technology Industries Association
Broadband communication subcommittee (WG16)	Japan Electronics and Information Technology Industries Association
Nomadic device subcommittee (WG17)	Japan Electronics and Information Technology Industries Association
Cooperative system subcommittee (WG18)	Highway Industry Development Organization
System human interface subcommittee	Society of Automotive Engineers of Japan
ITS international business support team for vehicle-to-vehicle and road-to-vehicle communication	Society of Automotive Engineers of Japan

*1: Comprises approximately 30 members including manufacturers, consumers, and intermediaries
*2: Comprises approximately 30 members including subcommittee chairpersons, business team leaders, liaison officers, and technical specialists
*3: Dormant since April 2004
*4: Dormant since October 2007

Source: “ITS Standardization 2013”, Society of Automotive Engineers of Japan

Fig. 3. ITS standards committee organization.

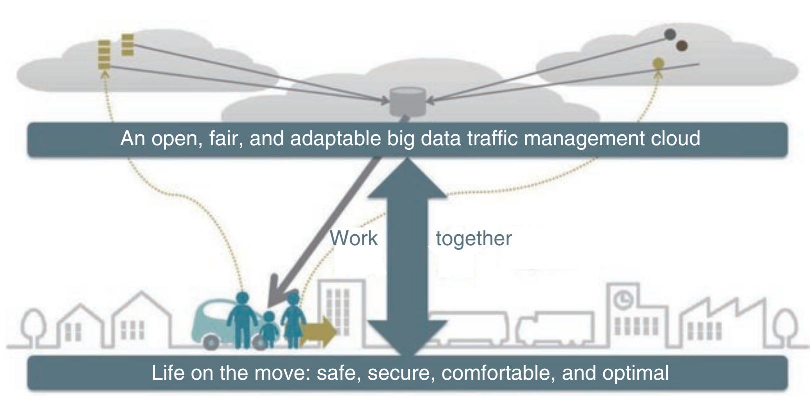


Fig. 4. Image of future efforts by NTT Group.

References

[1] H. Kawashima, "The Globalization of ITS and Its International Standardization," Morikita Publishing Co., Ltd., 2013 (in Japanese).

[2] Society of Automotive Engineers of Japan, "ITS Standardization Activities in Japan, 2013."
https://www.jsae.or.jp/01info/its/2013_bro_e.pdf

Tetsufumi Shoji

Manager, Strategic Business Creation Team, NTT Research and Development Planning Department.

He received the B.E. and M.E. degrees in mechanical engineering from Keio University, Kanagawa, Japan, in 1989 and 1991, respectively. In 1991, he joined NTT Applied Electronics Laboratories. He was engaged in research on nano-scale optical devices. He moved to his current department in February 2013. He is a member of the Information and Communication Engineers (IEICE), the Japan Society of Applied Physics (JSAP) and the Japan Society for Precision Engineering (JSPE).

Norifumi Yasue

Senior Manager, Business Solutions Department, Business Platform Division, NTT Resonant Inc.

He received the B.S. degree in mechanical engineering from Nagoya University, Aichi, Japan, in 1992. He joined NTT Telecommunication Field Systems R&D Center in 1992. He worked on control engineering using neural networks with the aim of applying them to robotics. In 2012, he joined NTT Research and Development Planning Department. He moved to his current department in April 2014. He is a member of the Society of Instrument and Control Engineers (SICE).

Seiji Kuno

Manager, Strategic Business Creation Team, NTT Research and Development Planning Department.

He received the B.E. and M.E. degrees in electrical and electronic engineering from Nagoya University, Aichi, Japan, in 1993 and 1995, respectively. He joined NTT in 1995 and was primarily engaged in corporate planning and service developing. He has held his current position since 2012 and has been engaged in producing services of voice recognition, and wearable sensing.

Kazunari Moriuchi

Chief Producer, Strategic Business Creation Team, NTT Research and Development Planning Department.

Since joining NTT in 1987, he has been engaged in IT service creation. He assumed his current position in 2011 and has been engaged in strategic service incubation.

New NTT Colleagues
—We welcome our newcomers to the NTT Group

Here, we welcome newcomers to the NTT Group. This is a corner of the NTT Technical Review where we introduce our new affiliate companies.

RagingWire Data Centers

RagingWire designs, builds, and operates mission critical data centers that deliver high-density power and 100% availability. The company has 650,000 square feet of critical data center infrastructure in Northern California and Ashburn, Virginia and is affiliated with the global network of 150 data centers operated by NTT Communications. RagingWire's patented power delivery systems and EPA ENERGY STAR rated facilities lead the data center market in reliability and efficiency. With flexible colocation solutions for retail and wholesale buyers, a carrier neutral philosophy, and the highest customer loyalty in the industry as measured by the Net Promoter Score®, RagingWire meets the needs of top enterprise, Internet, and government organizations. More information is available at www.ragingwire.com/.

Contacts
RagingWire Data Centers
http://www.ntt.com/aboutus_e/news/data/20140203_2.html

Arkadin

Arkadin Founded in 2001, Arkadin is one of the largest and fastest growing collaboration service providers in the world. With a vision rooted in the belief that progress emerges from people's desire to share, Arkadin offers a complete range of remote audio, web, and video conferencing and unified communications solutions. The services are delivered in the SaaS model for fast, scalable deployments and a high ROI. Its global network of 52 operating centers in 32 countries has dedicated local-language support teams to service its 37,000 customers. Further information: www.arkadin.com/

Contacts
Voice & Video Communication Service
NTT Communications Corporation
http://www.ntt.com/aboutus_e/news/data/20140122.html

Virtela

Virtela Technology Services Incorporated, recognized by more than 100 awards for outstanding service innovation, is the smart alternative for enterprise networking and virtualized IT services. The company delivers services via its Virtualized Overlay Network and Virtela Enterprise Services Cloud (ESC) platform, which enables the transition of branch office networking and security services to the cloud. Virtela gives enterprises the benefit of "asset-light" software-defined networking that delivers up to 80% cost savings in upfront capital and 30% savings in ongoing operating expenses.

Virtela's services suite includes global managed network, security, mobility, application acceleration and IT infrastructure services. Virtela is headquartered in Denver, Colorado, with global support centers. For more information, please call +1 (720) 475-4000 or visit www.virtela.net.

Contacts
Public Relations Department
NTT Communications Corporation
http://www.ntt.com/aboutus_e/news/data/20140110.html

Papers Published in Technical Journals and Conference Proceedings

An FPGA Based BMNoC Architecture Consisting of Hybrid Connections and Hierarchical Structures with a Router Soft-core

S. Lee, N. Togawa, Y. Sekihara, T. Aoki, M. Nakanishi, and A. Onozawa

Proc. of ITC-CSCC 2013, Vol. 1569754499, No. 1569754499, p. 1569754, Yeosu, Korea, 2013.

In this paper, we propose a busmesh network-on-chip (BMNoC) with parametric soft-core router architecture. The BMNoC adapts a hierarchical communication network consisting of clusters and mesh routers. The BMNoC is a scalable network based on a parametric router architecture to be used in the synthesis of customized low cost NoCs. The switch features allow the BMNoC to scale in several dimensions: node degree, channel width, and buffer depth. Such features allow the automatic customization of BMNoC in order to meet the application requirements.

Single-crystalline 4H-SiC Micro Cantilevers with a High Quality Factor

K. Adachi, N. Watanabe, H. Okamoto, H. Yamaguchi, T. Kimoto, and J. Suda

Sensors and Actuators, Vol. 197, pp. 122–125, 2013.

Single-crystalline 4H-SiC micro cantilevers were fabricated by doping-type selective electrochemical etching of 4H-SiC. With this method, n-type 4H-SiC cantilevers were fabricated on a p-type 4H-SiC substrate, and resonance characteristics of the fabricated 4H-SiC cantilevers were investigated under a vacuum condition. The resonant frequencies agreed very well with the results of numerical simulations. The maximum quality factor in first-mode resonance of the 4H-SiC cantilevers was 230,000. This is 10 times higher than the quality factor of conventional 3C-SiC cantilevers fabricated on an Si substrate.

High-reality Space Composition Using Stably-positioned Imaging and Acoustic Wave Field Synthesis

H. Takada, M. Date, S. Koyama, S. Ozawa, S. Mieda, and A. Kojima

Proc. of European Conference on Visual Perception 2013, Vol. 36, p. 224, Bremen, Germany.

We proposed a natural communication concept produced by high-reality space composition made by using the high fidelity position representation induced by a stably positioned imaging technology and acoustic wave field synthesis technology. It provides natural and comfortable communication by reproducing the distance and position of an image and sound without inconsistency.

We developed a high-reality space composition system using these technologies. We applied our system and reported on aspects concerning the perception and relationship between image and sound. The result indicated a synergistic effect of image and sound. This system can be applied to high-reality communication systems.

Digital Archiving of Tapestries of Kyoto Gion Festival Using a High-definition and Multispectral Image Capturing System

M. Tsuchida, A. Takayanagi, W. Wakita, K. Kashino, J. Yamato, and H. Tanaka

Proc. of the International Conference on Culture and Computing, Vol. 1, No. 1, pp. 204–205, Kyoto, Japan, 2013.

We report the archiving project of the Kyoto Gion Festival using a high-resolution multiband imaging camera. We have been developing a two-shot six-band image capturing system for recording the color and physical properties of early modern tapestries. In an experiment, an image of a tapestry whose image size was 2700 M pixels was synthesized. The resolution of the images was 0.02 mm/pixel. Accurate color under an arbitrary illumination and spectral reflectance can be reproduced from the six-band image using the Wiener estimation.

Bayesian Nonparametric Approach to Blind Separation of Infinitely Many Sparse Sources

H. Kameoka, M. Sato, T. Ono, N. Ono, and S. Sagayama

IEICE Trans. Fundamentals, Vol. E96-A, No. 10, pp. 1928–1937, 2013.

This paper deals with the problem of underdetermined blind source separation (BSS) where the number of sources is unknown. We propose a BSS approach that simultaneously estimates the number of sources, separates the sources based on the sparseness of speech, estimates the direction of arrival of each source, and performs permutation alignment. We confirmed experimentally that reasonably good separation was obtained with the present method without specifying the number of sources.

An Eleven-band Stereoscopic Camera System for Accurate Color and Spectral Reproduction

M. Tsuchida, K. Kashino, and J. Yamato

Proc. of Color and Imaging Conference, Vol. 21, No. 1, pp. 14–19, Albuquerque, USA, 2013.

For accurate color and spectral reflectance reproduction, we propose a novel eleven-band acquisition system using a nine-view stereo camera. The proposed system consists of eight monochrome cameras with eight different narrow band-pass filters and an RGB camera. To generate an eleven-band image, the shapes of the nine captured stereo images are transformed to correct registration displacement caused by stereo parallax. For the process of the correspondence search between stereo images, the POC (phase-only correlation) method is used. The most significant point of our method is that the captured RGB image is converted into narrow-band images for accurate correspondence search. The detected corresponding points are used to estimate the parameters of image transformation, and an eleven-band image is generated. A comparison of the experimental results with those of a conventional method showed that the accuracy of correspondence detection and spectral reflection estimation was improved.

Towards Modeling Ad Hoc Networks: Current Situation and Future Direction

J. Liu, H. Nishiyama, N. Kato, T. Kumagai, and A. Takahara

IEEE Wireless Communications Magazine, Vol. 20, No. 6, pp. 51–58, 2013.

The last decade has witnessed a tremendous increase in both the number of mobile devices and the consumer demand for mobile data communication. As a consequence, networking technologies are shifting from traditional highly centralized toward future organically distributed so as to meet such demand. As the most general networking architecture, the ad hoc network has long been regarded as the most challenging to design and quantify, due to the possible hybrid component settings and heterogeneous node behaviors there. Toward this end, we review the current state of the art of analytical models and techniques developed for performance analysis in ad hoc networks. Specifically, we discuss modeling techniques related to the fundamental topics in ad hoc network research: node mobility, wireless interference, node spatial distribution, and information delivery process. Besides discussions of advantages and limitations of available models, promising future research directions are also outlined.

A Fundamental Study of Efficiency of Information Processing in Emergency Operations Center

F. Ichinose, Y. Maeda, N. Kosaka, M. Higashida, M. Sugiyama, H. Takeda, T. Yamamoto, and H. Hayashi

Journal of Disaster Research, Vol. 9, No. 2, pp. 206–215, 2014.

Many of the emergency operations centers that take the initiative in commanding and controlling a disaster response still rely on inefficient manual information processing even though they have information and communications technology (ICT) systems at hand. This paper reports on three-year functional exercises in a local government using ICT to improve information processing.

Modal Crosstalk Measurement Based on Intensity Tone for Few-mode Fiber Transmission Systems

T. Mizuno, H. Takara, M. Oguma, T. Kobayashi, and Y. Miyamoto

Proc. of Optical Fiber Communication Conference (OFC) 2014, Vol. W3D. 5, pp. 1–3, San Francisco, USA.

We propose a novel method based on intensity tone for measuring modal crosstalk in few-mode fiber transmission systems. Our method can measure crosstalk for multiple modes simultaneously with a wide dynamic range of 40 dB.

12-core × 3-mode Dense Space Division Multiplexed Transmission over 40 km Employing Multi-carrier Signals with Parallel MIMO Equalization

T. Mizuno, T. Kobayashi, H. Takara, A. Sano, H. Kawakami, T. Nakagawa, Y. Miyamoto, Y. Abe, T. Goh, M. Oguma, T. Sakamoto, Y. Sasaki, I. Ishida, K. Takenaga, S. Matsuo, K. Saitoh, and T. Morioka

Proc. of OFC 2014, Vol. Th5B.2, pp. 1–3, San Francisco, USA.

We demonstrate dense SDM transmission of 20-WDM multi-carrier PDM-32QAM signals over a 40-km, 12-core, 3-mode fiber with 247.9-b/s/Hz spectral efficiency. Parallel MIMO (multiple-input and multiple-output) equalization enables 21-ns DMD compensation with 61 TDE taps per subcarrier.

A 204.8 Tbps Throughput 64×64 Optical Cross-connect Prototype that Allows C/D/C Add/drop

K. Takaha, Y. Mori, H. Hasegawa, K. Sato, and T. Watanabe

Proc. of OFC 2014, Vol. M2K.1, pp.1–3, San Francisco, USA.

We fabricate a subsystem modular 64×64 OXC with C/D/C add/drop capabilities. Its throughput reaches 204.8 Tbps at the channel speed of 40 Gbps. Transmission experiments verified the performance of the prototype.

Low Crosstalk Wavelength Tunable Filter that Utilizes Symmetric and Asymmetric Mach-Zehnder Interferometers

S. Takashina, Y. Mori, H. Hasegawa, K. Sato, and T. Watanabe

Proc. of OFC 2014, Vol. Th3F.6, pp.1–3, San Francisco, USA.

We propose a novel AWG-based wavelength tunable filter architecture that utilizes symmetric and asymmetric Mach-Zehnder interferometers for switching and filtering functions. A prototype was fabricated as a PLC and its good performance was experimentally confirmed.

Silica-based PLC 1 × N Switch for All Wavelength Bands

T. Watanabe, T. Mizuno, Y. Hashizume, and T. Takahashi

Proc. of OFC 2014, Vol. Th1I.5, pp. 1–3, San Francisco, USA.

We describe a silica-based PLC switch that operates over a wavelength range of 1260–1610 nm. The fabricated 1×15 switch exhibits a WDL of <1.1 dB and an isolation of >40 dB with a low power consumption of 0.51 W.
