Everything Is in Motion

Checker Vision Sensor and Time Resolution

At the Vision show end of 2007, INSPECT hosted a panel discussion on the topic "All you ever wanted to know about Machine Vision software". One of the panelists, Cognex co-founder and Senior VP Bill Silver focused in his closing remark on motion and time resolution as the next most important development in machine vision software. INSPECT recently visited Bill at Cognex headquarters in Natick, Massachusetts and found out what this is all about and how it is connected to Cognex's vision sensor product Checker.

INSPECT: Bill, I heard you claim that the Checker is not only a vision sensor but a new technology unmatched in the industry. Isn't that a somewhat wild claim for the low cost end of the product range?

B. Silver: It's very easy to look at Checker as Cognex's low cost easy to use vision system. You know, we took In-Sight, which is more expensive and does a million things and wanted something that was easier and cheaper, so we stripped out all this stuff and now all Checker has is a couple of simple tools, and its really easy to set up. Well of course that's the truth but it's not the whole truth. And the whole truth is, I think, more interesting. I think it really does represent a tremendous break from the past and one that – if I have anything to say about it – is going to be much more common in machine vision moving forward.

Regarding the break from the past, I remember you saying during the panel discussion we had in Stuttgart that there are two important technology developments coming up: one is 3D, and the other one is time resolution. So over a year later now, would you still forecast the same?

B. Silver: I would. So let me talk some about motion and time resolution, partly because I did make that probably the centerpiece of my closing remarks. I think it was probably the most surprising thing that people heard that day because nobody else was talking about it. I'm the only one, and here's why. Let's go back 25 years to the emergence of industrial machine vision. So, in the old days, all image analysis and industrial machine vision was binary. Threshold an image, and you've got a binary image and you do things like blob analysis and morphology, and that's the way it was all done. Grey levels were kind of controversial. I remember I gave a paper in 1984 for a conference in Boston on why grey level processing is a good idea. And there were articles saying it's a waste of time, who needs all these grey levels because after all it's either an object or its background. And what my paper showed, that it makes a difference even if the object is essentially binary, like let's say a character. It's supposed to be binary. It's either ink or not ink, right? But here's the problem and this problem is fundamental to machine vision accuracy, and one of the reasons I'm interested in motion today. You've got a pixel grid and if you've got a perfectly binary object its edge cuts through the pixel grid in random ways. That's why you need the grey levels, because if you want to get down to the sub-pixel level you've got to understand how these edges cut through the pixel grid. And the pixel grid fundamentally limits what you can do. In the early days, it was very controversial that you could get any sub-pixel accuracy. In fact there are some people even today who are skeptic, "You can't get sub-pixel, that violates the sampling theorem. How can you get more resolution than your pixel size?" Well of course in 1983 people didn't believe it. There were articles published in the trade press saying this whole sub-pixel thing is nonsense. Just like you were talking about the 40th of a pixel, was it just a myth…

"I remember I gave a paper in 1984 for a conference in Boston on why grey level processing is a good idea."

"The grid is the limiting factor for things like normalized correlation."

> *"You don't have to squeeze that same orange any more because now you've got fundamentally new information that you didn't have from that one image."*

It's a long way from binary images to the claim of sub-pixel accuracy of a 40th of a pixel.

B. Silver: We eventually had to prove that you really could do it, and the key was going from binary to grayscale. That was the key to sub-pixel resolution. And the reason is you need more information. When you have a binary image you've kind of thrown away all this information and the nature of the pixel grid is what kind of screws it up. So we went to grey levels, and we achieved sub-pixel resolution. The problem though with grey levels once again when we developed things like Search is that the patterns we were looking for with the search were represented by pixel grids, alright. So if you look at a correlation template even if it's grayscale it's represented by a pixel grid. With pixel grids it's easy to move one pixel by one pixel or five pixels. It's hard to move a pixel grid by 2.5 pixels or 3.7 pixels, it's hard to rotate a grid. When you rotate a grid, when you change the magnification of a grid, when you scale it, errors are introduced necessarily, and again the grid is the limiting factor for things like normalized correlation. So when we went to PatMax, what I call geometric pattern matching, the idea was to get away from the grid representation.

By using a camera you have no chance but starting off with a pixel grid.

B. Silver: Of course you have to start off with a grid because that's what the camera does, But if you can represent your model not as a grid, but as essentially a geometric shape that is at least conceptually real numbers then all of a sudden you can rotate it, you can do all sorts of things without losing accuracy, and that's why with these methods we were able to get more accuracy than correlation. We can argue about whether it's a 40th of a pixel. I'd be happy to have that discussion as to why I believe that, but that was the next step.

Now at this point, sub-pixeling long since solved, what is the next step, what is the next "PatMax"?

B. Silver: You know what? We have gotten as much information out of an image as we are going to get. We are not going to get anymore information out of an image. And here is the reason why I say this: when we went from binary to grey level and developed Search grey level normalized correlation it took us about three months to make that work. It was 10,000 lines of code. Some very clever algorithms were developed, and that got us from whole pixel accuracy to quarter pixel or whatever it is. PatMax was probably 100,000 lines of code and it took four years.

We are squeezing information out of that image and squeezing and squeezing and squeezing information out of the image and you know what? It's gone man. To get more information out of the image than PatMax gets, would take a million lines of code and 10 years and who knows whether it would really even work. It's like you ever squeeze an orange to get orange juice? At some point you stop and throw the orange away because the effort to get another drop of juice out of that orange isn't worth it. And I think that's where we are with images today.

> *"To get more information out of the image than PatMax gets, would take a million lines of code and 10 years and who knows whether it would really even work."*

So, how do you get more information, what do you do?

B. Silver: To me what you got to do is get past the limitations of the pixel grid, and these accidental alignments that I keep talking about. For every image, all the image features are aligned relative to the pixel-grid in an accidental way. Where was that object when the trigger came in? So here's the thing about motion. Motion eliminates accidental alignment because you can see a feature as it moves through the pixel grid. It will appear in one image in a certain way relative to the pixel grid, but the next image it will be moved a little bit, it will be rotated a little bit. So you get to see it multiple times. You get to see that feature as it cuts this pixel grid in different ways and all of a sudden it's like opening up a window, there's more information. You don't have to squeeze that same orange any more because now you've got fundamentally new information that you didn't have from that one image. You can watch this feature move through the image. To give you an example, let's say that I'm running a standard edge detector. I've got an edge that's a little bit under resolved. It's maybe a pixel wide, or maybe a pixel and a half. As this thing moves through the pixel grid, as this thing rotates, the basic characters of that edge – its strength, its direction, whether it even exists or not – change quite a bit based upon the accident of where it hit that pixel grid. And if you look at these algorithms you can see what happens when things get a little bit under resolved. The fundamental measures vary quite a bit. It's no longer reliable whether this edge exists, you can no longer tell where exactly it is anymore from a single image. But, from several images if you can identify that feature in these several images as Checker tries to do, – all of a sudden you can recover a tremendous amount of information. Accurate information about where it is, accurate information about whether it exists or not, accurate information about its angle.

Checker was the starting point for the use of new methods to get beyond the pixel grid?

"Checker is right now the first and only product on the market that does motion."

B. Silver: The methods I've been working on since Checker are to use the motion of objects through the pixel grid to fundamentally get additional and more reliable and more accurate information about everything from edges to complex patterns. And to me, if you want to do better than PatMax, if you want to take the next step, you've got to get more information and motion is one of the best ways I know to get that information. Everything is in motion – in an industry, in the world, everything is. Even if something stops in front of the camera it had to move to get there. So if the information is available in the process all you've got to do is build the sensors and the algorithms that can take advantage of it. And if you want to do better than what we've got now from a single image – to me that's one of the best ways of doing it. 3D is another way of getting additional information but again we have squeezed that one image as much as it's going to be squeezed. You got to do something different. And the reason I'm excited about motion is – as I said – everything is in motion, I don't have to create it, it's already there. The quality of the information you get, the accuracy and reliability of information that you can get goes up tremendously. I also like it because nobody else is working on it. I like things that nobody else is doing. Checker is right now the first and only product on the market that does motion.

Why is it that this method is only used in the lowest cost and easiest product of Cognex, and not in any other product so far?

B. Silver: Historical accident.

Ok, …

B. Silver: A lot of times the answer to the why questions is, it was an accident. Checker was a giant accident. We really were trying to make a low cost easy to use vision system. That was the original goal of the project and so we said „Ok, we need a low cost imager". What we also need since it's a vision system, is an imager with a global shutter. The problem was to find an imager with a global shutter that's cheap. Well cheap means CMOS. At that time, you could get plenty of CCD sensors with global shutters but they were very expensive and we were having a hard time finding what we wanted. Well, one of the hardware guys came up to me and said "I've found this sensor. It's very inexpensive, it's got a global shutter, but I'm sure you're not going to like it. I only brought it up because it's the only one I could find with a global shutter and is inexpensive. But the reason you're not going to like it Bill, is because it's only 128 x 100 and that's not enough". So I looked at it and I said, „Yeah that's really not a whole lot of pixels, but – oh isn't this interesting, it runs at 500 frames per second. What can you do with that?" Well immediately it occurred to me that no one else was thinking about what you could do with it. That's what gets me interested, what is it that no one else is thinking about. Now maybe they're not thinking about it because it's a stupid idea, but I was intrigued by this and I said „Well, is there anything you can do with 128 x 100 pixels in machine vision?" And it occurred to me that we compete with photoelectric sensors for inspection applications where people use five or six sensors to inspect something. Photoelectric sensors got one pixel. Surely with 13,000 pixels I can do something. I've got 13,000 times as many pixels as that guy does. Now I don't have a quarter of a million pixels like In-Sight does, but surely there's room in the world for a 13,000 pixel sensor, when we know that one pixel sensors are useful. And I also said to myself „Why should I make yet another quarter of a million pixel sensor?" The world is covered with quarter of a million pixel sensors. So, this accident got us to thinking – what can you do with this oddball thing that's only 13,000 pixels, but runs at 500 frames a second? That's why it's in Checker and not anything else. It was this kind of accident.

"Everything is in motion – in an industry, in the world, everything is."

"The only way to get something radically new is to start over again."

So the major innovation needed the open mindset for starters.

B. Silver: At that point the In-Sight project has been going on for 10 years. Major innovations like this never happen to a mature product line. Technology is like cement. When cement is wet you can do anything – it's early, it's new, you can pour it into any shape you want but once it dries you're stuck with it and technology is like that. When you start off with the technology you can mold it any way you want but after a certain amount of time, and particularly when it starts finding some success in the marketplace, it becomes hard like cement. You can no longer do big things to it. You can chip away a little bit at it, but the only way to get something radically new is to start over again. And of course with Insight, you're not going to do that. Its 10 years old. It's wildly successful. It's a great product. So it had to be something new. It's the history of how the technology was developed that determined why it ended up in Checker and not somewhere else.

Bill, it was fun talking to you and I'll be looking forward to the next historical accident that you might come up with.

► **Contact**
Cognex Corporation, Natick, MA, USA
Tel.: +1 508 650 3000
Fax: +1 508 650 3333
info@cognex.com
www.cognex.com