Interview

Invited Article

On Applying Signal Processing to Computational Art History: An Interview

Park Doing and C. Richard Johnson, Jr.

Park and Rick sit across a desk facing each other in Rick's office in Cornell's Engineering College. We see Park's back and Rick's face. They are dressed in winter clothes. Overseeing the proceedings, peering over Rick's shoulder at Park, is a six foot tall Terra Cotta Warrior that has made the journey from China.

The view is into Rick's office with a small window in the back wall offering a view of barren tree limbs and falling snow. Rick's office is situated between closed doors to offices on the same hall. The office on the left houses a professor expert in information theory. On the right a professor specializing in medical image processing.

Along the left wall in Rick's office stands a bookcase with treatises, dissertations, books, and volumes about control systems and signal processing and art history and conservation. The top of the bookshelves displays pictures of Rick's graduate students. Among the awards advertised on the wall above the bookcase is Rick's prized Eagle Scout certificate. The warrior is in the back left corner of the office.

Along the right wall of the office hang three full-size prints of Van Gogh's "Bedroom": color, false-color infrared, and raking light.

Park has been observing Rick for years, with Rick luring him along. Everyone, including Rick, knows Park has written about 'cultural' battles between physicists and biologists (not to mention technicians and administrators) in a particle physics laboratory, and about interactions between experts and 'lay persons' with regard to issues of science and public welfare. Rick had previously explained to Park how he was entering the world of art history, museums, curators, and conservators, and how much he was learning in the process about approaches (some successful) to cross-disciplinary research. Imagining the potential for insights from an expert observer, Rick offered Park access to a front row seat for viewing a real-time attempt at cross-cultural collaboration between technologists (represented by signal processors) and humanists (represented by art historians and conservators). Park couldn't possibly say no to such an offer. We are about to observe one of their bull sessions.

Park looks down from the Terra Cotta soldier's eyes to Rick's.

RJ – OK – how do you want to get started?

PD – I sent you that list of questions. We could go through them.

Park Doing (left) and Rick Johnson with the Chinese Warrior in Rick's office.
(Photo: Jessica E. S. Edmister, ECE, Cornell University)

Rick turns away from Park toward his computer monitor.

RJ – Sure. Sure. OK. Let me get that email.

Park interrupts this gesture.

PD – First though, let me just step back and ask you a big picture question.

Rick answers, still looking sideways.

RJ – OK. What's the question?

PD – Why did you get into the application of signal processing to problems in art history? What is interesting about it to you?

Pause, Rick still looking at screen rather than at Park.

RJ – What is interesting about it to you?

Park is not flustered by this push back. Both Park and Rick exhibit nothing but comfort with the flow evidently familiar from many previous such episodes of wandering banter.

PD – Well… I have to think a little… for me it comes down to a scene in Gabriel Garcia Marquez's 100 Years of Solitude. Toward the end of the novel, the story is told of a priest and a peasant playing chess beneath a tree, but the game can never end because each of them is playing with a different set of rules. I'm forever intrigued by that kind of situation.

Rick laughs.

RJ – That's the question? It's exactly your type of question. When there's no blueprint for picking the problems – how do you pick the problems?! This is especially critical in trying to bring two areas together that are deemed by all to have little to no overlap. In all of my research subjects, there are two languages to learn, one for the area of the exploited technical expertise and another for the domain of the subject to which it is to be applied. It has always been that way for me. My first crossover between control systems and digital signal processing (DSP) began during my PhD studies in the 1970s when the researchers in these two subjects occupied two distinct worlds, with separate journals, separate conferences, etc… This is difficult to visualize now as the two fields are intertwined in many ways with a large number of top researchers active in both communities. Back then, I saw issues in adaptive filtering popular among the DSP crowd that were being addressed in slightly different versions in recursive system identification research in the control systems community. I was one of a small group – at first – of researchers exploiting and explaining this interconnection. That worked out well.

PD – So even within engineering – those were two different worlds.

Scientists face language and cultural barriers.

Park's tone conveys that he is trying to push Rick to acknowledge the disunity in science and engineering, the bricolage of what appears from the outside to be monolithic endeavor unified by shared understandings – that scientists themselves face language and cultural barriers.

RJ – Absolutely. At the time.

PD – So, you've already done this crossover-thing before.

RJ – Well, Yes and no. This time is much more extreme in the differences encountered. The institutions – universities and museums in the US and Europe – are different. Their languages are different. Research conventions are different. Different worldviews. A lot to overcome in getting us on the same page.

PD – Let's talk about worldviews. I would say that you have a 'signal processing worldview'. You see the world in signals. You see the elements of those signals and how they can be broken down, rearranged, reformulated even.

"Without data no theoretical problem can be posed that will have practical impact."

RJ – I agree. Continuing that train of thought, the art expert's reliance on examination of images viewed as signals suggests that signal processing can somehow assist art investigation. Actually, that line of thinking gave me the confidence to seek out ways to enter the museum world to observe the users of technology within the museum, i.e. the conservators, in the hope of seeing where my expertise could be applied. My PhD minor in art history helped me speak at a basic level in their language. I was keenly aware from the start that I could not resort to mathematical language or thinking in describing to them what types of problems I could tackle.

PD – In addition to the cross-cultural challenges, I've heard you remark on the unexpectedly large amount of time required to obtain access to sufficiently-high-resolution digitized images of art works.

RJ – The time commitment to gain access to scientific quality data has proven formidable. It remains a high barrier for new entrants into this growing field from outside the museum community. Fortunately, this is beginning to change.

You run into this in pure engineering too – in the middle phase of my career we were working on receivers for terrestrial broadcast high definition television before it existed on a wide scale. For competitive reasons, companies with field data were reluctant to share it with perceived competitors. And without data it is nearly impossible to pose an academically appealing theoretical problem that will have practical impact. The downside of working with real data is that it comes with all the nasties in it that complicate the specific problem's solution. But, having the data lets you raise questions that you do not know to raise without the data. That's the whole point. Real data helps you ask the right questions, and get useful answers.

PD – Why do you think painting data was made available to you?

RJ – Basically, I was very lucky with my first museum contact: the Van Gogh Museum in Amsterdam. In 2006, when I had my first meeting with their research and conservation management, they agreed right away to grant me the access to observe their conservators in action in my hunt for promising issues for collaboration. In exchange I offered to organize an international workshop the museum would host that would bring image processing experts to talk – without using mathematics – to art experts about computer-based tools for brushwork style classification.

Eighteen months later, we had identified the thread counting problem. The Van Gogh Museum ultimately provided us scans of x-radiographs of all of their paintings on canvas by van Gogh. They also approached other museums with requests for scans of their x-radiographs of van Gogh paintings on canvas. I learned that museums are used to sharing data with other museums, but not with outsiders like me. The museums have something to offer each other, i.e. access to images of artworks in each other's collections. At first, I had nothing to offer.

PD – What other attitudes/procedures did you have to adapt to?

RJ – I promised not to ask for the three things I knew they did not have to offer: money, space, or staff time.

PD – Tell me more about getting started on thread counting and weave matching?

RJ – In the beginning, I went to the Van Gogh Museum for a 10-day visit every 3 months or so. One day they said we were going to count threads. I asked to see a document beforehand that tells about this procedure and got blank looks. The concept of standard procedures, i.e. detailed algorithms for capturing measured data in a standard way, was itself not standard to them. When they showed me the x-radiograph images and taught me to count the threads visible under magnification, I recognized this task as a measurement of period that could be done on a scale unimaginable manually with the use of a Fourier transform. I would be helping to answer a question they wanted to answer.

PD – And weave matching came out of that?

"We started with the goal of automating thread counting."

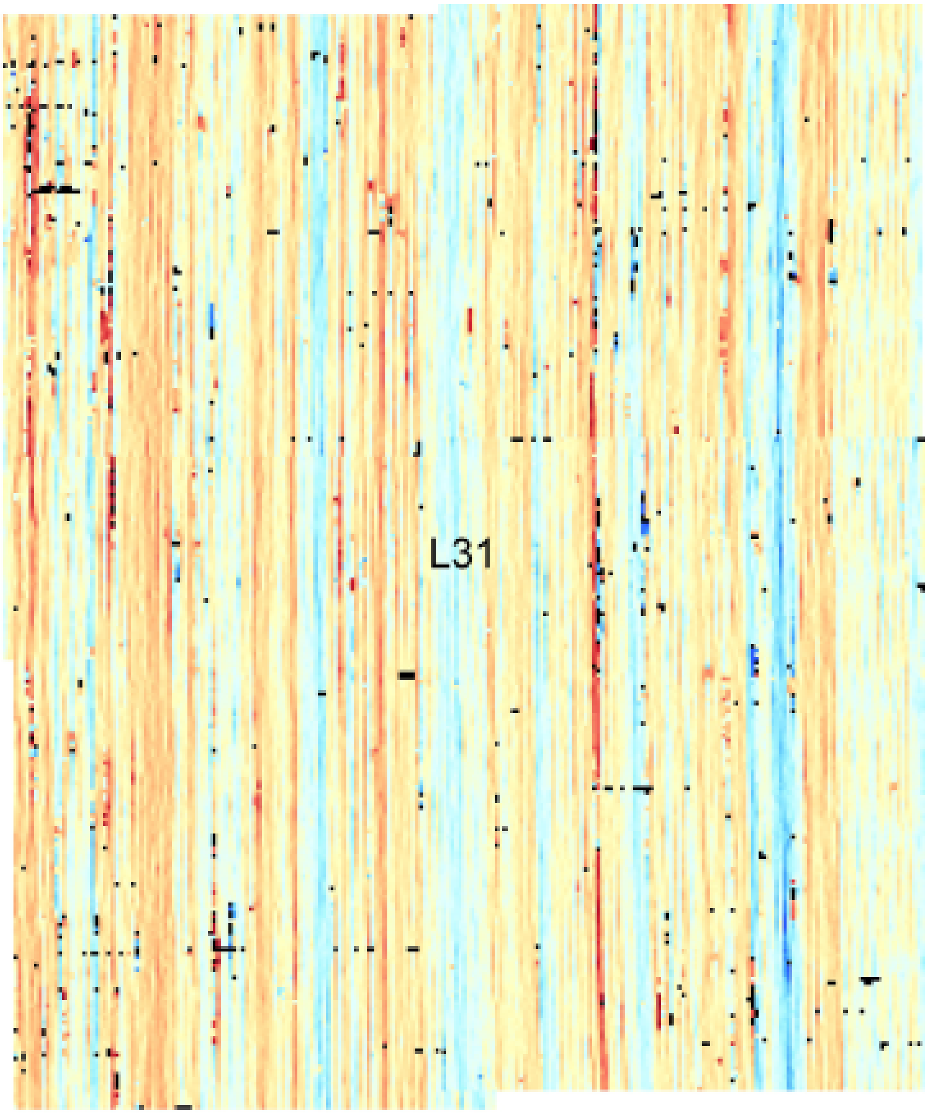
RJ – Well, yes. But nobody said, "We're going to invent a weave map." We started with the goal of automating thread counting. With that you can count not just the threads in a few sections of the painting, but in every section of the painting – and in every painting. So, after I got my first basic Fourier analysis program to work, I said to the Van Gogh Museum people that we could count all the paintings in their museum. They laughed. We started with x-radiographs for about 30 paintings. I realized later that they thought I was a funny guy – so American, wants to rush and do everything. I resolved to hew more to what I saw as the Dutch style of consensus decision-making where individuals are expected to suppress public display of their personal ambitions.

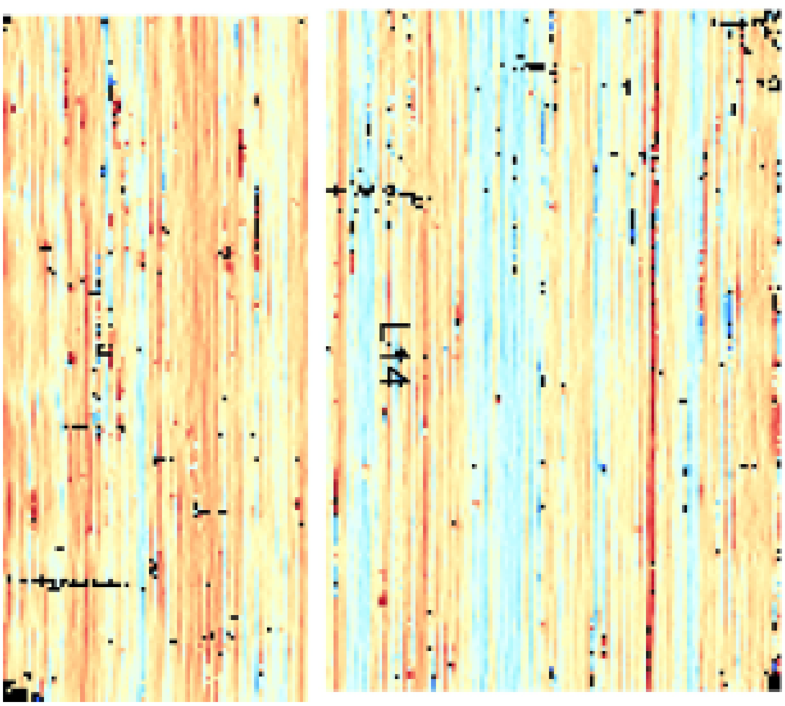
PD – So, a cultural difference beyond art and science?

RJ – Yes.

PD – Please continue.

The color-coded maps of local computations of weave density (threads/cm) reveal a matching stripe pattern in a pair of paintings by Vermeer.
(Image: Don H. Johnson, ECE, Rice University)
[For further details see W. Liedtke, C. R. Johnson, Jr., and D. H. Johnson, "Canvas Matches in Vermeer: A Case Study in the Computer Analysis of Fabric Supports," Metropolitan Museum Journal, vol. 47 (2012): 99-106.]





"We saw this pattern as a fingerprint for canvas from the same roll."

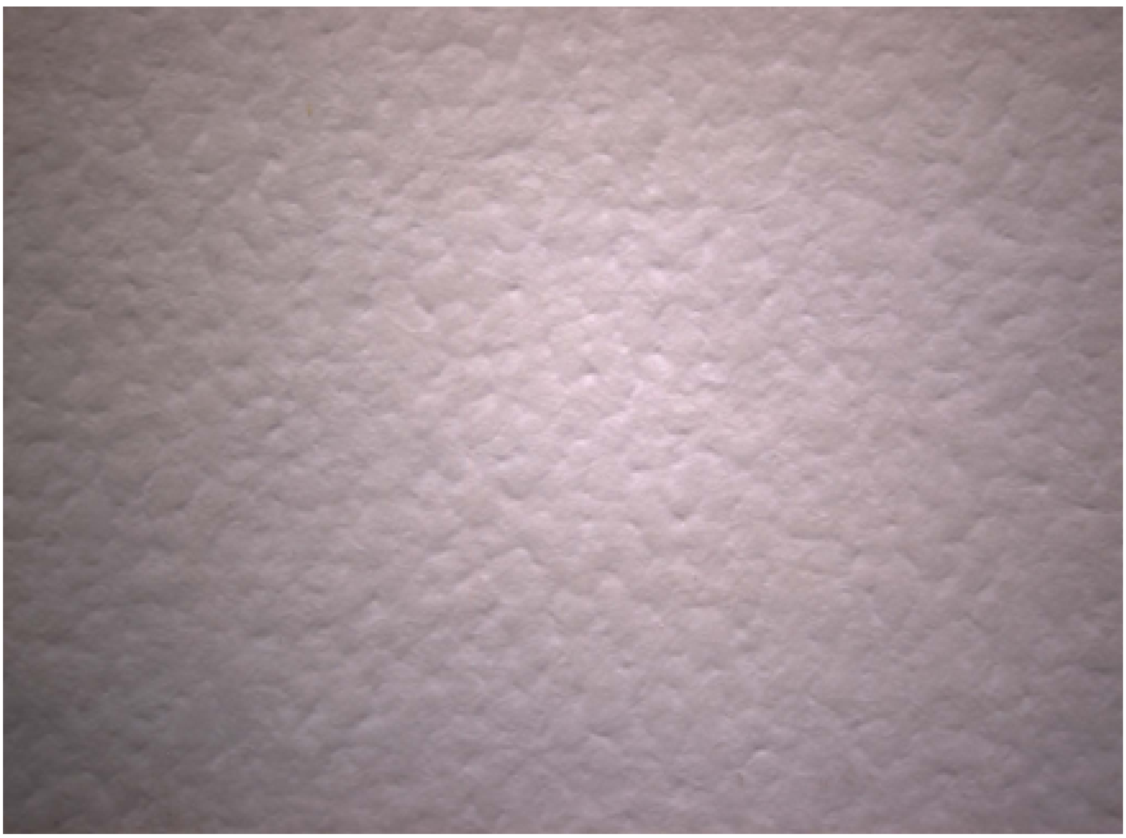
RJ – Weave maps emerged as we considered ways to present our count data. I knew a table with numbers was just about the worst thing we could do. That would go against what I was learning about how to communicate with art experts. It needed to be visual. I was working with people in the art community who are sublimely visually adept. If we could get it right, they would see it in a second. We color-coded the weave densities that were automatically counted and presented them as they covered the canvas. When we saw the vertical bands of color emerging we saw this pattern as a fingerprint for canvas from the same roll. You could see it clearly. I remember when I first presented an image of a match at a conference for conservators. I unveiled it and there was an audible gasp in the room. They got it right away. The weave map is now accepted as a new object with which to ask and answer questions about paintings on canvas.

> *"That encourages the pursuit of automating photographic paper classification."*

PD – After the canvas studies and weave matching, the photo paper analysis came about. How did that start?

RJ – For two months in 2010 I spent half of each week visiting the Museum of Modern Art looking for a task suited to the application of image processing. I met with their conservation scientists and paper conservators. I learned that photo paper is made for its texture and could be classified by observing the changes in reflectance as the paper sample was moved around under a bright light. Raking light is a standard illumination for revealing modest surface texture variations by their shadow pattern. We chose to collect raking light images of photographic paper at a microscopic scale. It took me over a year and a half to convince the paper experts that we needed images of some sets of paper for which the classification is known to allow us to build algorithms. While museums are most curious regarding the objects about which their knowledge is uncertain or simply lacking – to start we need images known to be a match. We need them in order to be able to design and test the accuracy of our candidate algorithms. Once we built such a dataset suitable for algorithm development and testing, the groups pursuing different textural strategies for classification were all able to show promise in using raking light images of historic photographic paper as a proxy for classification by metadata, i.e. manufacturer, surface finish, brand, pe-

Raking light images of 1.00 x 1.35 cm patches from two different 20th century black and white photographic papers displaying their distinct textures.
(Photo: P. Messier, Messier Reference Collection)
[For further details see C. R. Johnson, Jr. and others, "Pursuing Automated Classification of Historic Photographic Papers from Raking Light Images," Journal of the American Institute for Conservation, Vol. 53, No. 3 (2014): 159-170.]



riod of manufacture. Just last year we published a paper in the art conservation literature on this study that encourages the pursuit of automating photographic paper classification.

PD – And what about the chain line work?

RJ – That came about a little differently, since I had already done some useful work with the weave mapping and photographic paper, I didn't have to 'shadow' the art folks to scout out an opportunity. I was approached by a paper conservator at a digital humanities workshop. The task of identifying pieces of antique handmade laid paper made on the same mold from the impression in the paper left by the screen in the mold was proposed to me as being similar to thread counting and worth my consideration. At that time the standard approach to identifying moldmates was to match watermarks. But only about a third of, for example, Rembrandt's prints have a watermark or a fragment of one in the paper. But all laid paper exhibits chain lines. We decided to see if a simple description of the chain line pattern was enough to guide reduction of a library of paper samples to a manageable number of candidate matches.

Computational Art History or Digital Art History

Therefore, we skipped development of an automatic chain line marker, which is a difficult problem that will ultimately need to be solved in a real system, in order to get more quickly to testing the hypothesis of moldmate candidate discovery using just the chain line pattern. We observed that many paper samples had straight but non-parallel chain lines, which for some reason was a combination that had not been studied in the thin literature on automating moldmate identification for antique laid papers. From there a least squares fit did the trick. That was enough to establish the feasibility of using chain line spacing sequences to help find laid paper moldmate candidates.

A big issue was that the data to which we had gained access was collected for looking at watermarks. Thus, the images were taken of just a small part of the full print. Consequently, they typically contained too few chain lines. The chain line sequences were often just not long enough to be unique enough to sufficiently reduce the percentage of false matches. We needed full-print images of the prints. Luckily, last year we gained access to a trove of indexed full print images of etchings by Rembrandt. We should have much to report by the end of this year.

PD – Why do you call what you do a part of "Computational Art History" rather than "Digital Art History"?

RJ – I'm mimicking the currently fashionable use of "computational" as in computational biology or chemistry or fluid dynamics or linguistics. I want to imply that it's not just sorting and displaying images in large datasets, which is what is implied to me – perhaps incorrectly – by the label "digital art history". It's now much more than just managing digitized datasets. It extends to extracting information from the images, both forensic and contextual. It's modeling and simulation. Recently, I've begun to interpret most of the problems of current interest in applying signal processing to computational art history as some form of image feature mining.

PD – Feature mining?

RJ – Yeah – you see that article.

Rick points to his desktop where he has laid out an article with "image feature mining" in its title by researchers using facial recognition as their application.

RJ – The big deal with this paper is that the algorithm didn't know what features to look for ahead of time. It came up with the interesting features itself. In many of our problems the feature of interest is defined in the problem statement. The issue is locating/extracting/measuring this feature automatically.

PD – Do you think that is a direction for Computational Art History?

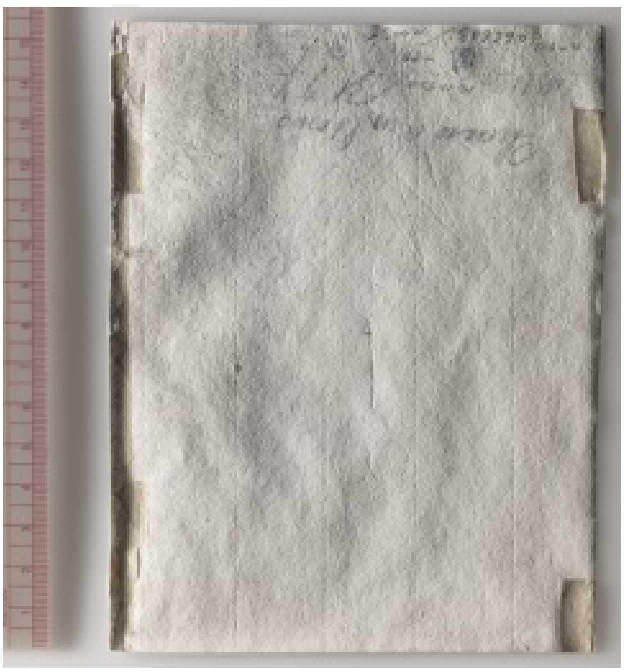
RJ – Yes. But, again, if the problems and questions being answered aren't coming from the art community – it's not going to be adopted.

PD – What is the biggest impediment to showing you can be of value?

RJ – Still I think it is typically the lack of quality data in sufficient quantities. But, this is definitely starting to change. In the beginning I had to use images they already had gathered. And very few were digitized in 2006. Conservation departments didn't have their own scanners. It took too many resources to digitize large numbers of images. You have to figure out what they can actually provide – can you get enough data to get started and convincingly demonstrate a potential positive impact by what you are developing? We managed to get enough and get something going. Data is everything. That's why one of my targets has been convincing museums to provide easy access to academic researchers of more and more images of art objects.

PD – Let's imagine that that the floodgates open up and the data issue fades – where do you see the future of computational art history going?

RJ – Rather than try to make long-term projections, I'll relate a recent relevant experi-



Vertical chain line impressions visible in raking light image of the back side of a Rembrandt etching "The Small Lion Hunt (with Two Lions)" on laid paper.
(Photo: David O. Brown/Herbert F. Johnson Museum of Art, Cornell University)
[For further details refer to C. R. Johnson, Jr. and others, "Hunting for Paper Moldmates Among Rembrandt's Prints," IEEE Signal Processing Magazine (Special Issue on Signal Processing for Art Investigation), (July 2015).]

"Newly institutionalized interactions are forming with art historians, curators, conservators, and engineers."

ence. As a guest editor for a forthcoming special issue of the IEEE Signal Processing Magazine on art investigation, I was trying to draft our editorial foreword about how what we are doing now in this nascent field relates to activity at the start of the 21st century when fewer signal processors were involved. I decided to divide the activities into image acquisition, manipulation, and feature mining. After consultation with my fellow guest editors, we decided that all of the articles in the special issue dealt with aspects of feature mining. Here we are using an inclusive definition of feature mining encompassing situations where the features are prescribed as well as instances where they are to be learned automatically – with a common primary objective being classification. This emphasis on feature mining contrasts sharply with the strong emphasis on image acquisition and manipulation around 2000. The current range of feature mining applications is quite broad, as evidenced by the topics addressed in the special issue, which include classifying ancient coins, facial recognition in Renaissance paintings, extracting and comparing visual stylistic features of paintings by a particular artist or school of artists, canvas thread counting, photographic and laid paper classification, and content based image indexing.

Imagine offering an art historian automatic labeling of content information in art works covering an artist's entire output – who knows what kinds of questions they would then ask? This is where I run out of my ability to predict the future. Uncertainty about the most fruitful future directions in such a young interdisciplinary field is a major reason for maintaining active cross-disciplinary collaborations in such projects. The domain is rich enough that I am convinced that some useful knowledge nuggets no one knows about now are waiting to be discovered. Thread counting and the subsequent weave density maps, thread angle maps, and roll layout capabilities form my current best example for this optimism.

PD – If the data is there…

RJ – Precisely! I think we are going to get to giant databases for images – I definitely think that is the direction. You'll load yours up to the cloud and within minutes you'll get a bunch of suggestions back about its mates. That's within reach. Within a decade, maybe. That should be a major target for our current collaborations. Again, my sense is not to try and predict – but get the data there and then we will see what happens – things I can't even guess now! Studying these mountains of data with feature mining tools seems like a very promising path to take.

PD – How do you see your contributions to this new field?

RJ – When I started I made a list of things that I wanted to accomplish. I wanted to convince more signal processors to look at these art investigation tasks. I wanted to convince art historians, curators, and conservators that the results from the signal processors will extend the scholarly reach of the art experts. I wanted to help establish an accessible archive of data and algorithms. I wanted to produce one textbook for both undergraduate engineers and graduate art and conservation students. And I wanted to give away software with a short course to conservation grad students. Basically, all of these targets were adopted as measures of my desire to accelerate the integration of signal processing and art history. All of these targets are now in hand or in sight.

I am very heartened – newly institutionalized interactions are forming with art historians, curators, conservators, and engineers all together at the start of interdisciplinary projects. The Netherlands Institute for Conservation, Art and Science and the Yale Lens Media Lab are recently inaugurated examples I have watched at close range as they took shape. I sit back sometimes and I think – it's really happening!

Sunshine is visible through the office's small window.

C. Richard Johnson, Jr. received a PhD in Electrical Engineering from Stanford University, along with the first PhD minor in Art History granted by Stanford, in 1977. Following 4 years on the faculty at Virginia Tech, he joined the Cornell University faculty in 1981, where he is the Geoffrey S. M. Hedrick Senior Professor of Engineering and a Stephen H. Weiss Presidential Fellow.

At the start of 2007, after 30 years of research on adaptive feedback systems theory and blind equalization in communication receivers, he accepted a 5-year appointment as an Adjunct Research Fellow of the Van Gogh Museum (Amsterdam, the Netherlands) to facilitate the interaction of art historians and conservation specialists with algorithm-building signal processors. In 2013, Professor Johnson was appointed a Scientific Researcher of the Rijksmuseum (Amsterdam, the Netherlands) and Computational Art History Advisor to the RKD - Netherlands Institute for Art History (the Hague, the Netherlands).

For a fuller description of Rick Johnson's research activities in computational art history, specifically using signal processing to match manufactured patterns in art supports, visit http://people.ece.cornell.edu/johnson/.

Park Doing earned B.S. and M. Eng. Degrees in Electrical Engineering and a Ph.D. in Science and Technology Studies, all from Cornell University. His book *Velvet Revolution at the Synchrotron: Physics, Biology and Change in Science* (MIT Press, 2009) analyzes interdisciplinary interactions between scientific fields, and between scientists and technicians. His subsequent research has centered on engineers as experts in dialogue with policymakers and the public. Most recently, he is focusing on applications of algorithmic processing to social issues and the humanities. He is currently a Lecturer in The Bovay Program in History and Ethics of Engineering at Cornell.