Constellations Podcast

Episode 159 – Challenges with Geo-Registration, Advanced Remote Sensing and Utilizing AI

Speaker: Ken Campbell, President, Edgybees – 25 minutes

John Gilroy: Welcome to Constellations, the podcast from Kratos. My name is John Gilroy, and I'll be your moderator. Today, we'll talk about geo-registration and artificial intelligence with our guest, Ken Campbell. Ken is the President of Edgybees Incorporated, the U.S. Division of the Israeli-based Edgybees Limited, a global leader in geo-registration software. Ken, we're going to jump right in here. What the heck is geo-registration anyway?

Ken Campbell: John, first of all, thank you for having me come join you today. Geo-registration is, in short, the process of aligning or matching digital maps and images to the real-world locations on the Earth's surface. It's like putting a transparent overlay of a map on top of the satellite image or photos, so that the features in the map perfectly match the corresponding features in the image.

You got to think of it as fitting a puzzle together, the digital map and the image have different shapes, but geo-registration makes sure that all the features align on top of each other. This is crucial for various applications such as running this kind of data into geospatial information systems, for navigation, for additional remote sensing, and many other fields where precise location information is absolutely essential.

John Gilroy: Just for the benefit of our audience, so we define this topic here, are there different methods used for geo-registration?

Ken Campbell: There are. There's many, but let me just talk to two of them just for time's sake. The first choice, or the generally accepted practice, or well-known practice, is what we call using ground control points. Let me back up just a little bit here. First and foremost, the method is very dependent on the available data, the accuracy required, and the nature of the digital mapper image that you're trying to geospatially reference.

Most common and widely used method for this geo-registration then of course is ground control points, like I mentioned. What are ground control points? These are reference points on a digital mapper image that has already been processed, and each of those points has known geographic coordinates, and it's digital, so that as you cursor over it, you can see the latitude and longitudes associated with the pixels.

These reference points are usually, they're easily identifiable features on the Earth's surface, such as road intersections and bridges and buildings, and this sort of thing. You take those ground control points and you align them with the imagery. This has to be done typically in geospatial information system software like Esri's ArcGIS or ArcPro or something like that. There's actually a toolkit application for doing geo-registration in Esri.

Ken Campbell: You can use those kind of software applications. There's a number of others, ERDAS IMAGINE and ENVI and a few others. For the most part, that's one of the main accepted practices, if you will. Now, the only problem with that is that it typically takes time. It can take anywhere from minutes to hours to spatially reference an image. This is where we get into other techniques and applications, some of which are definitely being accelerated by the application of artificial intelligence processes. I would hope that that helps the audience understand a little bit about geospatial referencing, and some of the techniques that are used. We can get into how AI is being used to accelerate that process and improve accuracy. In fact, that's where Edgybees comes in.

John Gilroy: Yeah. I have a friend who's involved in agricultural, and he talks about precision agriculture and how important water is, and water distribution. I keep thinking, this is one way to optimize inputs and outputs for the agricultural world, but many, many other applications besides that. Give us just some general applications that this may-

Ken Campbell: For any remote sensing, you're going to need to do geo-registration because as you just described, you want to know what's the state of the water sources, where the head water's at, how's the floodplain looking? You want to be able to capture that not only from sensors, but then be able to geo-registrate all that data into digitized maps or images, and then push out to your farmers for civil engineering processes, and this sort of thing, so they can then use those along with their surveying capabilities and whatnot on the ground, to get a very accurate picture and know the location of where things are, so they can get there when they need to.

Ken Campbell: Cartography and mapping, it's essential, for emergency management and disaster response, it's essential. Also, as you can imagine, for military planning, it's essential.

John Gilroy: Ken, you and I both live in the Washington D.C. area, and I get my daily executive order from the White House, and at 9:00 a.m. this morning, breaking news, I have to mention artificial intelligence every 22 minutes. It's the top of the hour here, so I've got to ask you about artificial intelligence, because everyone's talking about it now. Everything's artificial intelligence.

I guess this works hand in hand with geo-registrations, because there's a certain limit to the amount of information a human can compute in a certain amount of time, and at least the artificial intelligence may not be as good as humans, but it can do it quicker. Where does AI fit in the discussion here?

Ken Campbell: Yeah. Let me just talk if I might about Edgybees' solution, because it's one that I know best of late. We employ artificial intelligence and machine learning, we develop algorithms that sit behind our software, that in essence take the raw image and a base reference layer. If you think about bringing the raw image into GIS software, our algorithms quickly compare all the feature elements in that raw image data to the reference layer.

They're doing this at thousands of calculations per second. With that particular technique, we can take for example, an image strip of, let's just say, multiple square kilometers and process that image strip in seconds, versus minutes or hours that other processes might take. Now, the other interesting thing is that, that raw image, it might come off of a sensor whose raw accuracy is maybe 10 to 15 meters from the actual target.

Ken Campbell: In some cases, that kind of accuracy might be good. In other cases, it's not helpful or useful at all. Our technology aligns that image with a base reference layer, as I mentioned. If that base reference layer is sitting at a one-meter level accuracy, in other words, each pixel is within one meter of the latitude and long that it's supposed to be, we can make that raw image take on the accuracy, the positional accuracy of that image.

Now, we can take sensor data that is maybe even further off, 80 meters off, and snap it right to that reference layer, do it in seconds, and provide that kind of positional accuracy, which then now has prepared the image data for other processes that you might want to apply, including additional artificial intelligence, machine learning processes, specifically for things like auto object detection, becomes very important for a variety of applications, whether you're in commercial industry or in the military. Counting cars in a port or identifying exquisite military items.

If that image does not have good positional accuracy, you can imagine, it's garbage in, garbage out. It's not very helpful.

John Gilroy: Yeah. That's a topic I want to bring up here is this whole idea of tracking capability, and then what, move up to the next level, warnings, alarms, automatically saying something, so that we could actually take this and go from listening, understanding, to acting, because we'll have this information available, something changed that we're waiting for. That's another application.

Ken Campbell: Absolutely. This gets into the whole discussion about not only auto object detection, but change detection. Being able to identify, for example, from a military perspective, troop movements or massing of equipment, and this sort of thing, becomes a little less labor-intensive, when you have this kind of technology, that's helping to take the cognitive load off the analysts, and give them more time to actually consider, now, what is it that I'm looking at in that change? Is that what's really happening? Is something being spoofed?

The bottom line is, when you're able to speed up that whole process, you're now giving the analysts back time to do that level of analysis, to get to a better answer regarding what is going on, on the ground.

John Gilroy: Yeah. 1440 minutes in a day, analyst is working eight hours. They can spend six hours of that filtering information, or they can spend six hours actually analyzing and using their evaluation capabilities. That's great.

John Gilroy: Earlier this morning, I spent a half hour on the phone with a director at the MIT Media Lab up there in Boston, and over MIT, if you look what the articles they published, they published an article a couple years ago about a need to devise best practices for synthesizing classified and open source intelligence, and for sharing this information with allies, partners and the public. Synthesizing or fusing global imagery, I don't know how you do it, seems overwhelming. Are there industry best practices yet for how to share this kind of data and how to collect it in real time?

Ken Campbell: Yeah. Oh boy. John, that’s the question.

John Gilroy: An overwhelming question, I know.

Ken Campbell: Yeah. Because you've got all kinds of legal and policy and business practice issues you got to sort through before you ever even get into the processing dynamics and sense-making, to make sure that you now have information from the data that's worth sharing. Let me try and ease into that just a little bit, try and make sense of this answer I'm going to give you.

First and foremost is the issue of what I call truth versus spoof. That is where minimally processing and exploitation efforts, they've got to be able to assess the veracity and relevance of the data and information, whether it's derived from open sources or sensitive sources. It's got to happen fast, since the sense-making piece can't start until one has a certain level of confidence in the information. This is where I think commercial imagery really comes into its own as a source of high-confidence information, principally because you can trace the pedigree or its origination.

Ken Campbell: Additionally, since it is commercial and collected through known means and this sort of thing, not derived from sensitive sources, there's rarely any issue or limitation being placed on its shareability, particularly when needed to support humanitarian relief, disaster response efforts, or for needs associated with national security. Meanwhile, in the context, the question regarding industry best practices for sharing commercial imagery, it depends on the imagery producers and the downstream consumers.

There's no one cookie cutter protocol for how sharing is done. That said, probably the best organized efforts at scale are those that are involving humanitarian assistance, disaster relief efforts. For example, the geospatial intelligence community. It's got organizations like the U.S. Geospatial Intelligence Foundation or Information Foundation, which is actively involved in disaster response efforts and promoting best practices and this sort of thing.

The National Geospatial Intelligence Agency is a key agency that helps with these things. There's another organization, Humanitarian Open Street Map Team, which helps with this. There's actually an International Charter Space and Major Disasters Organization that's made up for the most part of overseas companies involved in space infrastructure and Earth imaging that volunteer their services and time to help when these things occur.

By and large, it's pretty much an effort of mutual cooperation, and again, very dependent upon the providers and the consumers. Then, of course, those that are in the policy lens. I hope that helps. It's convoluted.

John Gilroy: Yeah. No, but there's so much data out there. There are whole companies that work on privacy enhancing technologies. It's called PET, it's got its own acronym, and how much data can be released and what can't be released. That's mostly with text-based data.

Ken Campbell: Right.

John Gilroy: Identifiable information. What about geographic information? How would you know, at what level ... It's harder to tag something, I get, than a specific fact, like 5027, rather, some kind of an image. That is a whole new world we're going to have to live in the next 20 years with these advances in technology, like your company.

Ken Campbell: If I might add, this is why you see so many up-and-coming space imaging companies. It's getting cheaper and cheaper to put those birds up. That said, to maintain accuracy at a high level, say below five meters, is very expensive. A lot of these companies don't have the funds to put $50 million worth of wheels and weights on their birds.

This is where solutions like Edgybees' SKY software come into play. It is very easy. It’s very easy to slip into the workflow and it plays nice with the software, and it can do that processing and provide a level of accuracy that is only limited by the level of accuracy of the base reference layer that is being used. Again, I use that as just an example.

John Gilroy: Now accuracy, you brought up earlier, we talked about meters, five meters, one meter and pixels and everything else. I think that someone, an outsider like me, I would think that of course the satellite's going to know right down to the number of buttons on my shirt, and it can see what kind of shoelaces I have. I don't think that's always the case, and I think people don't understand that the precision can be ... What's the number? It could be off 150, 200 meters, can't it?

Ken Campbell: Yeah. The one thing you got into is there's positional accuracy versus the resolution issue. You can certainly zoom in and see the writing on your glasses, but the positional accuracy of where you're actually located may come through in the raw data as tens of meters off. Again, getting back to positional accuracy, the things that make that difficult, why the raw data sometimes can be, in your question, I think you asked, why can it be off as much as 200 meters?

We've got atmospheric distortions. Again, you've got, depending on the camera, frame camera, line camera, it doesn't really matter, you're still running these problems. You've got atmospheric distortion from the earth's atmosphere. You've got orbital errors just from the satellite, because the satellite's moving, right? Okay. Then, there's the sensor calibration.

All these things factor in, and they all have to work together to get that good picture. Even then, you're still going to be a little off. Then there’s the topography of the Earth, and the data transmission and there's just a whole bunch of things that you have to factor in.

John Gilroy: Now, wait a minute, Ken. That's not how it works in the movies. In the movies, you go right down and you can see the watch on Ken's arm and tell what time it is. Come on now, you got to watch more movies, Ken.

Ken Campbell: Yeah, I have it all coming in right into my special spy reading glasses.

John Gilroy: Yeah, it does make it look easy. The reality behind this is it's really complex.

Ken Campbell: What's cool though is that, the technology today, with the workflows and the architecture and everything else, actually move things really, really fast. Architecting the solution is one thing, but then dealing with the electromagnetic spectrum to move all that data is another. That can be an impediment throughput too. Yeah, it's complicated.

John Gilroy: Yeah. Let's talk about architecture and enterprise architecture. Get down to a practical nuts and bolts, kind of an engineering type question. High-res images take up a lot of bandwidth, we know that, especially when pulled down from the sky. Is this a problem, or what can you do to overcome this type of a challenge of bandwidth?

Ken Campbell: Do you have stock in Amazon Web Services?

John Gilroy: I was at their tradeshow, but that's the only connection I've got to AWS.

Ken Campbell: Yeah. That's another cool thing about today. Cloud computing and parallel processing really come in their own. The ability to rapidly use these computational processes to crunch all that data and everything are just orders of magnitude better than they ever have been. I remember when I was going to university in my undergrad, using the old data cards to program the computer for a simple question. It's crazy, right?

John Gilroy: Every night, you get the answers in the morning. I remember.

Ken Campbell: Now, I talk to my son and he talks in foreign language, C++, GeoJSON, I'm doing this, I'm doing that, spitting this stuff out really fast with a laptop. Compute processing and the computer ecosystems have just really accelerated the ability to deal with these massive amounts of information.

In addition, with artificial intelligence, machine learning, continuous improvements are being had through the development of algorithms designed to crunch and do the sense-making of all this data, and then the inference engines in which the algorithms work. If you think about algorithms as bullets in a handgun, and the handgun is like the inference engine, everything just works really smoothly, and then you're spitting out information, right? It's probably not the best metaphor, but it's just one that I was thinking about.

John Gilroy: The Constellations Podcast was launched back in 2017. It was a small step for man, but a giant leap for podcasting. Today, thousands of people from all over the world listen to Constellations, and thanks to you, we've grown into more than just a podcast. Sign up for the Constellations newsletter to receive articles on current industry issues, podcast summaries, and contributed blog posts at constellationspodcast.com.

John Gilroy: Let's focus more on this power and AI and ML and the compute power that's available now. Does this actually, does this lead toward real-time processing? What's required to make this work?

Ken Campbell: It's incredible. First of all, you have to have the compute power. This is where the world has moved to graphic processing units, GPUs. Nvidia is sitting on the top of that particular technology. Having GPUs, having the power of parallel processing and the ability to store, to tag, to recall information, this sort of thing, is all a part of that.

When it comes down to the deck plates, what are the things that can actually be done? I mentioned one which is auto object detection. As that raw data is coming down from the bird, you stream it through the inference engine and algorithms, and it's automatically identifying with high confidence if the models have been trained correctly, objects of interest.

In some cases, and the ones that I'm most familiar with, these could be objects of concern, and then that metadata is stripped off, and it goes to analysts, or it goes to weapon systems, or it goes to any one of a variety of downstream clients that need to see that data, that information, and need to see it now for the purpose of making some course of action decision.

That's a big one. The other thing is land use, land cover, and making sense of what's actually on the ground. Semantic segmentation, where the algorithm can perform processing of each pixel and classify them into categories. This is very helpful, particularly when you're talking about things like agriculture and water use, identifying crops and assessing their health, all those kind of things.

The last one I'll just touch on is high resolution. We talk about accuracy, but we also want to be able to see all the fine detail. With the advances being made in compute ecosystems and the algorithms, you can get what's sometimes referred to as super resolution, where the actual image coming off the sensor might be at a half meter, but through additional processing, you could take it down to 15 centimeters.

John Gilroy: Wow, that's incredible.

Ken Campbell: That's pretty cool. Yeah.

John Gilroy: Tomorrow, I'm doing a couple interviews with some IT folks, and inevitably, we talk about data. Data in rest, data in transit, data in process, and data this, and data that, and data security. What do you do with all this data? Do these CSPs, do they provide enough storage for you? Or, can you just securely, data at rest, can you securely store it in the Cloud?

Ken Campbell: Yeah. This gets into the whole concept of what we call knowledge management. Certainly, there are, in Cloud services and whatnot, the ability to bin lots and lots of information and data. The real key though is being able to recall that data and information when you need it and have it at your fingertips.

It's like if you look, if I didn't have my image distorted, you could see my old library right there. I don't know how many times I go look for a book and it'll take me 15 minutes, because I don't know where the hell I put it. In comparison, you not only want to be able to store that data, you have to be able to get at it faster. It's just not useful.

Metadata tagging is really a big one. Data organization is super important. We talk about Cloud-based storage to store that stuff. With the metadata tagging, we get into data indexing and cataloging and then replication. Then, we get into the whole concept of compression and data formats, and tiered storage approaches where you want to have the stuff that you need now, always up at the top, and this sort of thing. All that gets architected by the wizards of code writing and data science.

John Gilroy: Yeah. We're both in the D.C. area, I mentioned that earlier. The big dog for standards is NIST. We all know NIST up in Gaithersburg, we've been to their campus there, we know what's going on. I would imagine, I don't know, but are there standards that companies follow in this geo-referencing model? Do map companies have some kind of a standard so that data can be shared besides longitude and latitude?

Ken Campbell: Yeah, there've been standards developed, and the standards typically define the common data formats, data models and protocols that allow for that sharing and ingestion into other systems, so that it can be used to do that sense-making and generate information for decision making, this sort of thing.

John Gilroy: Ken, I went to your LinkedIn profile and I saw quite an interesting background. If you want to be entertained, follow Ken Campbell on LinkedIn. I have to ask you this, you had a pretty diverse career here. What personally got you interested in this type of technology?

Ken Campbell: Yeah. Back in 2004 at the beginning of the global war on terrorism, I was asked to support a special defense project to evaluate a series of new technologies, designed specifically to identify human behavior in a spatial context. One of the companies in particular, and it's on my LinkedIn profile, it's called Spatial Data Analytics or SPADAC. They had developed a geospatial modeling solution that could identify location preferences for different types of human activities, such as meeting sites, residential preference areas, weapons storage sites, this sort of thing, those things that we commonly refer to as pattern-of-life activities.

They developed a technology that could help model pattern-of-life activities. You can imagine what we might use that for, given that it was a quantitative model, crunching lots and lots of information, and then spitting out a spatial preference for certain types of activities we want to learn more about. I actually helped the team develop the methodology for really using that technology well.

Shortly after we got in sync, or my analyst got in sync with the geospatial analysts, we started to get some really, really exquisite result. Shortly after that, the CEO of that company, young man named Mark Dumas, who I'm still very, very close to, invited me to come work for the company, and I never looked back. I've had years now in the GIS space infrastructure intelligence world.

John Gilroy: Quite a storied career. It's kind of interesting when we go through it. Ken, I think what you've done today is given our listeners a real good understanding of what geo-referencing is all about, so they know where it applies for their specific problem. I'd like to thank our guest, Ken Campbell, President of Edgybees Incorporated. Thanks, Ken.

Ken Campbell: All right, John, thank you.