National Institute of Informatics News ISSN 1883-1966 (Print) ISSN 1884-0817 (Online)

NII Today

70
June. 2019

National Institute of Informatics News

NII Interview

Looking Into Other Worlds by Designing Light

Selecting light to suit a purpose—applications in food testing and medicine

Imari Sato [Professor, Digital Content and Media Sciences Research Division]

Why Are People Able to See and Recognize Things? Designing Mathematical Models to Find Out

Akihiro Sugimoto [Professor, Digital Content and Media Sciences Research Division]

Uncovering Objects Invisible to Humans Using Computer Vision

Yinqiang Zheng [Associate Professor, Digital Content and Media Sciences Research Division]

Aiming For Practical Application of High-Precision 3D Reconstruction Techniques

Satoshi Ikehata [Assistant Professor, Digital Content and Media Sciences Research Division]

Feature

The Forefront of Computer Vision Research

Capturing the Invisible With Computer Vision

This English language edition NII Today corresponds to No. 84 of the Japanese edition

Looking Into Other Worlds by Designing Light

Selecting light to suit a purpose—applications in food testing and medicine

Imari Sato
Professor, Digital Content and Media Sciences Research Division, National Institute of Informatics /
Professor, School of Multidisciplinary Sciences, The Graduate University for Advanced Studies

Interviewer: **Kyoko Takita**
Deputy Managing Editor, Yomiuri Shimbun, Tokyo Headquarters

Computer vision technologies that analyze objects invisible to the human eye in images, or reproduce the real world in virtual space, are a familiar presence in applications such as facial recognition in digital cameras, movie computer graphics (CG), and 3D maps. Recently, the range of applications has been expanding into areas such as quality testing and medical diagnosis using artificial intelligence (AI) based on machine learning. How will the smart eyes of computers change our lives? I asked NII's Professor Imari Sato, who is working at the forefront of this research.

Takita How has research on computer vision evolved?

Sato It began with a desire to give robots eyes like humans. We humans understand the real world through visual information. Where is the flower? What color is it? Who is this person I have met? Implementing this process of seeing and understanding through a robot's camera is the goal of computer vision.

In the early 1990s, I spent a year abroad in the US at Carnegie Mellon University, where pioneering research on artificial intelligence was in progress, and afterward, I had the opportunity to study there as a visiting scholarship student. During that time, I served as an assistant to Professor Takeo Kanade, director of the world's first robotics institute and became interested in the computer vision technologies that were being developed at the institute, such as autonomous driving, exploration robots, and distance measurement using stereo cameras.

Imari Sato

Real-world analysis technologies based on computer vision are also used in CG in games and movies, and the texture of characters and objects has improved dramatically thanks to developments in modeling technology based on real measurements. This is because the modeling allows "the wealth of apparent differences in the real world," such as the way a particular object looks when the light comes in through a certain window, to be reproduced using a computer. Recently, it has become possible to replicate the real texture of human skin and the like by incorporating technologies that analyze and model the propagation of light on the surface of objects.

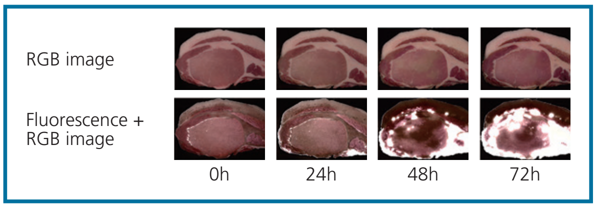
Research on computer vision has expanded from imitating the human eye to reproducing the real world on a computer. Now, however, I am trying to analyze objects that are invisible to the human eye through computer vision.

Takita How does a computer see things that are invisible to the human eye?

Sato Objects reflect sunlight. Humans can only detect the reflected light that is in the visible range. Furthermore, we perceive the world via our photoreceptor cells as three primary colors: red, green, and blue. However, insects can detect ultraviolet light, and snakes have a sensory system that allows them to detect infrared radiation. The human eye cannot distinguish, for example, between a male and female butterfly, but the male absorbs ultraviolet light and so appears black to a butterfly's eyes. The female reflects light, and so it appears white. Also, when water that looks transparent to a human being is observed through near-infrared light, the water absorbs the light, and the brightness varies depending on the water depth.

Takita So, by manipulating light, we can see new worlds?

Sato This is also useful in food testing. To determine the sweetness of a melon, it is possible to visualize the sugar content by inserting a filter that blocks other wavelengths and then observing visible



Photograph 1 Visualization of the degree of meat spoilage using bacteria absorption wavelength (bottom). Areas that appear white are where bacteria are proliferating.

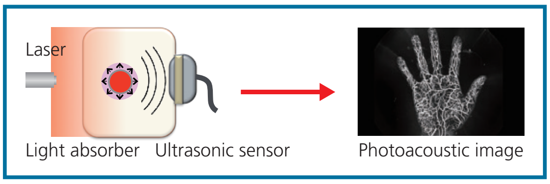
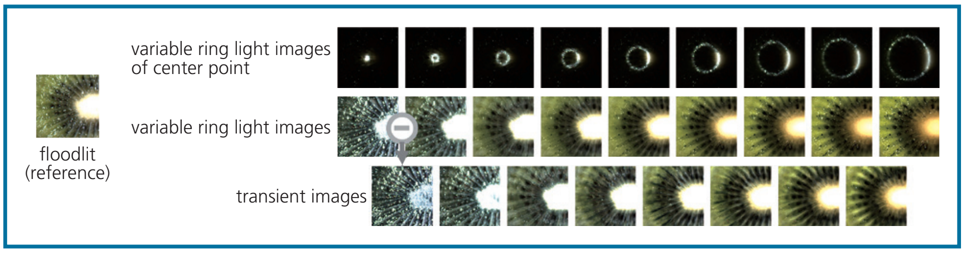
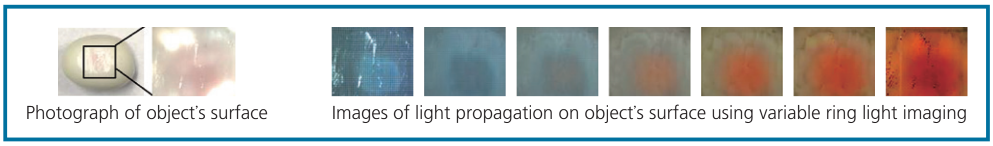


Figure 1 Photoacoustic imaging. Even the capillaries are clearly visible.



Photograph 2 Visualization of light propagation process using ring light (kiwi). Cross-sectional image of kiwi projected using normal lighting (left); appearance under ring light illumination (top row); images visualizing the interior by illuminating kiwi cross-section via the ring light of each radius in the top row (middle row); images sorted according to light propagation distance (bottom row). The further left the image, the shorter the light propagation distance (close to the surface); the further right, the longer the light propagation distance (shows the appearance inside the kiwi).



Photograph 3 Example of visualization of light propagation process (boiled egg). From left to right, the surface shell, the egg white, and the yolk become visible.

light with a wavelength of 676 nm, which correlates with sugar content. To accomplish this, we measure the state of reflection/absorption at various wavelengths, and then if we can determine a wavelength that correlates highly with sugar content, we can just look at that wavelength from then on.

I am interested in finding out which light can be used to obtain the information that we need to know most. I want to engage in image processing that unites sensing and analysis.

I am also looking at fluorescence, which is the emission of light by a substance that has previously absorbed light. For example, when green light is shone on olive oil, the light changes to orange, but when it is shone on sunflower oil, the light stays green. The oils appear to be similar in color, but if you use a green laser pointer, you can soon distinguish between them. The color of fluorescence emitted by Scotch whiskey and other whiskeys is also different, and it is also possible to examine the extent of meat spoilage by using a wavelength that is well absorbed by the proliferating bacteria (Photograph 1). Since the absorbed wavelength and the color of the emitted light differ depending on the living organisms and the material, we are unraveling large amounts of image data using machine learning and are publishing algorithms to determine the optimum wavelength distribution of irradiation light for identification.

Takita Is it possible to examine the inside of foods without cutting or breaking them?

Sato Research aimed at seeing inside food is progressing using ring lights (Photograph 2). When a small ring light is cast on a boiled egg, only images of the surface reflection can be taken, but when a large ring light is used, images containing reflections of the egg yolk inside can be seen (Photograph 3). When the differences between the images taken using a large ring light and those taken using a small ring light are compiled, the layered structure of the object is revealed. It resembles a CT scan. But, when the skin is thick, like a melon's, the light is absorbed by the surface, and the inside cannot be seen. The issue is how deep inside the light can reach.

Takita Could there be applications in medicine?

Sato I am participating in the Cabinet Office's ImPACT Program and working on supporting medical diagnostics using machine learning technology by advancing the analysis of photoacoustic images. I am attempting to apply this technology to cancer diagnostics. The technology makes use of a mechanism in which, when a living body is irradiated with near-infrared light, the light is absorbed, the blood vessels expand, and ultrasonic waves occur (Figure 1). Since the blood vessels around cancer cells are characterized by their length and thickness, I am analyzing and verifying images of blood vessels in an effort to target breast cancer. Light of this wavelength has little effect on the body, so there is no radiation exposure risk like with an X-ray examination. I hope that this research will result in a next-generation medical device.

Takita Could you tell me about the future of computer vision research?

Sato I am working on developing a portable device that allows the analysis to be performed outside of the lab, and I hope to make it usable for non-invasive diagnostics in medicine. In the future, you might be able to check the degree of freshness or flavor of foods in supermarkets with a familiar device such as a smartphone.

(Photography by Yusuke Sato)

A Word from the Interviewer

In the world of astronomy, the new ways in which the universe can now be observed using wavelengths other than visible light have expanded the frontiers of knowledge. It seems that by manipulating invisible light, we will also be able to see other worlds in our own kitchens and refrigerators! In an aging Japan, I hope that this will help to alleviate elderly peoples' concerns about no longer being able to see as well as they used to.

Kyoko Takita

Deputy Managing Editor, Yomiuri Shimbun Tokyo Headquarters
In 1989, graduated from the Faculty of Foreign Studies, Sophia University, and joined the Yomiuri Shimbun. In 2000, graduated from the Graduate School of Journalism, University of California, Berkeley. From 2002, responsible for science and technology policies, IT, space exploration, environment, disasters, etc., in the Science News Department. After serving as Editorial Writer and Science News Department Director, assumed her current role in 2018.

Interview

Why Are People Able to See and Recognize Things? Designing Mathematical Models to Find Out

Images, text, and audio can finally be handled together

Akihiro Sugimoto

Professor, Digital Content and Media Sciences Research Division, The National Institute of Informatics / Professor, School of Multidisciplinary Sciences, The Graduate University for Advanced Studies

Inputting text into a machine (computer) describing a scene you want it to draw and then having the machine output a picture, and having a machine find camouflaged objects hidden in images that are difficult even for humans to discern are both research subjects that NII's Professor Akihiro Sugimoto is pursuing for fun. His motivation is the same: "What does it mean to see things? Why are people able to identify what they see?" To answer these questions, he explores the mechanism by which people recognize things, designs a mathematical model, validates the model by running it on a machine, and then repeats this series of process.

We can do something that was impossible before

"Sheep by sheep, eating grass with a mountain behind tree, another sheep eating grass, the sky has a cloud." When this sentence is input, the machine draws a picture of a flock of sheep in a meadow. This is a typical example of a research area called "text-to-image generation," in which images are generated based on meaning.

Combining this with speech recognition technology allows a person's speech to be turned directly into pictures. Conversely, it is also becoming possible for a machine to recognize an existing picture or photograph and to output text describing the picture or photograph. This requires technology that handles images, text, and audio all together. Once known as "multimedia processing," this has finally become a technological reality. We have now reached the stage of being able to unify the technologies of image recognition, speech recognition, and natural language processing, which developed entirely separately.

Akihiro Sugimoto

To draw the sheep picture correctly, we developed a method for inferring the appropriate location and size (or layout) of each object from the relationships between various objects that appear in the text. Using this method, the computer can draw the sheep in the meadow, rather than floating in the sky. A deep learning based network is used to infer the layout.

First, the text to be input and a picture drawn according to that text (i.e., the ground-truth data) are prepared. The ground-truth data concerning objects and the layout of the objects can be obtained using image recognition technology that detects and recognizes various objects in an image, and this ground-truth data are used to train the network. Combining this network with an existing method for extracting words from texts and deriving the relationships between the words, which is known as a "scene graph," makes it possible to infer the layout from the input text (Figure 1).

I am also working on other subjects in image generation using deep learning. For example, we can give a machine an image of flowers and a written instruction such as "Make only the white flowers yellow," and have the machine identify and redraw only the corresponding flowers among all the flowers in the input image. In this case, the machine needs to identify the foreground and background of the image separately, so two deep neural networks for discrimination are prepared. This enables the device to redraw only the white flowers identified as the foreground and leave the background as it is. Conversely, it is also possible to redraw the background only.

I have also worked on "image style transfer." If Van Gogh or

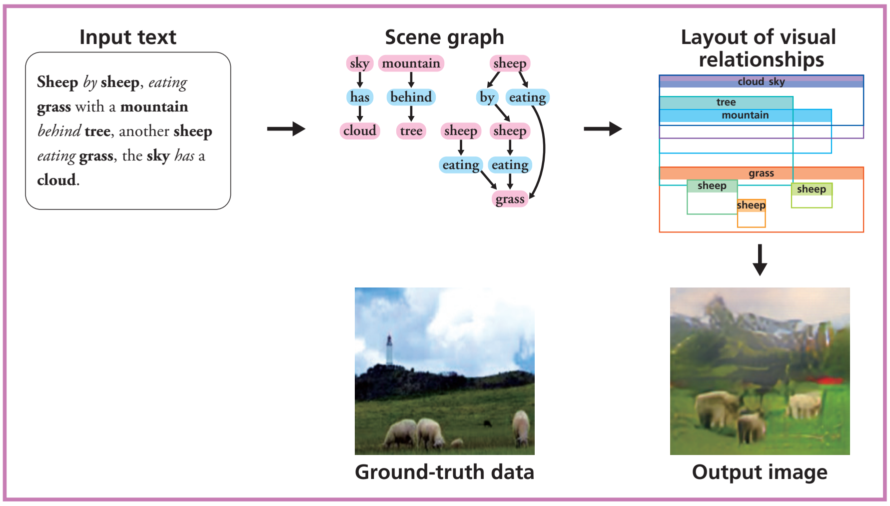


Figure 1 Example of image generation from text. Using a deep neural network, the appropriate layouts of objects in images are learned, and a picture is drawn from an input text.

Hokusai were alive today, how would they draw modern scenes? To find out, I fed photographs, pictures, video scenes, and painting styles into a machine and attempted to convert them into the styles of Van Gogh or Hokusai.

Various attempts at image style transfer have been proposed in the past, but my idea was to prepare a deep learning based network for recognizing content characteristics, such as scenery, and a separate network for learning Van Gogh's artistic style from his paintings, in other words, his particular style characteristics. This is because the roles of recognizing scenery and learning painting styles are different.

Combining the two networks not only makes possible a well-balanced image style transfer process but also makes it possible to make flexible adjustments, such as accentuating the painting style or accentuating the original picture. Image style transfer can be performed by feeding numerous Van Gogh or Hokusai paintings into just one network, but that way does not allow the flexible adjustments to resulting pictures.

The point is grasping the true nature of phenomena

As you can see from the image generation and image style transfer examples, when we are making a machine do something, we need to design models that consider what functions are required and how they should be combined. When this is described using mathematical expressions, it is called a mathematical model.

The true nature of various phenomena can be expressed using mathematical models. This type of modeling is the whole point of mathematical engineering—racking one's brain to design a model. As long as you succeed in that, the only remaining serious problem is how to run the model, but it is something that anyone can do.

When designing a model that recognizes image contents, I think about what humans do. For example, how can I make a machine recognize camouflaged objects that even humans cannot see right away? Without a prior explanation, human beings do not readily notice objects hidden in images, but if they are told that an object is hidden in the image beforehand, they will find it relatively quickly.

Therefore, when we added a "function that determines whether there is an object" to a deep neural network that detects salient objects from images, the machine was able to detect camouflaged objects hidden in the images.

We are also developing a learning based network that detects objects drawing peoples' gaze in still images or videos. This allows the machine to detect and recognize objects even if they are overlapping with each other. Since each object stands out in different ways, we can make use of that to detect the most eye-catching object first and move on to detecting the next object.

"Seeing" and "understanding" are profound research topics

My motivation for working on this line of research came from my curiosity about what it means to see things and why humans can recognize what they see. I began thinking about this when I was very young, and I've remained curious about it ever since. I explore how humans perceive things and how they recognize them, and then turn it into designing a model, which I evaluate on a machine. If this yields promising results, even slowly, it should be possible to uncover the mechanism by which people see and recognize objects.

Once models are built, they can be improved by accumulating experimental results. This is what science is all about. By preparing large amounts of images and their ground-truth data by human recognition tests, and then feeding these data into a deep neural network to train, we have become able to produce results that surpass the human recognition rate.

We would be foolish not to make use of this. However, improving the model haphazardly, without a design concept, entails the risk of just starting a competition in which money is thrown at increasing the computer power and preparing vast amounts of data. This may improve the performance level, but no insightful knowledge would be gained.

I have pursued a variety of research, but the question of "Why can humans recognize what they see?" is genuinely profound and little understood. I intend to continue following my interests without forgetting the curiosity I had as a child.

(Interview/Report by Nobuyuki Yajima
Photography by Yusuke Sato)

A Word from the Interviewer

Researchers are expected to have both the curiosity of a child and the calm, composed manner of a scientist or engineer. The curiosity we have as children tends to disappear as we become adults, but Professor Sugimoto has held onto his. At the same time, he maintains the scientist's approach to building a mathematical model, validating the results, and improving the model—always with a design concept in mind. Relying on the computer power and thinking that everything is fine as long as the performance becomes better is childish and does not interest Professor Sugimoto.

Nobuyuki Yajima

Senior Researcher, Nikkei BP Intelligence Group, Nikkei Business Publications, Inc.

Born in 1960. Studied mathematics at university with the aim of becoming a computer engineer, but joined Nikkei McGraw-Hill, Inc. (now, Nikkei Business Publications, Inc.) in 1985 and became a reporter for the *Nikkei Computer* magazine. Worked as editor-in-chief of *Nikkei Computer* from 2009. Assumed his present post in 2016.

Uncovering Objects Invisible to Humans Using Computer Vision

Developing software that realizes the performance of an expensive, specialized camera using a general-purpose silicon camera

Yinqiang Zheng

Associate Professor, Digital Content and Media Sciences Research Division, National Institute of Informatics / Associate Professor, School of Multidisciplinary Sciences, The Graduate University for Advanced Studies

Recently, with rapid progress in image recognition technology and the popularization of AI terms such as “machine learning” and “deep learning,” interest from industry regarding information sensing technology in the field of computer vision is growing. Silicon cameras are the most versatile “computer vision” tools and NII’s Associate Professor Yinqiang Zheng is working to enhance their accuracy via software improvements. I asked him about what today’s computer vision is capable of seeing and how it will develop in the future.

Clear video recording in dark conditions using an ordinary camera

—— What have you been researching recently?

Zheng Human eyesight can only perceive visible light and is inadequate in terms of “sensing information.” Human beings can see very little, particularly in the dark or underwater. However, computer vision, such as an electron-multiplying charge-coupled device (EMCCD) camera can display a variety of information types as images, even in dark conditions. Costing several million yen each, EMCCD cameras are extremely expensive, so they are not readily available to ordinary companies or people. Therefore, we are working on software that will allow us to implement a multiplying function for taking clear images at night or in dark conditions, even with the ordinary silicon camera mounted in a smartphone. Our primary goal is to create software capable of maximizing the capability of an average camera in dark conditions.

—— What method are you using in your research?

Zheng First, we prepare two (RGB[1] format) cameras and take videos in the same place, using the same timing. This is done to output images taken at the same time and location as pairs consisting of a “bright image” and a “dark image.” These videos are essential (Figure 1). The videos are taken by splitting the light from a single source via a relay lens and a beam splitter before it reaches the sensor of each camera. Additionally, a neutral density (ND) filter is fitted to the camera taking the “dark image” to reduce the amount of light entering the lens. With a video frame rate of 30 fps and a recording time of about 5 hours, approximately 200 gigabytes of data are obtained in one recording.

Then, we use this paired dataset to conduct deep learning repeatedly. For the deep learning environment, we employ a neural network for segmentation called “U-Net,” and we generate models until the machine can output a bright image when a dark image (frame) is input. Implementing algorithms acquired in this way allows us to obtain beautiful, high-resolution images using an ordinary silicon camera. We use the open-source “PyTorch” machine learning library as the deep learning framework.

—— What applications do you expect this technology to have?

Zheng There are three main applications. The first is in surveillance cameras. There are already surveillance cameras capable of filming at night, such as IR cameras that use near-infrared radiation, but these are made for industrial use and are not geared towards the general public. Installing a function that enables filming in dark conditions into an ordinary camera will allow it to be used as a surveillance camera. Zooming in and out is now also possible, and there are high hopes for practical application of this technology.

The second application is in ordinary smartphones. Currently, people around the world upload vast numbers of images to social networks every day, and the desire to be able to take

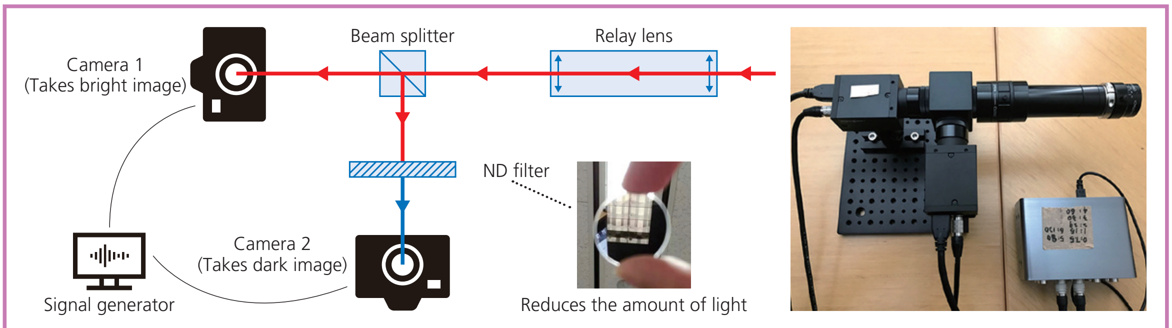


Figure 1
A camera system that takes a “bright image” and a “dark image” at the same time and location (right image), and the path of light in the system (left image).

beautiful photos in dark conditions keeps on growing. The software-based multiplying function could meet this need quickly and inexpensively. However, smartphone users won't be satisfied unless images can be output as soon as they were taken, so making the technology available in real time is one challenge on the path towards practical application.

The third application is in the IoT, which includes drones and self-driving cars. Bright lights are currently necessary if one wishes to fly a drone or run a self-driving car at night, but installing powerful lamps can result in draining the available electric power. In response to this, our latest technology allows power consumption to be drastically reduced.

One other challenge is that the sensor gradation is currently 8 bit (256 gradations), but we would like to increase this a little. We are also investigating ways to increase the sensitivity using a CMY[2] filter. This would also increase the accuracy of the algorithm and increase the likelihood of obtaining a clearer image, thus bringing us closer to our goal of "maximizing the capability of the camera."

Application in optical fiber cable and solar panel inspection

—— Could similar experiments be done using still images?

Zheng Yes, of course, we are doing that too. So far, I have talked about our research aimed at obtaining clear images in dark conditions, but we are also working on bringing the performance of a silicon camera closer to that of an indium gallium arsenide (InGaAs) camera that realizes short-wave infrared (SWIR) imaging, and for that we are using still images. The experimental method is almost the same, but an RGB-blocking filter is used.

—— Does that mean that RGB images are not needed to bring the performance of a silicon camera closer to that of an InGaAs camera?

Zheng Yes, that's right. To make things that the human eye cannot see visible, it is, conversely, better to remove RGB information. When RGB is deliberately excluded, things that cannot be seen in visible light emerge.

—— Where would this research be most useful?

Zheng The most promising field is the inspection of optical fiber cables laid on the seabed. Generally, silicon cameras are designed to capture wavelengths between 400 and 1,000 nm, but wavelengths of approximately 1,200 nm are used to inspect for damage to optical fiber cables. In our experiments, we succeeded in examining the state of a cable by applying this filter, but the images are still inadequate, so it will be necessary to increase accuracy levels by accumulating a little more training.

I think that this research could also be applied to checking solar panels. Scanning parts that cannot be approached and checked by human beings, such as solar panels or optical fiber cables, requires a sensor with a performance equivalent to an InGaAs camera, rather than an RGB sensor, so we are aiming for a performance that will supersede the existing expensive cameras.

Computer vision and human vision play different roles

—— I understand that you have also been involved in developing a 3D camera for shooting videos underwater. Do you think that computer vision has surpassed human vision?

Zheng Yes, it definitely has. Like I said at the beginning, human eyesight is extremely limited. Computer vision is clearly superior in terms of sensing information. However, that does not mean that humans are inferior to machines. Human eyes not only detect information but also instantly relay it to the brain, which "understands" the information that has been captured as sight. This is an area in which machines are no match at all. I think it is probably best that machines be given the chores they're best suited for and humans be encouraged to focus on areas where we excel.

—— What motivates you to continue this research?

I majored in 3D reconstruction at the Tokyo Institute of Technology and then joined a company where I researched computer vision and algorithms. It was during that time that I started to think, "I want to use the best camera to make things that are invisible visible, and I'm going to have to develop it myself." There are many challenges, but especially now, in this data-driven age, I want to get closer to achieving my goal of obtaining beautiful images, but not with expensive, specialized cameras and filters. Instead, I want to use ordinary cameras that are available to anyone.

(Interview/Report by Akiko Gomi
Photography by Takuya Furusue)

Notes

[1]The three primary colors of light, R: Red, G: Green, B: Blue.
[2]The three primary colors expressed as reflection/absorption of light, C: Cyan, M: Magenta, Y: Yellow.

Yinqiang Zheng

Aiming For Practical Application of High-Precision 3D Reconstruction Techniques

Combining accumulated knowledge and deep learning for applications in real estate and medicine

Satoshi Ikehata

Assistant Professor, Digital Content and Media Sciences Research Division, National Institute of Informatics /
Assistant Professor, School of Multidisciplinary Sciences, The Graduate University for Advanced Studies

We perceive depth by using retinal images to construct a 3D world in our brain. What mathematical model can represent this process? This is a question that interests NII's Assistant Professor Satoshi Ikehata, a computer vision researcher. Recently, he has been working on 3D reconstruction using the photometric stereo method, which is a method of performing 3D reconstruction based on shading. He is currently aiming to apply this in various fields by making use of deep learning. I asked him about the potential of this technology.

Reconstructing 3D information from 2D images

During his master's program, Assistant Professor Ikehata worked on creating 3D building models using a "multi-view stereo method" that constructs models using "corresponding points" (physically identical points) in images from multiple cameras. His goal was to use photographs to reconstruct, in 3D, not only static buildings but also moving objects, such as the bride and groom at a wedding.

Upon entering his doctoral program, he changed the theme of his research to the photometric stereo method. Unlike multi-view stereo, photometric stereo uses only one camera. Instead of multiple cameras, multiple images are taken while the object is illuminated from different directions, and 3D shapes are reconstructed based on the resulting multiple shading patterns.

The simplest reflectance model used in 3D reconstruction is the Lambertian reflectance model, which deals with rough surfaces that exhibit diffuse reflection. For example, the surface of a plaster figure can be approximated by Lambertian reflectance. However, dealing with specular reflection from, for example, a polished table or a plastic or metal surface is difficult.

Satoshi Ikehata

With those points in mind, Assistant Professor Ikehata proposed a groundbreaking method for handling non-Lambertian materials in the photometric stereo problem. First, in 2012, he proposed a technique for dealing with specular reflections observed only in limited areas on an image (called "sparseness") as outliers and excluding their effects. This technique is still used today because it is practical and easy to implement.

However, sparseness cannot be assumed for materials that produce specular reflection widely, such as wood. In such cases, an accurate physical model representing specular reflection is necessary. While this generally requires dealing with a nonlinear nonconvex problem[1], which is extremely difficult to solve, this problem can also be simplified by putting in place conditions that limit the target space. In 2014, Assistant Professor Ikehata discovered that the problem could be replaced with a more straightforward task by assuming that most materials in the natural world exhibit isotropic reflection[2]. He also showed that it is theoretically possible to deal with any material using this technique (convex optimization).

Objects that can be handled in this way have gradually increased, but a challenge remains. That challenge is shadows. Photometric stereo uses the intensity of "shade." However, a "shadow" is an area formed on a surface that light has not reached because some obstacle has blocked it, which means that there is no information to handle. Bowl shapes, in particular, block light.

In 2018, Assistant Professor Ikehata proposed, for the first time ever, using photometric stereo for nonconvex objects utilizing deep learning. In deep learning, it is generally necessary to provide a fixed number of input images. However, in photometric stereo, that number is not set in advance. Therefore, Assistant Professor Ikehata came up with an original idea regarding the input data. In his method, the inputs are all projected onto one image, called an "observation map," that shows how the light was observed. The photometric stereo method is then performed by training a machine using these maps, which compress and encode the light information.

As a result, photometric stereo can represent correct shapes more precisely than a 3D scanner. In the future, Assistant Professor Ikehata would also like to try generating surface detail using images taken from multiple viewpoints.

"Human beings can reconstruct correct shapes from one image. This is because they make automatic assumptions about which direction the light is coming from. Therefore, it should be possible to train a computer with the knowledge needed to make such assumptions using deep learning. I want to combine that with a multi-view method."

Possible real estate application

Assistant Professor Ikehata is currently focusing on applying 3D reconstruction techniques to real estate, which is a project he began working on during his post-doc days in the United States. The use of IT is growing in the US real estate industry, but there are many old properties whose floor plans have not been adequately maintained and so the demand for modeling is high.

Capturing the interior of a property using a commercial laser scanner provides a 3D point cloud. However, a point cloud alone does not provide the scene context, such as the building floor plan or the ceiling heights. Despite these problems, Assistant Professor Ikehata succeeded not only in creating 3D models from point clouds but also in creating property information that can be linked to computer-aided design (CAD) by providing semantic information to the model[3].

There are two key points to consider. One is a "structure graph," which is a novel 3D information representation proposed by Assistant Professor Ikehata. This is a graph representation in which nodes correspond to the structural elements of the property, such as rooms, walls, and ceilings, while edges correspond to their containment/connection relationships. Since each node contains all of the geometric and topology (spatial position and connection relationships) information required for 3D reconstruction, a structure graph can be guaranteed to produce a 3D model that preserves the connection relationships between walls and has no gaps between planes.

The other key point is the "structure grammar" used in the structure graph reconstruction. Since buildings are designed to be easy for humans to live in, human intentions intervene in their design. Based on these intentions, when the structure grammar—a set of algorithms such as "Disassemble the room into walls and a floor"—is applied, the 3D point cloud is automatically divided into the structural elements inside the building and reconstructed (Figure 1).

The most significant advantage of structure graph representation is its high degree of freedom. For example, a model without a ceiling can be created instantly by deleting only the ceiling nodes from the graph. The number of rooms can be reduced or walls can be added at will. Combining this with deep learning, it is also possible to distinguish which rooms are which. Floor plans can be created, and the shortest escape routes can be calculated automatically.

There is also a function called "Inverse-CAD" that allows the finished 3D model to be easily edited, which was only made possible by the creation of a graphical structure where each node possesses 3D information. This kind of 3D reconstruction has never existed before.

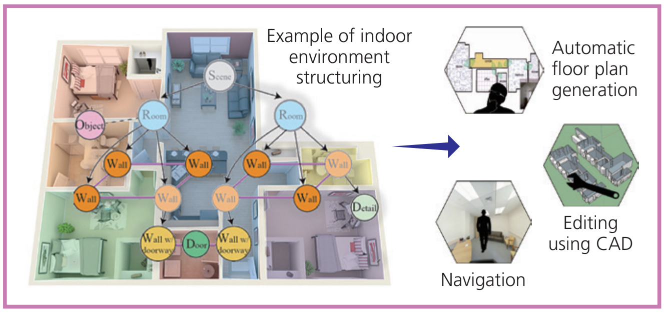


Figure 1 Structured indoor 3D model and its applications

Bridge-building between humans and machines

Assistant Professor Ikehata is also working on reconstructing 3D information from panoramic photographs taken using a smartphone. Currently, 3D reconstruction from one panoramic photograph is a difficult indeterminate problem, but one that many researchers are already tackling, and Assistant Professor Ikehata thinks that it could be achieved within a year.

Applications in medicine, such as 3D reconstructions using microscopic or radiographic images, are also progressing. In particular, expectations for the application of deep learning in 3D reconstruction are growing. "I think that deep learning can be used in areas where 3D reconstruction is difficult using conventional techniques but easy for humans."

Each technique has its strengths and weaknesses. "It is important that the appropriate technique be used for the appropriate problem. That's what is interesting," says Assistant Professor Ikehata, who thinks that we have reached the point of being able to combine all of the separately developed techniques.

"Human beings use a variety of information in order to perceive depth, including texture and binocular disparity, motion parallax, and shading information. Computer vision also uses different techniques. The application of deep learning is blurring the boundaries between these techniques, but combining knowledge that has been physically accumulated with deep learning is especially interesting because it allows solutions to be narrowed down using our knowledge of physics, rather than just correspondence between inputs and outputs."

While, technically, various things have become possible, the practical side is lacking. Once these technologies are applied in practice, they are expected to develop even more substantially.

Initially majoring in perceptual psychology, Assistant Professor Ikehata is also interested in differences between the human perspective and the computer's perspective. For example, *trompe l'oeil* images are only visible to humans, and pursuing this leads to research on how human beings see the world. Problems with multiple interpretations also exist in 3D reconstruction.

"Inadequacies in computer vision can be overcome by looking at human beings. Research that bridges the human aspects and information processing may be possible."

(Interview/Report by Kazumichi Moriyama
Photography by Takuya Furusue)

Note

[1]Finding the minimum value of a convex function, such as a quadratic function, that is convex downward is relatively easy, but solving a nonconvex function is difficult.
[2]A type of reflection where rotating the object does not change the brightness of the reflected light.
[3]https://www.youtube.com/watch?v=RZU8w3uvenU

NEWS 1

Hosting NII Week: Introducing NII's research and initiatives in one shot

The National Institute of Informatics (NII) designated the week from May 27 to June 01 as "NII Week" and held a variety of academic events at the National Center of Sciences (Chiyoda-ku, Tokyo).

Japan Open Science Summit 2019 (JOSS2019)


Panel discussion at JOSS2019

The Japan Open Science Summit 2019 (JOSS2019), which is an event that disseminates the latest open science trends in Japan and the world, was held on May 27 and 28. It was hosted jointly with the Japan Science and Technology Agency (JST), the National Institute for Materials Science (NIMS), the National Institute of Science and Technology Policy (NISTEP), the National Institute of Information and Communications Technology (NICT), and the Academic Repository Network (Re*poN).

Open science is a new approach to research that promotes publishing and sharing, not only research papers but also research data and software, with society at large. Discussions about the management and sharing of research data, in particular, have become animated recently, and there is a growing recognition that open science is a framework that academia should implement. NII established the Research Center for Open Science and Data Platform (RCOS) (Center Director: Professor Kazutsuna Yamaji, Digital Content and Media Sciences Research Division, NII) in April 2017, and is constructing and operating academic infrastructure to support open science activities at universities and research institutions.

JOSS2019 was attended by various interested parties such as policymakers, businesses, NPO staff, and citizen scientists, as well as researchers in a wide range of areas from universities to private companies, research supporters such as university librarians and research administrators (URA), and IT infrastructure researchers and developers. Enthusiastic sessions unfolded on topics such as citizen science, government policy, and data management at libraries and universities.

Academic Information Infrastructure Open Forum 2019


NII's Deputy Director General and Director of Cyber Science Infrastructure Development Department, Shigeo Urushidani, giving the keynote address

The Academic Information Infrastructure Open Forum 2019 was held on May 29 and 30. The aim of this forum is to share a detailed image of the education and research environments at universities and research institutions made possible by NII's Science Information Network (SINET5) with relevant parties on a timely basis, and to advance academic information infrastructure together with its users.

The sessions held over the two days were divided into tracks, including a cloud computing track, a SINET track, and an authentication track. Lectures and lively discussions about the future of academic information infrastructure ensued at each venue. In the plenary track on the first day, NII's Deputy Director General and Director of Cyber Science Infrastructure Development Department, Shigeo Urushidani, gave the keynote address titled "NII's Service Expansion from 2020," in which he described the next-generation academic information platform based on SINET and NII's research data cloud. Next, there was a discussion between participants and NII researchers on the topic of "Towards Co-Designed/Co-Created Academic Information Infrastructure." The participants put forward various opinions and questions, including, "I would like to see the development of services that younger generations, such as high school students, can use" and "What should be done to give researchers an understanding of research data management?"

Open House 2019


NII Computer Science Park


NII Research 100 finale


Professor Ullman giving a special lecture


Kengo Sakurada giving the keynote address

Open House 2019 (public exhibition/presentation of research results), an event to inform the wider public about NII's research findings and business activities, was held for two days on May 31 and June 1.

On day one, Professor Emeritus Jeffrey D. Ullman of Stanford University, a leading authority on database research, was invited to give a special lecture titled "Computing an Enormous Join of Relations." At the opening ceremony that followed, NII's Director-General Masaru Kitsuregawa gave an activity report and presented NII's main initiatives in its two pillars of "research" and "service." There were two keynote addresses. First, from the business world, Kengo Sakurada (Chair of the Japan Association of Corporate Executives/Group CEO of Sompo Holdings, Inc.) talked about such topics as digital (transformation) strategies in the Sompo Group. Then, from NII, Professor Ken Satoh gave a talk titled "The Application of AI to Law: Current State and Future Trends."

On day two, many children and their parents gathered at NII Computer Science Park to experience a programming approach without using computers. Activities included learning how to make objects move through a number game and actually making robots move. This allowed the participants to experience programming and informatics while having fun.

There was also a lot of excitement at the popular "NII Research 100," which is coming to an end this year. At this event, 10 NII researchers each presented 10 research results, making a total of 100 research results. Celebrity and engineer Ayaka Ikezawa appeared as MC again this year, and together with NII's Professor Hideaki Takeda and Associate Professor Ikki Ohmukai, brought additional excitement to the event.

In the poster presentation and demonstration/experience corner, NII researchers spent the entire day in front of their posters answering visitor's questions about their research. NII researchers also acted as guides to give visitors tours of the poster presentation and demonstration/experience corner, which proved to be very popular as well.

Information Session on the Department of Informatics, The Graduate University for Advanced Studies

An information session on the Department of Informatics in the School of Multidisciplinary Sciences, The Graduate University for Advanced Studies, was held in conjunction with Open House. NII, which is involved with The Graduate University for Advanced Studies, established the Department of Informatics in the School of Multidisciplinary Sciences. The department provides graduate school education through five5-year and 3-year doctoral programs. At the information session, as well as an overview of the Department of Informatics and an explanation of how to apply, students enrolled in the department talked on topics such as their own student life and research, and features of the department.

NEWS 2

2019 MEXT Commendation for Science and Technology awarded to Associate Professor Megumi Kaneko and GakuNin Development Group

The awards ceremony for the 2019 Commendation for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology (MEXT), which recognizes individuals who have achieved outstanding results in research, development, and promotion of understanding in science and technology, was held in the Ministry auditorium on April 17.

From NII, Associate Professor Megumi Kaneko of the Information Systems Architecture Science Research Division was awarded the Young Scientists' Prize for research results that increase the usage efficiency of wireless resources in mobile systems. In addition, a group made up of Yasuo Okabe (Professor, Kyoto University/Visiting Professor, NII), Ken Nishimura (Project Researcher, NII), Hiroyuki Sato (Associate Professor, University of Tokyo), Hideaki Goto (Associate Professor, Tohoku University/Visiting Associate Professor, NII), and Noboru Sonehara (Professor, Tsuda University/Visiting Professor and Professor Emeritus, NII) was awarded the Science and Technology Prize (Development Category) for development of the Academic Access Management Federation (GakuNin) for inter-university cooperation.

With beaming smiles, having been awarded their certificates of commendation, Associate Professor Kaneko and the other winners looked newly determined to advance their research even further.

NEWS 3

Wanted: People to discover graphs that can efficiently connect one million CPUs!

Competition to find future supercomputer network configuration

NII has launched its Graph Golf 2019 competition. In this competition, complex network configurations used in supercomputers are replaced by simple graphs, and entrants compete to find graphs with simple structures that will lead to the efficient design of networks inside and between CPU chips.

This year, a large graph with one million vertices was added as a new problem. Submissions will be accepted via the dedicated website (http:// research.nii.ac.jp/graphgolf) until October 14. Outstanding graph finders will receive their awards at CANDAR 2019, an international symposium on computing and networking to be held in November in Nagasaki City.

Agenda

2019 Public Lectures "The Forefront of Informatics" to Start in July

The 2019 program has been decided for the Public Lectures "The Forefront of Informatics," in which NII researchers explain the forefront of various subjects relating to informatics to the general public. For details and registration, please visit NII's website via the link below.
https://www.nii.ac.jp/event/shimin/

1st Lecture July 2, 2019 (Tue)
Speaker : Yu Yokoi (Assistant Professor, Principles of Informatics Research Division)
Theme : Making Everyone Happy!? The Mathematics and Computation of Matching
—How to Determine Smart Allocations

2nd Lecture September 10 (Tue)
Speaker : Satoshi Ikehata (Assistant Professor, Digital Content and Media Sciences Research Division)
Theme : Amazing! 3D Sensing
—Evolving Computer Vision

3rd Lecture November 7 (Thu)
Speaker : Ichiro Hasuo (Associate Professor, Information Systems Architecture Science Research Division / Director, Global Research Center for Systems Design and Mathematics)
Theme : Introduction to Theoretical Computer Science: Between Finite and Infinite
—From Mathematical Theory to AI and Autonomous Driving

4th Lecture January 21, 2020 (Tue)
Speaker : Satoshi Iwamoto (Professor, Research Center for Advanced Science and Technology, University of Tokyo / Research Member, Science of Hybrid Quantum Systems in Scientific Research on Innovative Areas)
Theme : Manipulating Light Using Topology
—Does Light Distinguish Between a Ball and Doughnut?

SNS

"Hey, this is great!" Hottest articles on Facebook and Twitter (March 2019-May 2019)

* Some text edited/omitted.

National Institute of Informatics, NII (official) Facebook
www.facebook.com/jouhouken/

The MCs for "NII Research 100" on day two of Open House will be Professor Hideaki Takeda, Associate Professor Ikki Ohmukai, and Ayaka Ikezawa. 10 NII researchers each present 10 research results—a total of 100 research results—in this event, which will be held for the final time this year!! (05/17/2019)

National Institute of Informatics, NII (official) Twitter
@jouhouken

NII Computer Science Park
"Learn About the Structure of Computer Systems With Digi-Comp II"
Digi-Comp II is an educational computer kit made of wood that can perform basic arithmetic operations using marbles. Learn about information representation in computing and the structure of calculations while having fun! (05/05/2019)

Bit on Twitter! Twitter
@NII_Bit

The MCs for "NII Research 100" on day two of Open House will be Professor Hideaki Takeda (@takechan2000), Associate Professor Ikki Ohmukai (@i2k), and Ayaka Ikezawa (@ikeay). 10 NII researchers each present 10 research results—a total of 100 research results—in this event, which will be held for the final time this year!! (05/16/2019)

Essay

"NII Research 100" and Passion for Research

Hideaki Takeda

Professor, Principles of Informatics Research Division, National Institute of Informatics
Professor, School of Multidisciplinary Sciences, The Graduate University for Advanced Studies

Professor Takeda as MC at NII Research 100. Appearing in a white coat and "cat ears" that move with brain waves (June 2019).

NII Research 100 has been held during NII Open House for five years, from 2015 to 2019. At this event, 10 NII researchers each introduce 10 items of their own research. Ten people multiplied by 10 items of research comes to a total of 100 research items. You might think, "I cannot listen to 100 research presentations!" but the event has an important rule. Each person has just 7 minutes and 30 seconds of presentation time. In other words, only 45 seconds for each piece of research. Ordinarily, presenting research in such a short amount of time is impossible, but making the researchers give presentations under such extraordinary conditions is the real intention behind this event.

Generally speaking, if you let a researcher talk about their research, they will speak for an entire evening. Giving them seven-and-a-half minutes to talk about 10 items of research causes them to get frustrated due to the marvelous combination of their desire to speak and the looming time constraint. As a result, they unintentionally expose their genuine true opinions and personalities.

NII Research 100 continued for 5 years, which means that a total of 50 researchers appeared on the stage. I participated either as an MC or a research presenter during this time, which means that I listened to 500 pieces of research from 50 NII researchers (although one of them was me!). NII is a small research institute, but its researchers are diverse. The scope of my research is relatively broad (I describe myself as an omnivorous researcher on my Twitter account), but even so, there is a lot of research that I know nothing about and I am not interested in. Nevertheless, I was always genuinely moved by the passion of the researchers for their investigations, which the superfast barrage of 100 presentations gave a glimpse of. It was amazing to realize that what we call "research" is made up of researchers' passion for their subject.

Incidentally, the style of presentation of NII Research 100 is not original but was created by an enterprise called NicoNico Gakkai Beta. Inspired by NicoNico Douga, NicoNico Gakkai Beta was started by Koichiro Eto as a place to collect "research" in a broad sense that goes beyond the existing academic framework. A curious coincidence meant that I participated in NicoNico Gakkai Beta from its early days. This was a great experience for me as I encountered a research world that I would never have come across normally. (Would an information researcher normally meet a doctor researching poop or an amateur researcher studying mushrooms?). This has been useful, directly and indirectly, in my own research, and I think it has successfully broken down the boundaries I had regarding research. Information is ubiquitous throughout the world, and in that sense, the whole world is a subject of informatics research.

Well, that's enough of my nonsense. If any readers are interested, please take a look at the NII Research 100 videos or the original NicoNico Gakkai Beta videos. They are all available on YouTube[1] or Niconico[2].

Note

[1] https://www.youtube.com/user/jyouhougaku
[2] https://ch.nicovideo.jp/jouhouken

Future Schedule

July 02 | "The Forefront of Informatics" 1st Public Lecture "Making Everyone Happy!? The Mathematics and Computation of Matching—How to Determine Smart Allocations" (Speaker: Assistant Professor Yu Yokoi, Principles of Informatics Research Division)For registration, go to **https://www.nii.ac.jp/event/shimin/**

July 13 | 2nd Karuizawa Saturday Salon "The Origins of the Universe and The Birth of Stars" (Speaker: Hitoshi Murayama, Former Director, Kavli Institute for the Physics and Mathematics of the Universe, University of Tokyo / Professor, University of California, Berkeley)

August 07–08 | Children's Day for Visiting Kasumigaseki (exhibitor). For registration, go to **https://www.nii.ac.jp/event/2019/0807.html**

September 10 | "The Forefront of Informatics" 2nd Public Lecture "Amazing! 3D Sensing—Evolving Computer Vision" (Speaker: Assistant Professor Satoshi Ikehata, Digital Content and Media Sciences Research Division)

Notes on cover illustration Notice the apples in the robot's hands. They appear to be perfect, but looking through the computer's eyes (the screen), we can see that the apple on the left is rotten. The latest computer vision not only mimics human vision but also reveals worlds that are invisible to the human eye by making use of light and deep learning.

Weaving Information into Knowledge

NII

National Institute of Informatics News [NII Today]
No. 70 June 2019
[This English language edition NII Today corresponds to No. 84 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
Publisher | Masaru Kitsuregawa **Editorial Supervisor** | Ichiro Satoh
Cover illustration | Toshiya Shirotani **Copy Editor** | Madoka Tainaka
Production | MATZDA OFFICE CO., LTD., Sci-Tech Communications Inc.
Contact | Publicity Team, Planning Division, General Affairs Department
TEL | +81-3-4212-2028 **FAX** | +81-3-4212-2150 **E-mail** | kouhou@nii.ac.jp

Bit (NII Character)

https://www.nii.ac.jp/en/about/publication/today/