ISSN 1883–1974
No. 28 February 2009
(This English language edition of NII Today corresponds to No.42 of the Japanese edition)

NII Today

National Institute of Informatics News

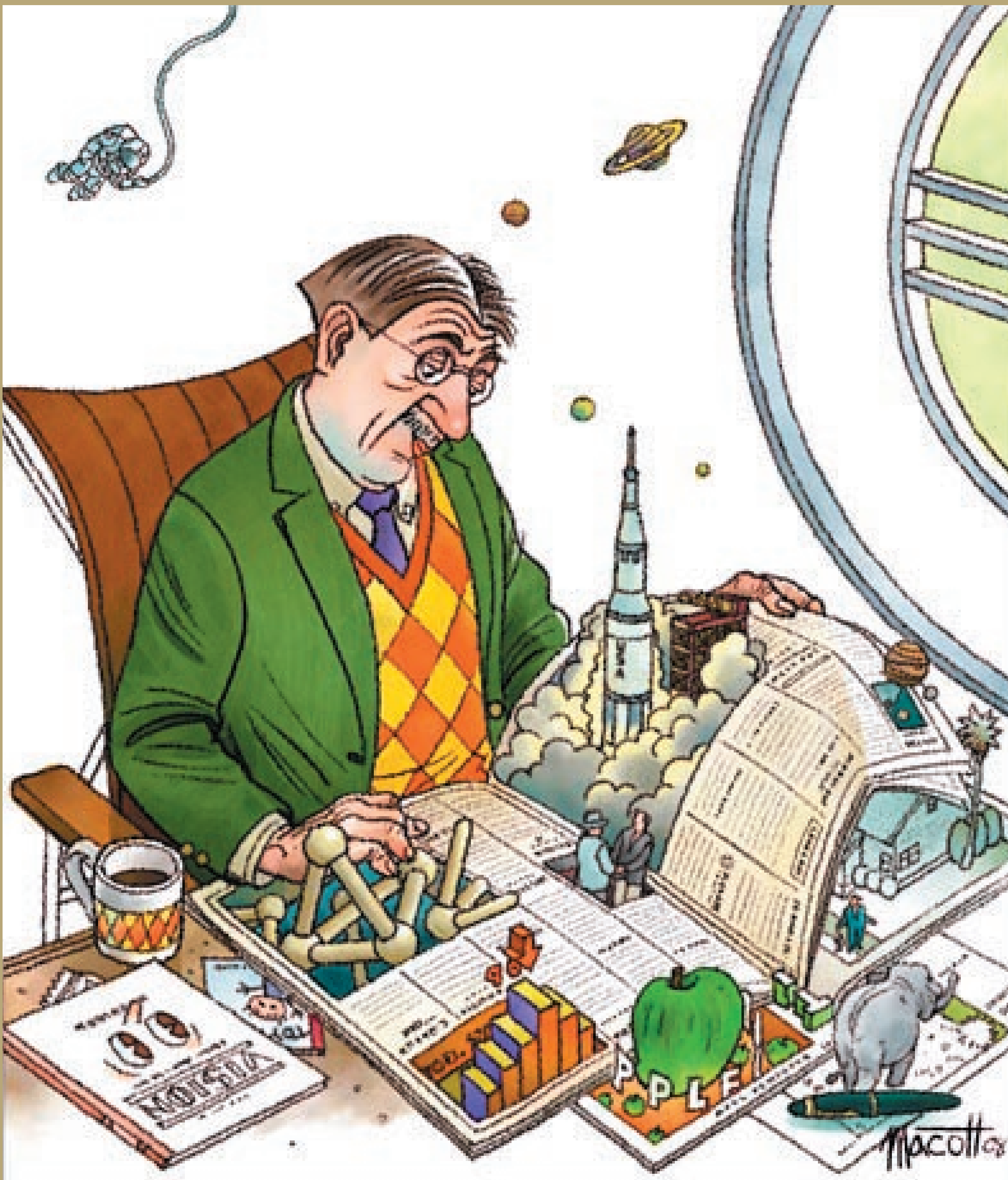
NII SPECIAL **Computer Vision**

Developing a 'anywhere projection display' —that incorporates characteristics of the human visual system

Challenge to guaranteeing accuracy of reconstructed three-dimensional images

Reproducing an 'appearance' with a new viewpoint

Japan-France informatics collaboration research begins



Imari Sato

Associate Professor, Digital Content and Media Sciences Research Division, NII

NII Interview: Imari Sato+Junichi Taki

Developing a 'anywhere projection display'

that incorporates characteristics of the human visual system

NII SPECIAL

COMPUTER VISION

Taki: First of all, tell me about your research objectives.

Sato: The main objective of my research is modeling the shape and complex appearance of real objects and capturing real-world illumination for synthesizing photorealistic appearances of real objects under natural illumination. Producing highly realistic images has a long history, and a variety of pictorial techniques, from the invention of the camera to the development of computer graphics (CG), have appeared.

Taki: Are you using CG in your research?

Sato: I am interested in producing images with a similar realism, but my main interest is computer vision, or CV (*1). CV can ascertain the real-world structure of photographic subjects.

CG synthesizes an image of a model of a scene as seen from the viewpoint of an imaginary camera. The question one asks is "If there was a camera there, what would the scene look like?" CV does the inverse process. Given an image of that scene, CV apprehends the world of the photographic subject and acquires its model.

CG can synthesize amazingly realistic images. However, talking about realism, I think people often feel that a world synthesized by CG is somewhat different from ours. I think that we could generate more realistic images with CG by using a model of a real object acquired by CV.

Taki: Tell me more about your approach to your research.

Sato: The real world is extremely complex and rich. I examine the real world carefully and find its essence from observations. Specifically, I'm working on technologies for modeling the lighting environment of a scene and the shape and appearance of objects in the scene on the basis of observations of the real world.

For modeling appearances, my colleagues and I proposed to incorporate the sampling theorem (*2) for determining a set of lighting and viewing directions. This method efficiently samples the complex appearance changes of a real object. For modeling shapes, we developed a technique for determining an object's shape on the basis of the similarity of radiance changes observed at points on its surface under varying illumination.

These techniques don't require expen-

(*1) Computer vision: A research domain that attempts to determine the structure of real-world scenes from images of those scenes.

(*2) Sampling theorem: An analog signal can be perfectly reconstructed from its samples if the signal has been sampled at proper intervals.

Junichi Taki
Staff Writer, Science & Technology News Department,
the Nihon Keizai Shimbun

sive, special devices and are feasible. I'm also attempting to model larger spaces that surround the subject.

Taki: I have heard that you've developed projector technology that can project images on any surface. In other words, an 'anywhere projection display'.

Sato: Yes. For example, if you project an image onto a wall covered with polka-dot wallpaper, you can see the pattern of the wallpaper through the image. But if a proper compensation process that incorporates the properties of the human visual system is applied to the projected image, it appears as though the image is being projected onto a white screen.

In fact, the polka dots on the wallpaper haven't disappeared, but the input to the projector has been processed so that the user doesn't notice these patterns. This process is based on the fact that the human eye is not so sensitive to smooth brightness variation.

Taki: So in other words, you reduce the dark tones at the edges in the patterns?

Sato: Yes, that's right. In general, the human eye is not so sensitive to differences in brightness in areas of high spatial frequencies such as areas of complex textures.

Taki: So it works with wallpaper or walls with patterns, but what about a wall with a poster or calendar that has writing on it?

Sato: Of course, there are some limits on photometric compensation. For example, a black wall doesn't reflect light, so photometric compensation itself is impossible, and things with a very sharp color or shapes are also difficult to compensate. We are still able to get better projections than with no correction though... Our compensation technique works best with natural patterns such as concrete and wood-grain.

By studying the human eye scientifically, we'll learn the right way of processing and presenting images...

High - performance projectors are getting smaller and cheaper, and there's a growing need for projectors with a variety of uses. There are still issues to resolve, but I think this brings us a step closer to an 'anywhere projector display' that allows images to be projected on any surface such as walls, curtains, and so on. Thinking ahead, wouldn't it be interesting to project advertisements on a big wall surface like a building or something?

Taki: Your research suggests some broad hints for making images that look natural and easy to understand doesn't it?

Sato: Yes, I've realized from this research that people see what they need to see, and do not pay attention to what seems not so important. I feel that by returning to the starting point of computer vision, and by studying the human eye scientifically, we'll learn the proper way of presenting visual information as well as find future directions for image processing research.

Taki: What are you aiming for next?

Sato: I'm interested in technologies that enhance the quality of everyday life. Instead of focusing only on performance and efficiency, I would like to develop computer vision and image processing techniques that are truly human-oriented, and which really meet people's needs.

A word from the Interviewer

There isn't a simple formula for human sensibility that says "lots equals rich". I'm sure that image processing techniques that do not carefully consider human perceptions won't last.

When I interviewed her, I thought Imari Sato is a researcher who really cares about users' perception of the quality of images. I thought that the technologies developed by her research group were practical.

NII SPECIAL

COMPUTER VISION

Challenge to guaranteeing accuracy of reconstructed three-dimensional images

The technology for reconstructing three-dimensional shapes and movements of objects in a computer is improving, and it's becoming possible to reconstruct them without any significant feeling of incongruity. However, there is still one issue remaining, that is, of quality concerning reconstruction accuracy.

Akihiro Sugimoto
Professor, Digital Content and Media Sciences Research Division ,NII

Let's consider reconstructing a mug on a table three-dimensionally in a computer (see the figure on the next page). You capture images of the mug using a digital camera and download its two-dimensional data into the computer. It is preferable to take lots of images from different viewpoints. This is because you will not be able to know what is on the back of the mug if you take only one single image from the front. There may be a dimple there for example. By integrating these two-dimensional images, you reconstruct the three dimensional model of the object. This is the typical approach to three-dimensional image reconstruction in the research field of computer vision(∗).

The question here is, how many images are required to reconstruct the mug completely? Or conversely, if you take images from only two different viewpoints, what percentage can be assigned for the quality of the reconstructed three-dimensional image?

Why does this become an issue? Although we've said "reconstruct completely" above, in fact a mug cannot be reconstructed completely in a computer. As we will explain later, the reason is that although the actual object is analog, it is expressed digitally in the computer. So however many images you use, or however many cameras you use, you cannot reconstruct the object perfectly. Therefore we need a criterion to evaluate reconstructed results based on the number of cameras used and also on their arrangement. Or in other words, we need a guarantee of accuracy concerning reconstructed results.

(∗)Computer vision: A research domain that attempts to determine the structure of real-world scenes from images of those scenes.

Discrete geometry is anti-intuitive

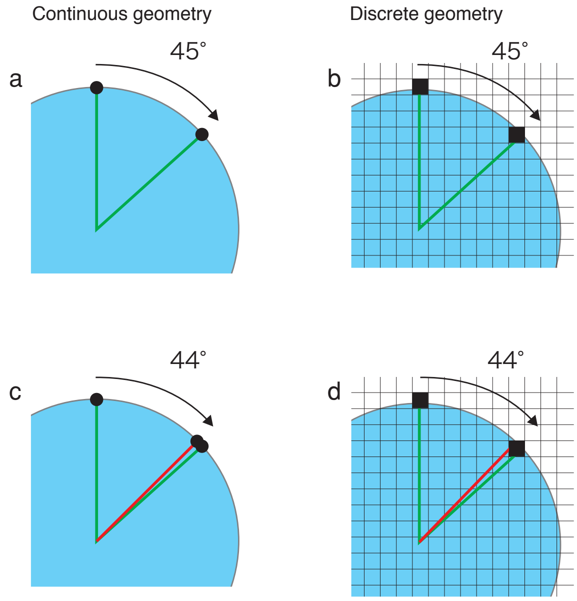
For about three years, Professor Akihiro Sugimoto of the NII Digital Content and Media Sciences Research Division has been conducting research into guaranteeing accuracy of reconstructed three-dimensional shapes. He has long felt the importance of guaranteeing accuracy, "But there was no effective methodology available, and I didn't know how to go about it in concrete terms. It was then that I first heard about discrete geometry in a lecture", says Prof. Sugimoto.

The Euclidean geometry that we are most familiar with is continuous (analog) geometry. But because digital computers have only discrete values in them, it is easy to imagine that a 'discrete' geometry will be useful.

However, this was the start of Prof. Sugimoto's difficulties. "Discrete geometry is completely different from continuous geometry. First of all, it was really difficult to get to grips with the unfamiliar idea of discrete geometry", he says.

So how different are discrete geometry and continuous geometry? Take two-dimensional rotation as an example. In the figures on the next page, a and c are the continuous geometry that we are all familiar with, while b and d are discrete geometry.

The needle pointing straight up is rotated clockwise by 45 degrees. There doesn't appear to be any dif-



a and c show continuous geometry, while b and d show discrete geometry. There doesn't appear to be any difference even when the needle is rotated clockwise by 45 degrees (a and b). When the needles, rotated by 44 degrees, are superimposed, the needles are pointing to different positions in a and c, but in d, they're pointing to the same place as the 45 degree rotation.

ference between continuous and discrete cases. However, let's superimpose them when the needle is rotated by 44 degrees (Figures c and d). On the left, with continuous geometry, the needles are pointing to different positions, while on the right, with discrete geometry, they are pointing to the same place. In other words, in the discrete world shown in this example, 44 degrees and 45 degrees are indiscriminative.

If you model a three-dimensional object from lots of its two-dimensional images, there is no way of avoiding the issue of these discrete peculiarities. The fact is that, however high you make the resolution of the images, you cannot avoid this as far as you are handling them digitally. Currently, even if you reconstruct something fairly well in 3D, there is no guarantee of its quality (accuracy).

Many difficulties ahead

Whereas continuous geometry can specify points, discrete geometry can only specify ranges, or pixels in terms of digital images. “Even though I got it with my head, I couldn't shake a certain wooly feeling, and it bothered me for days”, says Prof. Sugimoto.

Moreover, discrete geometry is not yet completely established. Two dimensional cases are fairly well studied. But three dimensional cases cannot really be mastered by simply extending two dimensional cases, which makes it quite a formidable opponent.

An actual object (analog, 3D) forms several two-dimensional digital images. These are then processed by a computer and reconstructed into digital 3D images using computer vision techniques. And so using the three-dimensional discrete geometry is inevitable.

Says Prof. Sugimoto, “In the research field of computer vision, the main stress is put on technologies for making three-dimensional images that look nice and natural to the human eye, while research from a mathematical perspective, like guaranteeing accuracy, is in the minority. Also, discrete geometry is a minor field even in pure mathematics, and I think there are hardly any Japanese researchers involved in it. But if you reconstruct 3D images without any guarantee of accuracy, it's like making a product with unknown specs. So users can't really use them with any confidence, can they?” Even though Prof. Sugimoto recognizes the importance, he's fighting a lonely battle.

Is discrete geometry promising to the future of computer vision?

Since this has all the appearance of exploratory research, its social usefulness in the future is unknown. However, it can, for example, answer the question of how cameras must be placed, and where, to achieve a 99% accurate reconstruction with maximum efficiency.

All cameras used for reconstruction are supposed to have the same spec in the literature, but as the research proceeds, it may show that efficiency can be improved by skillfully combining cameras with different specs. Conversely, it may be useful when less accuracy is required and you want to reduce the number of cameras.

There are still heaps of other issues to think about, such as how to handle videos rather than still images, or colors and patterns rather than just shapes. Prof. Sugimoto is hopeful; “I think it'll be nice if, in the future, we can make 3D images with guaranteed quality that you can use with confidence, using only the power of a computer”.

(Written by Tomoaki Yoshito)

Reproducing an 'appearance' with a new viewpoint

Can we convert a two-dimensional image of an object into an image seen from a new viewpoint? This has become possible thanks to improvements in image processing technology, but there are still many problems to overcome. Up to the present, research has sought to establish a framework of what is possible and what isn't, and the theoretical support behind it.

Kazuya Kodama
Associate Professor, Digital Content and Media Sciences Research Division, NII

Suppose that you want to view an object in a two-dimensional image taken with a digital camera from a different viewpoint. In order to reconstruct the image seen from any viewpoint that you want it would be good if you could reproduce its original three-dimensional form of the object.

However, if there's only one original two-dimensional image, it's theoretically impossible to reproduce its original form. That's because there isn't enough depth information. So if we take depth into account and add another image taken from the side of the object, is it possible to reproduce the three-dimensional image from the two two-dimensional images? We can easily see that this won't work either. That's because there isn't any information about the back of the object. So, how many two-dimensional images do you need to reproduce its original form of the object?

In fact, it isn't possible to reproduce the form completely without images taken from all directions, so it isn't easy to realize. That just leaves improving the technology for reproducing the original form of the object as far as possible using the limited available two-dimensional images. This sums up the thinking so far.

Considering only the 'appearance'

Here, a groundbreaking idea was suggested by Associate Professor Kazuya Kodama of the NII Digital Content and Media Sciences Research Division. Until now, the following three-step process was employed; 1) Take two-dimensional images of an object, 2) reproduce its original three-dimensional form, 3) reconstruct the desired two-dimensional image. However, Associate Professor Kodama says, "I gave up trying to reproduce the original form and decided to convert the two-dimensional image of the object directly into a two-dimensional image from another viewpoint."

He reached this position because he realized the limits to pursuing the 'original form'. If you simply combine two-dimensional images, there isn't enough information to reproduce its original form of the object. For example, if the computer cannot accurately identify a shadow, it may make an image with a person's nose appearing as a hollow rather than a bump. In that case, it would take human intervention to correct the contradiction. Specifically, that would entail programming assumptions (*1) into the image processing. For example, noses stick out, and ears have holes.

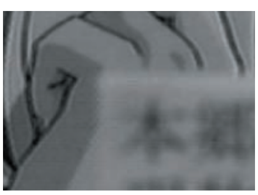
Associate Professor Kodama's idea is that for automatic image processing by a computer without human assistance, not to attempt reproducing the original form of an object

(Figure 1) Generating an arbitrary focus image
(a) Image with near focus, (b) image with far focus. Using these two images, an image with near/far in focus (c), or with near/far blurred (f) can be made. Additionally, arbitrary images such as (d) and (e) can be made.

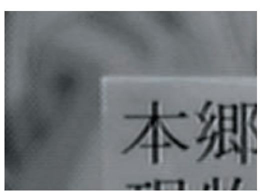
a Original image (near focus)

c All-in-focus image (near/far focus)

d Arbitrary focus image (suppressed near blur)

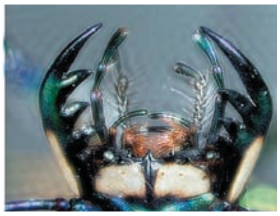


b Original image (far focus)

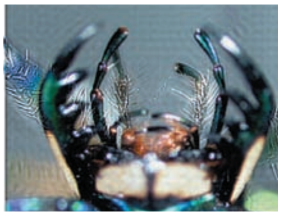


e Arbitrary focus image (emphasized far blur)

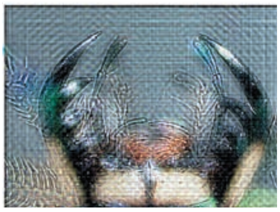
f Arbitrary focus image (emphasized near/far blur)



(8 images)

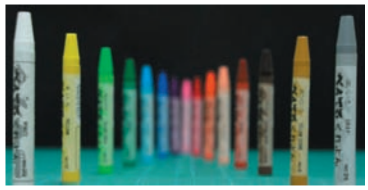


(16 images)



(30 images)

(Figure 2) All-in-focus images reconstructed from microscopic images. The true result cannot be reproduced simply by increasing the number of images.



Original image (near focus)

Reconstructed image (observation slightly from the right)

Original image (far focus)

Reconstructed image (observation slightly from the left)

(Figure 3) Generating a free-viewpoint image. When 64 images are taken with various focuses, they can be combined to make an image from a different viewpoint.

is practical.

Focal bokeh control

First, Associate Professor Kodama researched the subject of focal bokeh (*2) and depth of field as a simple model. (Figure 1) shows two images, one with a near focus, and the other with a far focus. By combining these two images with some effects, it's possible to create two variations of the image "automatically", one with both areas in focus, and the other with neither in focus. If a third image is added with a different focus, an image with more variations can be generated. If you gradually increase the variations on the original two-dimensional image in this way, it should be possible to create images with a rich range of variations.

However, when 33 microscopic images were taken (provided by Prof. Kenji Kohiyama), and a number of images were selected to reconstruct an image, it was found that increasing the number of images made it difficult to reconstruct an image (Figure 2). However, since this resulted from the difficulties of calculation, it proved possible to improve this by revising the method used for calculation.

As a result, it became possible to reconstruct an image using 64 two-dimensional images without any problem and to create images with a different viewpoint by naturally controlling the bokeh (Figure 3).

'Truth' is something beyond our reach

To express the shift in paradigm that favors 'appearance' without worrying about the 'original form', Associate Professor Kodama resorts to terms developed by philosophers. "The original form is what Kant called 'Ding an sich' (thing-in-itself) or what Plato called 'eidos' (idea). Husserl applied 'epoché' (bracketing) to the pursuit of the thing-in-itself, and conceived a phenomenology based on observation. This suggests that there's an approach that's concerned only with appearance and another which seeks after reality."

People can't help pursuing the original form, the thing-in-itself, or to put it another way, the 'truth'. However, if for example we're asked, "What is an apple?", we're stuck for an answer. If we say, it's red, it's round, it's sweet, that only describes the color, shape, and taste of an apple. So we put aside the original form which is not likely to provide an answer, and chose instead the appearance, which can provide an answer. It's fascinating that an information scientist has reached the same conclusion as the philosophers, through a completely different approach.

Satisfying acute vision

As for future applications of this approach, the most obvious seems to be entertainment. For example, if several cameras are placed in a concert hall or sports stadium, it's possible to use those images to create an image as seen from the chosen seat of the user. Another interesting application would be microscopes. The technology will be useful as a visual aid for designing and processing semiconductor devices made up of a number of layers.

The requirements for human visual information are demanding. If the pitch of a sound is slightly off, many people won't notice it, but if for example the surface of a tennis ball is out by just one degree, most people would notice the abnormality. Associate Professor Kodama wants to achieve automatic image processing on a computer, without preprogrammed assumptions, that "doesn't disrupt the viewer's dreams (the common sense that a ball is round)".

(Written by Tomoaki Yoshito)

*1 Preprogrammed assumptions: Assumptions programmed into computations involve knowledge obtained through experience, such as that balls are round and books are rectangular. Although there's a tendency to think that the more knowledge is available, the closer you can get to reality, it isn't so simple. For example, if the list of assumptions includes the knowledge that some balls are square, the computer cannot determine whether balls are round or square, and it stops computing.

*2 Bokeh: Areas that are deliberately blurred using a lens effect. Bokeh is a form of aesthetic expression originating in Japan. The Japanese term *'bokeh'* is rendered as *'bokeh'* in English.

In search of a new form of collaboration

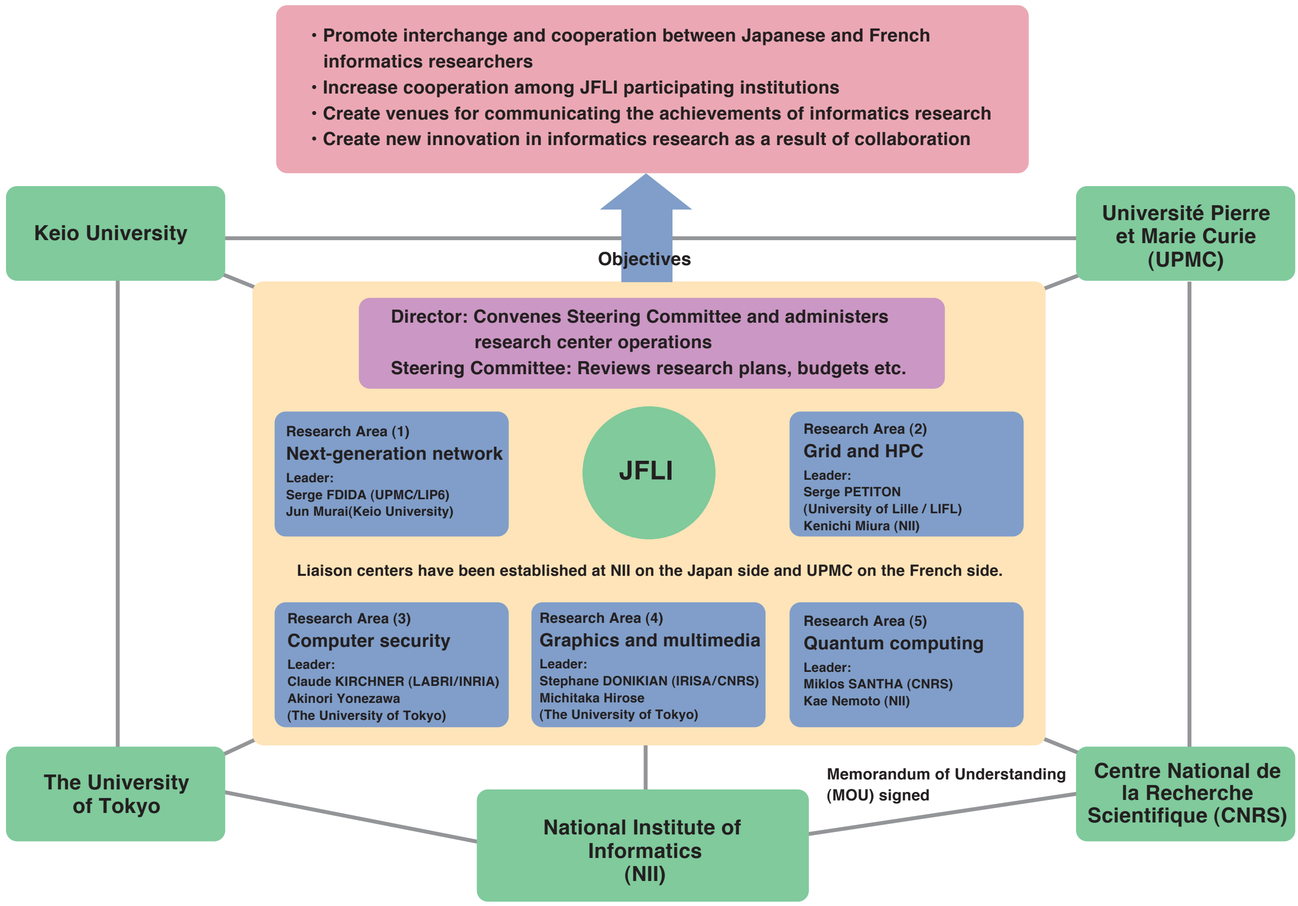
Japan-France informatics collaboration research begins

NII has produced numerous major achievements through collaborative efforts with many research institutions, companies, universities and other entities. In December 2008, the Japanese-French Laboratory for Informatics (JFLI) begins a new kind of collaborative effort between Japan and France. In the following pages, we will examine the ways in which JFLI represents a new form of collaboration and the goals that it aims to achieve.

The Japan liaison center of the Japanese-French Laboratory for Informatics (JFLI), a new organization to promote joint research between Japan and France, has been established on the 12th floor of the NII building and has begun full-fledged operations. NII, the University of Tokyo and Keio University will conduct cooperative research within a framework in which the Centre National de la Recherche Scientifique (CNRS) plays a leading role (Figure 1).

Overall management of the JFLI will be conducted by NII on the Japan side and CNRS on the French side, and liaison centers have been established at NII and Université Pierre et Marie Curie (UPMC). Both sides will designate leaders to grapple with five major areas in informatics. Professor Akinori Yonezawa of the University of Tokyo, who studies programming languages and information security, will assume the post of leader for computer security research on the Japan side. Professor Michitaka Hirose, also of the University of Tokyo and well-known for his research into virtual reality, will lead research into graphics and multimedia. Professor Jun Murai of Keio University, who has worked to establish an Internet infrastructure, will lead research into next-generation networks. Professor Kenichi Miura of NII will lead research into HPC and establishing grids for networks that link computers for high-speed data sharing, and Associate Professor Kae Nemoto, also of NII, will lead research into quantum computing (for more information about Professor Nemoto, see the article in NII

Figure 1 JFLI organization



Jun Adachi
Professor and Director, Cyber Science Infrastructure Development Department, NII

Philippe Codognet
Professor, Research Institute for Digital Media and Content, Keio University

Henri Angelino-
Acting Director, Global Liaison Office, NII

Today No. 27). Professor Jun Adachi of NII and Professor Philippe Codognet, CNRS staff member and researcher at Keio University, will serve as directors in charge of administration and coordination of the overall operations of JFLI.

More dynamic collaboration

NII has concluded cooperative research agreements with many French research institutions, including the CNRS, Institut National de Recherche en Informatique et en Automatique (INRIA), the UPMC, and Nantes University, and is promoting joint research and mutual research exchanges as well as accepting interns and so on. Moreover, Prof. Hirose, who has been appointed as leader of graphics and multimedia research, has already had interchange with France's Université Louis-Pasteur and other institutions. In recognition of the significance of such exchanges, the Japan Science and Technology Agency (JST) is providing financial assistance for exchanges as a Strategic International Cooperative Program.

In 2006, CNRS proposed that institutions with researchers who were already cooperating individually in research projects with French institutions conduct "more dynamic" collaboration. Prof. Codognet, the CNRS/UPMC staff member who proposed this collaboration, described the reasons leading to the proposal for an organization like the JFLI as follows. "Individual collaborative efforts between Japan and France in the field of informatics are on track. To ensure their continuation, we need stable bi-lateral relationships."

Established in October 1939, the CNRS is the largest governmental institution for basic science research in France, employing some 26,000 researchers and engineers and operating more than 1,300 research centers and laboratories in France alone (Figure 2). The research conducted at these locations covers various fields ranging from physics to the humanities and social science. The fact that CNRS has grown as large as it has is due not only to its own laboratories but also its active efforts to establish joint laboratories with universities and other research institutions. The advantages of joint laboratories is that it is easy to create many research centers as well as to incorporate talented researchers in joint research projects.

CNRS has expanded its circle of collaborative alliances primarily in EU nations. Recently, however, it has also begun to focus on Asia, and now has research centers in countries such as China, South Korea, Vietnam and Thailand as well. In Japan, CNRS has established joint laboratories in five locations, among them the University of Tokyo (for microelectronics research), the National Institute of Advanced Industrial Science and Technology (AIST) (for robotics research), and the High Energy Accelerator Research Organization (KEK) (for particle physics research). In some cases, CNRS also sets up joint laboratories at foreign companies, provided that an agreement is reached regarding rights and interests. The organization is a flexible one that is prepared to participate in any attractive research project.

In this sense, CNRS could be called a collaboration expert. CNRS has now invited NII, the University of Tokyo and Keio University to join it in creating JFLI as a completely new type of collaborative research organization.

For the future of informatics research

Henri Angelino, formerly chancellor of the Institut National Polytechnique de Toulouse (INPT) and counselor for the French Embassy in Japan, and currently

Figure 2 Centre National de la Recherche Scientifique (CNRS)

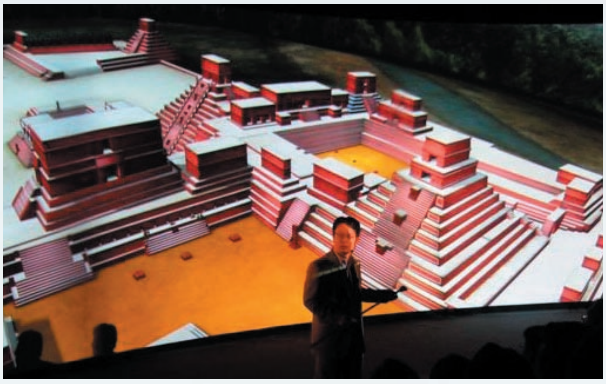
- Largest governmental basic science research institution in France, established in October 1939
- Approximately 26,000 employees (around 11,000 researchers and 15,000 engineers)
- Comprises 1,300 research centers in France alone
- Major research fields: physics, mathematics, atomic physics, particle physics, space science, engineering, chemistry, life science, humanities and social science
- Liaison offices established in 10 locations around the world
- Research is conducted by individual research units. In some cases, independent CNRS research units are established. In others, research units are established jointly with universities or other research institutions.
- President: Catherine Bréchignac Director General: Arnold Migus

Michitaka Hirose
Professor, Department of Mechano-Informatics, Faculty of Engineering, The University of Tokyo

Shinichi Satoh
Professor, Digital Content and Media Sciences Research Division, NII

Acting Director of NII's Global Liaison Office, thinks that the fact that CNRS, an institution in France, has established connections among three Japanese research centers makes this an extremely novel collaboration framework. This marks a first even for CNRS, indicating the truly unusual nature of this collaborative research organization. As a result, there are a variety of expectations on both the Japanese and French sides. Foremost among these is the hope that valuable research will be conducted. But Prof. Codognet says that for young researchers, the chance to see various research institutions will be a plus for their research careers. Accordingly, he thinks that more researcher interchange should be conducted between Japan and France. Prof. Jun Adachi of NII, the director on the Japan side, says that researchers who want to produce achievements in a specific area of research should gather researchers who are strong in that area and have them collaborate on the project. "The reason that JFLI invited researchers in a wide variety of fields is because the goal is to produce medium- and long-term achievements," he says. The hope is that the five research areas will mix with one another and that their chemistry will give birth to something new.

Initially, the central focus will be on collaboration in the area of financing — cooperating with one another to request funding and thinking of the way to most effectively use research funds. From a long-term perspective, however, no one yet has a clear understanding at present of how this new collaborative organization will function.



Restoration of historical legacy through virtual reality

Media research in the spotlight

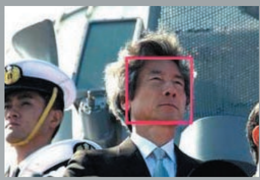
What kind of research will actually be conducted within the JFLI framework? For some of the five research areas, the specific content has not yet been determined. Prof. Hirose says that since the collaboration will be "French-style," a lot is still not known on the Japanese side. "However, the establishment of the JFLI framework has created the opportunity to have informational exchanges several times a year," he says. "These exchanges will undoubtedly produce new research topics." He welcomes the contact with numerous researchers that will result from this organization.

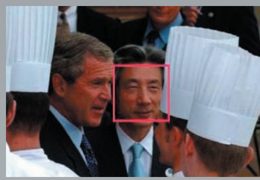
The defining characteristic of the discipline of informatics is that even the single area of graphics or multimedia research involves content that is truly diverse. One example is haptic technology. The word "haptic" means pertaining to the sense of touch. It indicates the reaction force and the feeling of smoothness or roughness when the surface of a hard or soft object is touched. The vibration function of a mobile phone and the bodysonic device in a train simulator (which creates the illusion that the floor is shaking) are the result of developments in haptic technology. This technology that utilizes the sense of touch has applications in the transmission of information to sight-impaired persons.

Some researchers see haptic technology as art. They feel that, for example, a floor that makes a scratching sound when a pen is used to write characters on the floor constitutes a type of artistic expression. The scratchy feel when the characters are written has a psychological effect, producing certain feelings and images in the writer. By fusing the informatics and engineering aspects with artistic and psychological aspects, haptic technology seems likely to become an even more intriguing field of study.

Olfactory sensor

Although still largely unexplored, senses other than touch such as smell





Example showing automatic detection of a specific person's face in an image on the Internet. The more famous the person is, the easier it is to detect that person's face.

Sport, walking / running

Beach, mountain, car

Plane

Face

Results of automatic meaning categorization of images. Each image is automatically assigned a label that indicates its meaning category (sport, plane, mountain, car etc.).

and taste are also the focus of media research. The same smell and taste are perceived differently depending on the color with which they are presented. In this way, smell and taste are very interesting as tools for the transmission of information.

Assimilating one other's research

Approximately 20 years ago when the concept first became known, virtual reality was a technology that enabled people wearing goggles and gloves to feel as if the objects in an image were actually real. Now, however, virtual reality has come to the point at which it could recreate the real world within a computer with exact precision. In actuality, however, there will never be enough time to create every one of the objects that make up our world. This problem would be solved if we had a miraculous scanner that could make exact copies of things, right down to their texture and uses.

A photograph can now copy the scenery exactly as it is, but this does not constitute a virtual reality. The flowers and tables in the photograph are nothing more than collections of dots that do not embody the meaning of a flower or table. For this reason, viewing a collection of photographs and selecting only those that show a flower is easy for a human being with the sense of sight, but it is very difficult for a computer. If there were a way to invest photographs with meaning, computers would also be able to locate photographs containing flowers. This would lead to the technology that could create the miraculous scanner that is able to recognize a flower and recreate it down to its texture and uses.

Prof. Satoh of NII, who will participate in the JFLI's graphics and multimedia research projects, is conducting research on this very topic: finding a way to invest images with meaning. In its earliest stages, research into image recognition focused on numbers and letters, and this technology is already at work in the form of the address recognition scanners in use at the post office and so on. More recently, it has become possible to recognize human faces with considerable accuracy. It is quite possible that Prof. Hirose's virtual reality research and Prof. Satoh's image recognition research will lead to new and groundbreaking advances. Already at Prof. Hirose's laboratory, students are working to develop a "Who's That? System" that will instantly identify the person standing in front of you. In this area of research as well, the encounter between these students and Prof. Sato seems destined to lead to major technical breakthroughs.

Creating a stir in the world of informatics

Although they barely knew one another, Prof. Hirose and Prof. Sato were extremely interested in each other's research, and as time went on their conversations became forums for informational exchange. They saw with their own eyes how an encounter between researchers could provide a major stimulus to both parties.

Informatics is a field that is poised to grow further in the coming years, but as a discipline it is not yet firmly established. "Informatics is a field that does not produce many major discoveries in terms of fundamentals, so it's difficult to get people to see how important it is," says Prof. Hirose. For example, the growth of the Internet was a major event related to informatics, but in the background of this event are factors such as improvements in semiconductor technologies that led to the increasingly compact size and lower prices of personal computers and their increasing use in individual households. This characteristic of informatics has led some to feel that it is best left to companies. "Yet it is when pure research institutions challenge a variety of topics, without getting caught up in the quest for profit, that new fields of research are produced," says Prof. Hirose.

Prof. Adachi wants to achieve further development for informatics in Japan by utilizing his experience as a director involved with management and operation. "NII is not a very large organization. So in order to cover the entire field of informatics, I think we should draw in many other domestic research institutions and create a framework like JFLI." The new collaboration organization of JFLI that emphasizes exchanges among researchers is expected to produce major changes in the discipline of informatics.

(Written by Akiko Ikeda)

NII ESSAY

Do You Open Your Mobile While You're Walking?

Kenro Aihara
Associate Professor, Digital Content and Media Sciences Research Division, NII

We hear much recently about something called the 'Galapagos phenomenon.' The term refers to the way that Japan's technology is developing in its own way, isolated from the rest of the world market, in a manner comparable with the evolution of life forms on those Ecuadorian islands. Mobile phones are a typical example of the phenomenon.

The way that mobiles in Japan are used more as IT terminals than they are in other countries is another peculiarity of our country. On the streets of Japan's towns, you can see many people holding their mobiles in front of their faces, staring intently at the screen and punching keys. And recently, many people are actually to be seen using their mobiles while they walk down the road. It goes without saying, however, that this can lead to all sorts of dangers.

'e-Spaces' Delivering the Information Users Want

The ways in which a mobile can be used are somewhat limited in situations like this, situations in which the screen is not easy to look at. Interaction is possible through sound and vibration, but it is still difficult to effectively get across information in contexts other than phone calls. So what can we do in order to get information as we walk?

Research is underway into providing users in the real world with the information they need in an appropriate format. One such example is the 'e-Space' project being launched by the Ministry of Economy, Trade and Industry (METI), which seeks to install sensors all over towns through which users will be provided with information according to their situation. The development and practical testing of the field service will begin by the end of 2008. This has now led to R&D concerning what information should be transmitted, and how it should be transmitted, to meet the objectives and interests of the users. How can the circumstances and intentions of the users walking through the town be ascertained? How can they be sent the necessary information? It would be feasible to obtain user data from profiles, and the cameras and sensors situated throughout the town. It is also possible to acquire data from the built-in sensors on mobiles and wearable bio-sensors, as well as peripheral sounds and images and the users' access logs. And using this data, it might be possible to predict the circumstances and intentions of users.

Cultivating Technologies that Surpass the Limitations of Mobiles

However, the problem of transmitting the information to the users remains. Even if the information is sent to the user as e-mail, reading mail while you walk is not exactly practical. Interaction that surpasses the limitations of mobiles, something that does not depend on screen displays and key-punching, becomes necessary.

That's what I would like to see from the e-Spaces. In an e-Space, rather than depending entirely upon transmissions sent to mobiles, information aimed at individual users is built in to the space where they are, and the data is supplied by monitors and other devices located on street corners and in stores. There are still unresolved issues, such as the matter of privacy, but this can certainly be described as one challenging approach to the questions of going beyond the mobile phone.

The ideal of the e-Spaces, that unspecified numbers of people will walk through, is that as well as letting users obtain useful information they will enable people to spend their time peacefully, and let those sharing the space neatly harmonize with each other. In order not to encourage the further isolation of individuals, some ingenuity will be required of the design - such as building in information aimed at the environmental aspects, in other words the data transmission to each mobile and the space itself. I hope above all that e-Spaces will be nurtured into a technology that goes beyond the Galapagos phenomenon.

Weaving Information into Knowledge

NII

NII Today No.28, February 2009 (This English language edition of NII Today corresponds to No.42 of the Japanese edition)
Published by: National Institute of Informatics, Research Organization of Information and Systems
Address: National Center of Sciences 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430
Chief editor: Yoh'ichi Tohkura Cover illustration: Makoto Komori Photography: Shuichi Yuri Design: Kotaro Suzuki Production: Sci-Tech Communications Inc.
Contact: Publicity and Dissemination Team, Planning and Promotion Strategy Department TEL:+81-3-4212-2135 FAX:+81-3-4212-2150 e-mail: kouhou@nii.ac.jp
http://www.nii.ac.jp/

Imari Sato

Associate Professor, Digital Content and Media Sciences Research Division, NII

NII Interview: Imari Sato+Junichi Taki

Developing a ‘anywhere projection display’

that incorporates characteristics of the human visual system

NII SPECIAL

COMPUTER VISION

Taki: First of all, tell me about your research objectives.

Sato: The main objective of my research is modeling the shape and complex appearance of real objects and capturing real-world illumination for synthesizing photorealistic appearances of real objects under natural illumination. Producing highly realistic images has a long history, and a variety of pictorial techniques, from the invention of the camera to the development of computer graphics (CG), have appeared.

Taki: Are you using CG in your research?

Sato: I am interested in producing images with a similar realism, but my main interest is computer vision, or CV (*1). CV can ascertain the real-world structure of photographic subjects.

CG synthesizes an image of a model of a scene as seen from the viewpoint of an imaginary camera. The question one asks is “If there was a camera there, what would the scene look like?” CV does the inverse process. Given an image of that scene, CV apprehends the world of the photographic subject and acquires its model.

CG can synthesize amazingly realistic images. However, talking about realism, I think people often feel that a world synthesized by CG is somewhat different from ours. I think that we could generate more realistic images with CG by using a model of a real object acquired by CV.

Taki: Tell me more about your approach to your research.

Sato: The real world is extremely complex and rich. I examine the real world carefully and find its essence from observations. Specifically, I’m working on technologies for modeling the lighting environment of a scene and the shape and appearance of objects in the scene on the basis of observations of the real world.

For modeling appearances, my colleagues and I proposed to incorporate the sampling theorem (*2) for determining a set of lighting and viewing directions. This method efficiently samples the complex appearance changes of a real object. For modeling shapes, we developed a technique for determining an object’s shape on the basis of the similarity of radiance changes observed at points on its surface under varying illumination.

These techniques don’t require expen-

(*1) Computer vision: A research domain that attempts to determine the structure of real-world scenes from images of those scenes.

(*2) Sampling theorem: An analog signal can be perfectly reconstructed from its samples if the signal has been sampled at proper intervals.

Junichi Taki

Staff Writer, Science & Technology News Department,
the Nihon Keizai Shimbun

sive, special devices and are feasible. I'm also attempting to model larger spaces that surround the subject.

Taki: I have heard that you've developed projector technology that can project images on any surface. In other words, an 'anywhere projection display'.

Sato: Yes. For example, if you project an image onto a wall covered with polka-dot wallpaper, you can see the pattern of the wallpaper through the image. But if a proper compensation process that incorporates the properties of the human visual system is applied to the projected image, it appears as though the image is being projected onto a white screen.

In fact, the polka dots on the wallpaper haven't disappeared, but the input to the projector has been processed so that the user doesn't notice these patterns. This process is based on the fact that the human eye is not so sensitive to smooth brightness variation.

Taki: So in other words, you reduce the dark tones at the edges in the patterns?

Sato: Yes, that's right. In general, the human eye is not so sensitive to differences in brightness in areas of high spatial frequencies such as areas of complex textures.

Taki: So it works with wallpaper or walls with patterns, but what about a wall with a poster or calendar that has writing on it?

Sato: Of course, there are some limits on photometric compensation. For example, a black wall doesn't reflect light, so photometric compensation itself is impossible, and things with a very sharp color or shapes are also difficult to compensate. We are still able to get better projections than with no correction though... Our compensation technique works best with natural patterns such as concrete and wood-grain.

By studying the human eye scientifically, we'll learn the right way of processing and presenting images...

High - performance projectors are getting smaller and cheaper, and there's a growing need for projectors with a variety of uses. There are still issues to resolve, but I think this brings us a step closer to an 'anywhere projector display' that allows images to be projected on any surface such as walls, curtains, and so on. Thinking ahead, wouldn't it be interesting to project advertisements on a big wall surface like a building or something?

Taki: Your research suggests some broad hints for making images that look natural and easy to understand doesn't it?

Sato: Yes, I've realized from this research that people see what they need to see, and do not pay attention to what seems not so important. I feel that by returning to the starting point of computer vision, and by studying the human eye scientifically, we'll learn the proper way of presenting visual information as well as find future directions for image processing research.

Taki: What are you aiming for next?

Sato: I'm interested in technologies that enhance the quality of everyday life. Instead of focusing only on performance and efficiency, I would like to develop computer vision and image processing techniques that are truly human-oriented, and which really meet people's needs.

A word from the Interviewer

There isn't a simple formula for human sensibility that says "lots equals rich". I'm sure that image processing techniques that do not carefully consider human perceptions won't last.

When I interviewed her, I thought Imari Sato is a researcher who really cares about users' perception of the quality of images. I thought that the technologies developed by her research group were practical.

COMPUTER
VISION

Challenge to guaranteeing accuracy of reconstructed three-dimensional images

The technology for reconstructing three-dimensional shapes and movements of objects in a computer is improving, and it's becoming possible to reconstruct them without any significant feeling of incongruity. However, there is still one issue remaining, that is, of quality concerning reconstruction accuracy.

Akihiro Sugimoto
Professor, Digital Content and Media Sciences Research Division ,NII

(＊)Computer vision: A research domain that attempts to determine the structure of real-world scenes from images of those scenes.

Let's consider reconstructing a mug on a table three-dimensionally in a computer (see the figure on the next page). You capture images of the mug using a digital camera and download its two-dimensional data into the computer. It is preferable to take lots of images from different viewpoints. This is because you will not be able to know what is on the back of the mug if you take only one single image from the front. There may be a dimple there for example. By integrating these two-dimensional images, you reconstruct the three dimensional model of the object. This is the typical approach to three-dimensional image reconstruction in the research field of computer vision(＊).

The question here is, how many images are required to reconstruct the mug completely? Or conversely, if you take images from only two different viewpoints, what percentage can be assigned for the quality of the reconstructed three-dimensional image?

Why does this become an issue? Although we've said "reconstruct completely" above, in fact a mug cannot be reconstructed completely in a computer. As we will explain later, the reason is that although the actual object is analog, it is expressed digitally in the computer. So however many images you use, or however many cameras you use, you cannot reconstruct the object perfectly. Therefore we need a criterion to evaluate reconstructed results based on the number of cameras used and also on their arrangement. Or in other words, we need a guarantee of accuracy concerning reconstructed results.

Discrete geometry is anti-intuitive

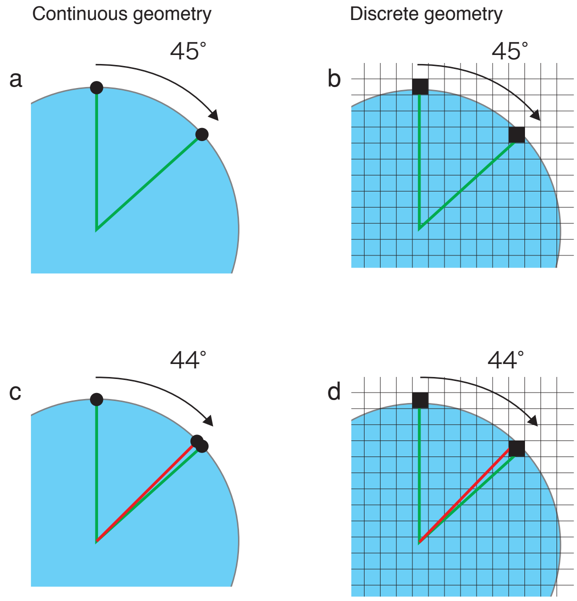
For about three years, Professor Akihiro Sugimoto of the NII Digital Content and Media Sciences Research Division has been conducting research into guaranteeing accuracy of reconstructed three-dimensional shapes. He has long felt the importance of guaranteeing accuracy, "But there was no effective methodology available, and I didn't know how to go about it in concrete terms. It was then that I first heard about discrete geometry in a lecture", says Prof. Sugimoto.

The Euclidean geometry that we are most familiar with is continuous (analog) geometry. But because digital computers have only discrete values in them, it is easy to imagine that a 'discrete' geometry will be useful.

However, this was the start of Prof. Sugimoto's difficulties. "Discrete geometry is completely different from continuous geometry. First of all, it was really difficult to get to grips with the unfamiliar idea of discrete geometry", he says.

So how different are discrete geometry and continuous geometry? Take two-dimensional rotation as an example. In the figures on the next page, a and c are the continuous geometry that we are all familiar with, while b and d are discrete geometry.

The needle pointing straight up is rotated clockwise by 45 degrees. There doesn't appear to be any dif-



a and c show continuous geometry, while b and d show discrete geometry. There doesn't appear to be any difference even when the needle is rotated clockwise by 45 degrees (a and b). When the needles, rotated by 44 degrees, are superimposed, the needles are pointing to different positions in a and c, but in d, they're pointing to the same place as the 45 degree rotation.

ference between continuous and discrete cases. However, let's superimpose them when the needle is rotated by 44 degrees (Figures c and d). On the left, with continuous geometry, the needles are pointing to different positions, while on the right, with discrete geometry, they are pointing to the same place. In other words, in the discrete world shown in this example, 44 degrees and 45 degrees are indiscriminative.

If you model a three-dimensional object from lots of its two-dimensional images, there is no way of avoiding the issue of these discrete peculiarities. The fact is that, however high you make the resolution of the images, you cannot avoid this as far as you are handling them digitally. Currently, even if you reconstruct something fairly well in 3D, there is no guarantee of its quality (accuracy).

Many difficulties ahead

Whereas continuous geometry can specify points, discrete geometry can only specify ranges, or pixels in terms of digital images. "Even though I got it with my head, I couldn't shake a certain wooly feeling, and it bothered me for days", says Prof. Sugimoto.

Moreover, discrete geometry is not yet completely established. Two dimensional cases are fairly well studied. But three dimensional cases cannot really be mastered by simply extending two dimensional cases, which makes it quite a formidable opponent.

An actual object (analog, 3D) forms several two-dimensional digital images. These are then processed by a computer and reconstructed into digital 3D images using computer vision techniques. And so using the three-dimensional discrete geometry is inevitable.

Says Prof. Sugimoto, "In the research field of computer vision, the main stress is put on technologies for making three-dimensional images that look nice and natural to the human eye, while research from a mathematical perspective, like guaranteeing accuracy, is in the minority. Also, discrete geometry is a minor field even in pure mathematics, and I think there are hardly any Japanese researchers involved in it. But if you reconstruct 3D images without any guarantee of accuracy, it's like making a product with unknown specs. So users can't really use them with any confidence, can they?" Even though Prof. Sugimoto recognizes the importance, he's fighting a lonely battle.

Is discrete geometry promising to the future of computer vision?

Since this has all the appearance of exploratory research, its social usefulness in the future is unknown. However, it can, for example, answer the question of how cameras must be placed, and where, to achieve a 99% accurate reconstruction with maximum efficiency.

All cameras used for reconstruction are supposed to have the same spec in the literature, but as the research proceeds, it may show that efficiency can be improved by skillfully combining cameras with different specs. Conversely, it may be useful when less accuracy is required and you want to reduce the number of cameras.

There are still heaps of other issues to think about, such as how to handle videos rather than still images, or colors and patterns rather than just shapes. Prof. Sugimoto is hopeful; "I think it'll be nice if, in the future, we can make 3D images with guaranteed quality that you can use with confidence, using only the power of a computer".

(Written by Tomoaki Yoshito)

COMPUTER VISION

Reproducing an 'appearance' with a new viewpoint

Can we convert a two-dimensional image of an object into an image seen from a new viewpoint? This has become possible thanks to improvements in image processing technology, but there are still many problems to overcome. Up to the present, research has sought to establish a framework of what is possible and what isn't, and the theoretical support behind it.

Kazuya Kodama
Associate Professor, Digital Content and Media Sciences Research Division, NII

Suppose that you want to view an object in a two-dimensional image taken with a digital camera from a different viewpoint. In order to reconstruct the image seen from any viewpoint that you want it would be good if you could reproduce its original three-dimensional form of the object.

However, if there's only one original two-dimensional image, it's theoretically impossible to reproduce its original form. That's because there isn't enough depth information. So if we take depth into account and add another image taken from the side of the object, is it possible to reproduce the three-dimensional image from the two two-dimensional images? We can easily see that this won't work either. That's because there isn't any information about the back of the object. So, how many two-dimensional images do you need to reproduce its original form of the object?

In fact, it isn't possible to reproduce the form completely without images taken from all directions, so it isn't easy to realize. That just leaves improving the technology for reproducing the original form of the object as far as possible using the limited available two-dimensional images. This sums up the thinking so far.

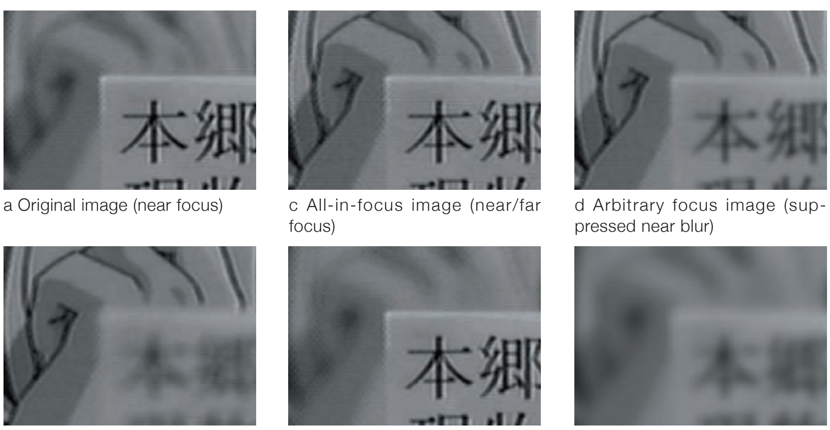
Considering only the 'appearance'

Here, a groundbreaking idea was suggested by Associate Professor Kazuya Kodama of the NII Digital Content and Media Sciences Research Division. Until now, the following three-step process was employed; 1) Take two-dimensional images of an object, 2) reproduce its original three-dimensional form, 3) reconstruct the desired two-dimensional image. However, Associate Professor Kodama says, "I gave up trying to reproduce the original form and decided to convert the two-dimensional image of the object directly into a two-dimensional image from another viewpoint."

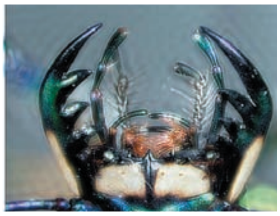
He reached this position because he realized the limits to pursuing the 'original form'. If you simply combine two-dimensional images, there isn't enough information to reproduce its original form of the object. For example, if the computer cannot accurately identify a shadow, it may make an image with a person's nose appearing as a hollow rather than a bump. In that case, it would take human intervention to correct the contradiction. Specifically, that would entail programming assumptions (*1) into the image processing. For example, noses stick out, and ears have holes.

Associate Professor Kodama's idea is that for automatic image processing by a computer without human assistance, not to attempt reproducing the original form of an object

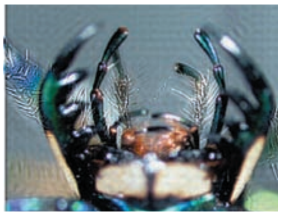
(Figure 1) Generating an arbitrary focus image
(a) Image with near focus, (b) image with far focus. Using these two images, an image with near/far in focus (c), or with near/far blurred (f) can be made. Additionally, arbitrary images such as (d) and (e) can be made.



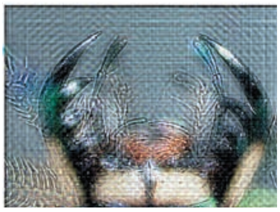
a Original image (near focus)
c All-in-focus image (near/far focus)
d Arbitrary focus image (suppressed near blur)
b Original image (far focus)
e Arbitrary focus image (emphasized far blur)
f Arbitrary focus image (emphasized near/far blur)



(8 images)

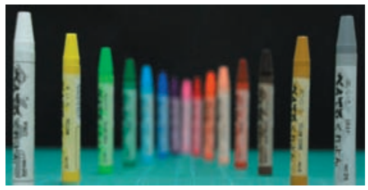


(16 images)



(30 images)

(Figure 2) All-in-focus images reconstructed from microscopic images. The true result cannot be reproduced simply by increasing the number of images.



Original image (near focus)

Reconstructed image (observation slightly from the right)

Original image (far focus)

Reconstructed image (observation slightly from the left)

(Figure 3) Generating a free-viewpoint image. When 64 images are taken with various focuses, they can be combined to make an image from a different viewpoint.

is practical.

Focal bokeh control

First, Associate Professor Kodama researched the subject of focal bokeh (*2) and depth of field as a simple model. (Figure 1) shows two images, one with a near focus, and the other with a far focus. By combining these two images with some effects, it's possible to create two variations of the image "automatically", one with both areas in focus, and the other with neither in focus. If a third image is added with a different focus, an image with more variations can be generated. If you gradually increase the variations on the original two-dimensional image in this way, it should be possible to create images with a rich range of variations.

However, when 33 microscopic images were taken (provided by Prof. Kenji Kohiyama), and a number of images were selected to reconstruct an image, it was found that increasing the number of images made it difficult to reconstruct an image (Figure 2). However, since this resulted from the difficulties of calculation, it proved possible to improve this by revising the method used for calculation.

As a result, it became possible to reconstruct an image using 64 two-dimensional images without any problem and to create images with a different viewpoint by naturally controlling the bokeh (Figure 3).

'Truth' is something beyond our reach

To express the shift in paradigm that favors 'appearance' without worrying about the 'original form', Associate Professor Kodama resorts to terms developed by philosophers. "The original form is what Kant called 'Ding an sich' (thing-in-itself) or what Plato called 'eidos' (idea). Husserl applied 'epoché' (bracketing) to the pursuit of the thing-in-itself, and conceived a phenomenology based on observation. This suggests that there's an approach that's concerned only with appearance and another which seeks after reality."

People can't help pursuing the original form, the thing-in-itself, or to put it another way, the 'truth'. However, if for example we're asked, "What is an apple?", we're stuck for an answer. If we say, it's red, it's round, it's sweet, that only describes the color, shape, and taste of an apple. So we put aside the original form which is not likely to provide an answer, and chose instead the appearance, which can provide an answer. It's fascinating that an information scientist has reached the same conclusion as the philosophers, through a completely different approach.

Satisfying acute vision

As for future applications of this approach, the most obvious seems to be entertainment. For example, if several cameras are placed in a concert hall or sports stadium, it's possible to use those images to create an image as seen from the chosen seat of the user. Another interesting application would be microscopes. The technology will be useful as a visual aid for designing and processing semiconductor devices made up of a number of layers.

The requirements for human visual information are demanding. If the pitch of a sound is slightly off, many people won't notice it, but if for example the surface of a tennis ball is out by just one degree, most people would notice the abnormality. Associate Professor Kodama wants to achieve automatic image processing on a computer, without preprogrammed assumptions, that "doesn't disrupt the viewer's dreams (the common sense that a ball is round)".

(Written by Tomoaki Yoshito)

*1 Preprogrammed assumptions: Assumptions programmed into computations involve knowledge obtained through experience, such as that balls are round and books are rectangular. Although there's a tendency to think that the more knowledge is available, the closer you can get to reality, it isn't so simple. For example, if the list of assumptions includes the knowledge that some balls are square, the computer cannot determine whether balls are round or square, and it stops computing.

*2 Bokeh: Areas that are deliberately blurred using a lens effect. Bokeh is a form of aesthetic expression originating in Japan. The Japanese term *'bokeh'* is rendered as *'bokeh'* in English.

In search of a new form of collaboration

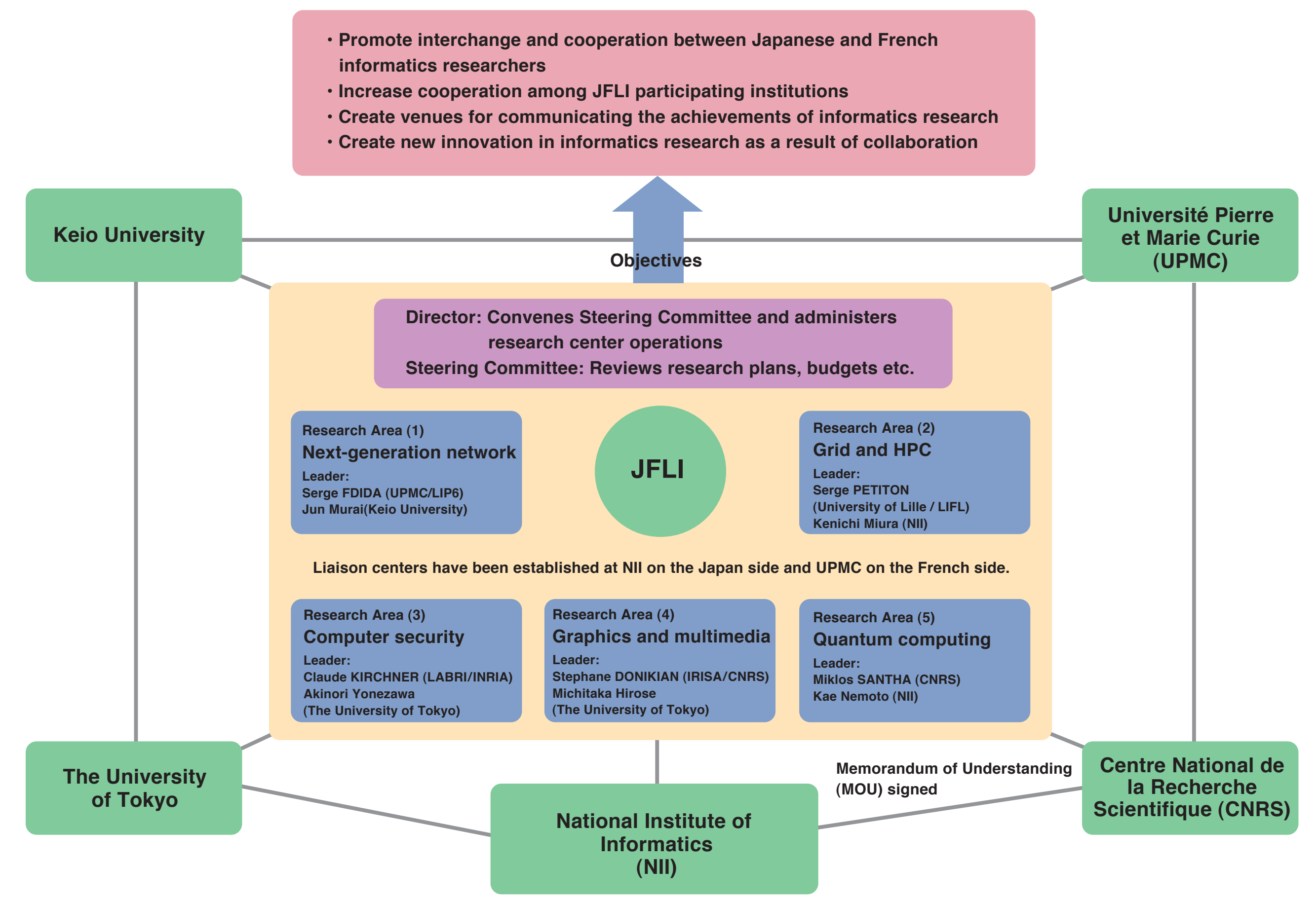
Japan-France informatics collaboration research begins

NII has produced numerous major achievements through collaborative efforts with many research institutions, companies, universities and other entities. In December 2008, the Japanese-French Laboratory for Informatics (JFLI) begins a new kind of collaborative effort between Japan and France. In the following pages, we will examine the ways in which JFLI represents a new form of collaboration and the goals that it aims to achieve.

The Japan liaison center of the Japanese-French Laboratory for Informatics (JFLI), a new organization to promote joint research between Japan and France, has been established on the 12th floor of the NII building and has begun full-fledged operations. NII, the University of Tokyo and Keio University will conduct cooperative research within a framework in which the Centre National de la Recherche Scientifique (CNRS) plays a leading role (Figure 1).

Overall management of the JFLI will be conducted by NII on the Japan side and CNRS on the French side, and liaison centers have been established at NII and Université Pierre et Marie Curie (UPMC). Both sides will designate leaders to grapple with five major areas in informatics. Professor Akinori Yonezawa of the University of Tokyo, who studies programming languages and information security, will assume the post of leader for computer security research on the Japan side. Professor Michitaka Hirose, also of the University of Tokyo and well-known for his research into virtual reality, will lead research into graphics and multimedia. Professor Jun Murai of Keio University, who has worked to establish an Internet infrastructure, will lead research into next-generation networks. Professor Kenichi Miura of NII will lead research into HPC and establishing grids for networks that link computers for high-speed data sharing, and Associate Professor Kae Nemoto, also of NII, will lead research into quantum computing (for more information about Professor Nemoto, see the article in NII

Figure 1　JFLI organization



Jun Adachi
Professor and Director, Cyber Science Infrastructure Development Department, NII

Philippe Codognet
Professor, Research Institute for Digital Media and Content, Keio University

Henri Angelino-
Acting Director, Global Liaison Office, NII

Today No. 27). Professor Jun Adachi of NII and Professor Philippe Codognet, CNRS staff member and researcher at Keio University, will serve as directors in charge of administration and coordination of the overall operations of JFLI.

More dynamic collaboration

NII has concluded cooperative research agreements with many French research institutions, including the CNRS, Institut National de Recherche en Informatique et en Automatique (INRIA), the UPMC, and Nantes University, and is promoting joint research and mutual research exchanges as well as accepting interns and so on. Moreover, Prof. Hirose, who has been appointed as leader of graphics and multimedia research, has already had interchange with France's Université Louis-Pasteur and other institutions. In recognition of the significance of such exchanges, the Japan Science and Technology Agency (JST) is providing financial assistance for exchanges as a Strategic International Cooperative Program.

In 2006, CNRS proposed that institutions with researchers who were already cooperating individually in research projects with French institutions conduct "more dynamic" collaboration. Prof. Codognet, the CNRS/UPMC staff member who proposed this collaboration, described the reasons leading to the proposal for an organization like the JFLI as follows. "Individual collaborative efforts between Japan and France in the field of informatics are on track. To ensure their continuation, we need stable bi-lateral relationships."

Established in October 1939, the CNRS is the largest governmental institution for basic science research in France, employing some 26,000 researchers and engineers and operating more than 1,300 research centers and laboratories in France alone (Figure 2). The research conducted at these locations covers various fields ranging from physics to the humanities and social science. The fact that CNRS has grown as large as it has is due not only to its own laboratories but also its active efforts to establish joint laboratories with universities and other research institutions. The advantages of joint laboratories is that it is easy to create many research centers as well as to incorporate talented researchers in joint research projects.

CNRS has expanded its circle of collaborative alliances primarily in EU nations. Recently, however, it has also begun to focus on Asia, and now has research centers in countries such as China, South Korea, Vietnam and Thailand as well. In Japan, CNRS has established joint laboratories in five locations, among them the University of Tokyo (for microelectronics research), the National Institute of Advanced Industrial Science and Technology (AIST) (for robotics research), and the High Energy Accelerator Research Organization (KEK) (for particle physics research). In some cases, CNRS also sets up joint laboratories at foreign companies, provided that an agreement is reached regarding rights and interests. The organization is a flexible one that is prepared to participate in any attractive research project.

In this sense, CNRS could be called a collaboration expert. CNRS has now invited NII, the University of Tokyo and Keio University to join it in creating JFLI as a completely new type of collaborative research organization.

For the future of informatics research

Henri Angelino, formerly chancellor of the Institut National Polytechnique de Toulouse (INPT) and counselor for the French Embassy in Japan, and currently

Figure 2 Centre National de la Recherche Scientifique (CNRS)

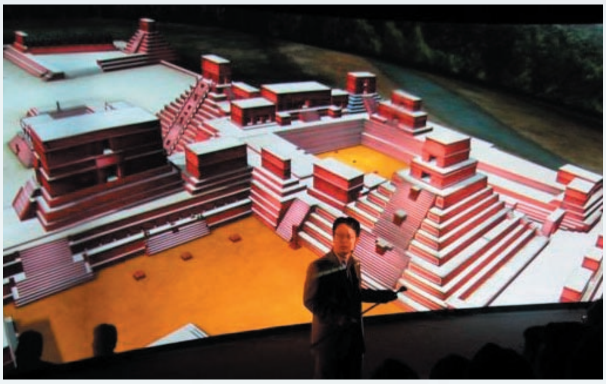
- Largest governmental basic science research institution in France, established in October 1939
- Approximately 26,000 employees (around 11,000 researchers and 15,000 engineers)
- Comprises 1,300 research centers in France alone
- Major research fields: physics, mathematics, atomic physics, particle physics, space science, engineering, chemistry, life science, humanities and social science
- Liaison offices established in 10 locations around the world
- Research is conducted by individual research units. In some cases, independent CNRS research units are established. In others, research units are established jointly with universities or other research institutions.
- President: Catherine Bréchignac Director General: Arnold Migus

Michitaka Hirose
Professor, Department of Mechano-Informatics, Faculty of Engineering, The University of Tokyo

Shinichi Satoh
Professor, Digital Content and Media Sciences Research Division, NII

Acting Director of NII's Global Liaison Office, thinks that the fact that CNRS, an institution in France, has established connections among three Japanese research centers makes this an extremely novel collaboration framework. This marks a first even for CNRS, indicating the truly unusual nature of this collaborative research organization. As a result, there are a variety of expectations on both the Japanese and French sides. Foremost among these is the hope that valuable research will be conducted. But Prof. Codognet says that for young researchers, the chance to see various research institutions will be a plus for their research careers. Accordingly, he thinks that more researcher interchange should be conducted between Japan and France. Prof. Jun Adachi of NII, the director on the Japan side, says that researchers who want to produce achievements in a specific area of research should gather researchers who are strong in that area and have them collaborate on the project. "The reason that JFLI invited researchers in a wide variety of fields is because the goal is to produce medium- and long-term achievements," he says. The hope is that the five research areas will mix with one another and that their chemistry will give birth to something new.

Initially, the central focus will be on collaboration in the area of financing — cooperating with one another to request funding and thinking of the way to most effectively use research funds. From a long-term perspective, however, no one yet has a clear understanding at present of how this new collaborative organization will function.



Restoration of historical legacy through virtual reality

Media research in the spotlight

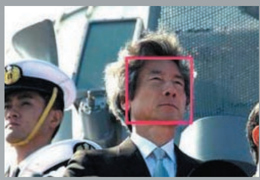
What kind of research will actually be conducted within the JFLI framework? For some of the five research areas, the specific content has not yet been determined. Prof. Hirose says that since the collaboration will be "French-style," a lot is still not known on the Japanese side. "However, the establishment of the JFLI framework has created the opportunity to have informational exchanges several times a year," he says. "These exchanges will undoubtedly produce new research topics." He welcomes the contact with numerous researchers that will result from this organization.

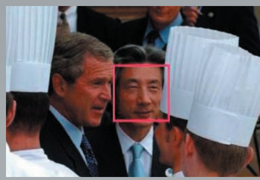
The defining characteristic of the discipline of informatics is that even the single area of graphics or multimedia research involves content that is truly diverse. One example is haptic technology. The word "haptic" means pertaining to the sense of touch. It indicates the reaction force and the feeling of smoothness or roughness when the surface of a hard or soft object is touched. The vibration function of a mobile phone and the bodysonic device in a train simulator (which creates the illusion that the floor is shaking) are the result of developments in haptic technology. This technology that utilizes the sense of touch has applications in the transmission of information to sight-impaired persons.

Some researchers see haptic technology as art. They feel that, for example, a floor that makes a scratching sound when a pen is used to write characters on the floor constitutes a type of artistic expression. The scratchy feel when the characters are written has a psychological effect, producing certain feelings and images in the writer. By fusing the informatics and engineering aspects with artistic and psychological aspects, haptic technology seems likely to become an even more intriguing field of study.

Olfactory sensor

Although still largely unexplored, senses other than touch such as smell





Example showing automatic detection of a specific person's face in an image on the Internet. The more famous the person is, the easier it is to detect that person's face.

Sport, walking / running

Beach, mountain, car

Plane

Face

Results of automatic meaning categorization of images. Each image is automatically assigned a label that indicates its meaning category (sport, plane, mountain, car etc.).

and taste are also the focus of media research. The same smell and taste are perceived differently depending on the color with which they are presented. In this way, smell and taste are very interesting as tools for the transmission of information.

Assimilating one other's research

Approximately 20 years ago when the concept first became known, virtual reality was a technology that enabled people wearing goggles and gloves to feel as if the objects in an image were actually real. Now, however, virtual reality has come to the point at which it could recreate the real world within a computer with exact precision. In actuality, however, there will never be enough time to create every one of the objects that make up our world. This problem would be solved if we had a miraculous scanner that could make exact copies of things, right down to their texture and uses.

A photograph can now copy the scenery exactly as it is, but this does not constitute a virtual reality. The flowers and tables in the photograph are nothing more than collections of dots that do not embody the meaning of a flower or table. For this reason, viewing a collection of photographs and selecting only those that show a flower is easy for a human being with the sense of sight, but it is very difficult for a computer. If there were a way to invest photographs with meaning, computers would also be able to locate photographs containing flowers. This would lead to the technology that could create the miraculous scanner that is able to recognize a flower and recreate it down to its texture and uses.

Prof. Satoh of NII, who will participate in the JFLI's graphics and multimedia research projects, is conducting research on this very topic: finding a way to invest images with meaning. In its earliest stages, research into image recognition focused on numbers and letters, and this technology is already at work in the form of the address recognition scanners in use at the post office and so on. More recently, it has become possible to recognize human faces with considerable accuracy. It is quite possible that Prof. Hirose's virtual reality research and Prof. Satoh's image recognition research will lead to new and groundbreaking advances. Already at Prof. Hirose's laboratory, students are working to develop a "Who's That? System" that will instantly identify the person standing in front of you. In this area of research as well, the encounter between these students and Prof. Sato seems destined to lead to major technical breakthroughs.

Creating a stir in the world of informatics

Although they barely knew one another, Prof. Hirose and Prof. Sato were extremely interested in each other's research, and as time went on their conversations became forums for informational exchange. They saw with their own eyes how an encounter between researchers could provide a major stimulus to both parties.

Informatics is a field that is poised to grow further in the coming years, but as a discipline it is not yet firmly established. "Informatics is a field that does not produce many major discoveries in terms of fundamentals, so it's difficult to get people to see how important it is," says Prof. Hirose. For example, the growth of the Internet was a major event related to informatics, but in the background of this event are factors such as improvements in semiconductor technologies that led to the increasingly compact size and lower prices of personal computers and their increasing use in individual households. This characteristic of informatics has led some to feel that it is best left to companies. "Yet it is when pure research institutions challenge a variety of topics, without getting caught up in the quest for profit, that new fields of research are produced," says Prof. Hirose.

Prof. Adachi wants to achieve further development for informatics in Japan by utilizing his experience as a director involved with management and operation. "NII is not a very large organization. So in order to cover the entire field of informatics, I think we should draw in many other domestic research institutions and create a framework like JFLI." The new collaboration organization of JFLI that emphasizes exchanges among researchers is expected to produce major changes in the discipline of informatics.

(Written by Akiko Ikeda)

NII ESSAY

Do You Open Your Mobile While You're Walking?

Kenro Aihara
Associate Professor, Digital Content and Media Sciences Research Division, NII

We hear much recently about something called the 'Galapagos phenomenon.' The term refers to the way that Japan's technology is developing in its own way, isolated from the rest of the world market, in a manner comparable with the evolution of life forms on those Ecuadorian islands. Mobile phones are a typical example of the phenomenon.

The way that mobiles in Japan are used more as IT terminals than they are in other countries is another peculiarity of our country. On the streets of Japan's towns, you can see many people holding their mobiles in front of their faces, staring intently at the screen and punching keys. And recently, many people are actually to be seen using their mobiles while they walk down the road. It goes without saying, however, that this can lead to all sorts of dangers.

'e-Spaces' Delivering the Information Users Want

The ways in which a mobile can be used are somewhat limited in situations like this, situations in which the screen is not easy to look at. Interaction is possible through sound and vibration, but it is still difficult to effectively get across information in contexts other than phone calls. So what can we do in order to get information as we walk?

Research is underway into providing users in the real world with the information they need in an appropriate format. One such example is the 'e-Space' project being launched by the Ministry of Economy, Trade and Industry (METI), which seeks to install sensors all over towns through which users will be provided with information according to their situation. The development and practical testing of the field service will begin by the end of 2008. This has now led to R&D concerning what information should be transmitted, and how it should be transmitted, to meet the objectives and interests of the users. How can the circumstances and intentions of the users walking through the town be ascertained? How can they be sent the necessary information? It would be feasible to obtain user data from profiles, and the cameras and sensors situated throughout the town. It is also possible to acquire data from the built-in sensors on mobiles and wearable bio-sensors, as well as peripheral sounds and images and the users' access logs. And using this data, it might be possible to predict the circumstances and intentions of users.

Cultivating Technologies that Surpass the Limitations of Mobiles

However, the problem of transmitting the information to the users remains. Even if the information is sent to the user as e-mail, reading mail while you walk is not exactly practical. Interaction that surpasses the limitations of mobiles, something that does not depend on screen displays and key-punching, becomes necessary.

That's what I would like to see from the e-Spaces. In an e-Space, rather than depending entirely upon transmissions sent to mobiles, information aimed at individual users is built in to the space where they are, and the data is supplied by monitors and other devices located on street corners and in stores. There are still unresolved issues, such as the matter of privacy, but this can certainly be described as one challenging approach to the questions of going beyond the mobile phone.

The ideal of the e-Spaces, that unspecified numbers of people will walk through, is that as well as letting users obtain useful information they will enable people to spend their time peacefully, and let those sharing the space neatly harmonize with each other. In order not to encourage the further isolation of individuals, some ingenuity will be required of the design - such as building in information aimed at the environmental aspects, in other words the data transmission to each mobile and the space itself. I hope above all that e-Spaces will be nurtured into a technology that goes beyond the Galapagos phenomenon.

Weaving Information into Knowledge

NII

NII Today No.28, February 2009 (This English language edition of NII Today corresponds to No.42 of the Japanese edition)
Published by: National Institute of Informatics, Research Organization of Information and Systems
Address: National Center of Sciences 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430
Chief editor: Yoh'ichi Tohkura Cover illustration: Makoto Komori Photography: Shuichi Yuri Design: Kotaro Suzuki Production: Sci-Tech Communications Inc.
Contact: Publicity and Dissemination Team, Planning and Promotion Strategy Department TEL:+81-3-4212-2135 FAX:+81-3-4212-2150 e-mail: kouhou@nii.ac.jp
http://www.nii.ac.jp/