

## Microsoft's Hong-Jiang Zhang: The Process of Product Innovation

*"If you're working on actual products you can't say that 90 percent is good enough and just move to something else."*

Hong-Jiang Zhang is managing director of the Microsoft Research Asia Advanced Technology Center (ATC) in Beijing, China. He received his Ph.D. from the Technical University of Denmark and his bachelor of science from Zhengzhou University in China, both in electrical engineering. Zhang will be the general co-chair of the 10th ACM Multimedia conference.

**UBIQUITY:** Tell us about your job at Microsoft Research.

**ZHANG:** I'm Managing Director of Microsoft Research Asia Advanced Technology Center (ATC) here in Beijing. Previously I was the Assistant Managing Director, and managed all the research activities around multimedia, Web search, data mining, etc. I'm still involved with research but with my new job I'm more focused on putting research into actual products, which is what the ATC was created for: to basically transfer research innovations into product.

**UBIQUITY:** Give us some examples.

**ZHANG:** Well, for instance, one thing we did involves video content summarization and editing, where the problem is to look at the video content and decide automatically which parts are most important and interesting, so that you can keep those parts and delete the less interesting parts. But of course all those decisions are made by a computer instead of by people looking at the content and deciding what's interesting and what is not

**UBIQUITY:** How do you decide what's interesting?

**ZHANG:** That's a very, very good question. What we do is define "interesting" in a way that it can actually be translated into numeric terms, so that, for example, you can look at the motion in the video and at the activities portrayed there and you can listen to the soundtrack and when there is, let's say, an explosion, you can presume there must be something big happening. Or if you on the video someone is laughing, that's presumably something interesting. If you see constant action and motion, that's probably something interesting, and you can identify many clues like this. If they happen all the time, that means the content is most interesting.

**UBIQUITY:** How many people are in your group?

**ZHANG:** Now I'm involved with a group of close to 150 people.

**ZHANG:** I have 150 people working for me.

**UBIQUITY:** You've got a big responsibility.

**ZHANG:** Well, yes. That's why this year has been kind of hectic.

**UBIQUITY:** And what are your biggest challenges?

**ZHANG:** I think number one is simply that it's never been easy to transfer research innovation into product, so that's really been a challenge. Researchers tend not to think about actual products, and when their solutions are 90 percent accurate and complete, they tend to think that's good enough, and they consider that the problem's essentially solved. But if you're working on actual products you can't say that 90 percent is good enough and just move to something else. So improving a research result which works *most* of the time to a product that has to work *all* of the time is the challenge, and it requires a mind shift. So that's the number one challenge. Number two, I think, is more my personal challenge: how do I best make the transition from managing very creative long-term fundamental research -- where there is no schedule most of the time -- move to a situation where there are very specific and very hard deadlines. When it comes to product development then you have to define and follow schedules that are so detailed that you often have to change them every week or even every day. And then of course my third challenge is learning a very new environment, because when you're in pure research you can use whatever programming language you want as long as you can solve the problem you can make it work, but when you're working on an engineering system you have to follow a rigorous engineering process to make sure that the bugs are minimized and that all of the security checks get done. So it's not just a mindset shift but also a skill-set shift, because the required management skills are very different now.

**UBIQUITY:** Has your group had to expand quickly?

**ZHANG:** Yes, very quickly. We had a huge growth in the last 12 months, going from a group of less than 30 to one that's now over 150. And it's very important to make sure new people fit into the culture and get their skill set right. And just think about interviewing all those people: that's definitely been a challenge for me.

**UBIQUITY:** What are the backgrounds of those people?

**ZHANG:** Mostly computer science related. Most of them have a master's degree in computer science or electrical engineering or some related field, and some of them have Ph.D.s in those fields. Their specialties range from software engineering to networking to mobility to multimedia to machine learning. It's a very broad range of computer science.

**UBIQUITY:** No marketing people in the group?

**ZHANG:** No, we're not doing marketing, but we do have a couple of people doing market research just to try to predict technology and market trends, but we don't do marketing as such because we're not doing end-product release. Our work goes into Windows, Office, and other products, such as MSN and Media Center, all those things. The whole range of Microsoft software. We do have a few projects focused on local users here in Asia, but mainly we are focused on global users,

**UBIQUITY:** What is the particular challenge of producing things for global markets?

**ZHANG:** Basically we are working with a large team and we are only one part of that team, so we have to communicate across oceans, across time zones, across language barriers.

**UBIQUITY:** And how often do you do that?

**ZHANG:** Well, we're doing it weekly, daily, by e-mails, teleconferences, video conferences, and by visiting each other. I travel to Redmond and people from Redmond travel to here.

**UBIQUITY:** What kinds of miscommunication problems have you had?

**ZHANG:** Any miscommunication usually comes from some cultural difference. Fortunately we typically have more than 15 people who are actually U.S.-trained and have U.S. experience, so that helps a lot. But in a Chinese culture there is a tendency not to say no to people, not to be very aggressive, not to disagree, and even if you disagree you do not express your disagreement in strong terms. So those are these kinds of cultural differences, and then some type of English language skill may sometimes become a barrier, but we always manage to overcome it.

**UBIQUITY:** How so?

**ZHANG:** We definitely have English training, and on top of that we have what you could call cultural training. Sometimes we have professionals come in do the training, but more importantly it's our own experienced employees who offer brownbag discussion and talk about how to handle various situations. And of course are people are very smart, they are top notch, and come from the best schools in China, and some of them from the best schools in the U.S., so they learn very fast. Between March and September we brought in 50 new-hires, most of them fresh out of school. And you can look at the demographics and look at the schools they graduated from, and you'll find that they are all from the best 15 schools in China, and most of them are from the top five computer science schools in China. And because of the intelligence and creativity of these employees the energy level at ATC is just phenomenal. Even though we are somewhere between product development and basic research, we are really like a startup, people are excited about coming to work, so they can be part of the newest and best research. On the other hand, also being able to turn those research results into products that will be used by many, many people -- this is something that is quite exciting. The energy level here is just phenomenal.

**UBIQUITY:** Who and where are the people with whom you communicate most frequently?

**ZHANG:** The ones we communicate with the most are back in the States in Redmond -- people back at Microsoft headquarters who are working on product, who are busy getting ready to ship the next Windows software or the next Microsoft Windows Media software. And we're also working with people who are going to work on MSN. Those are the product team, and on other side are our researchers. So we stand in-between research and product development.

**UBIQUITY:** Say something about the competitive nature of multimedia. In other words, how much do you think about what Microsoft competitors are doing other in that area?

**ZHANG:** Yes, that's a very good question. Multimedia has become absolutely essential to the PC business, and is a driving force for PC upgrades, which include both hardware and software upgrades. So naturally Microsoft is expending a lot of effort on improving the state-of-the-art of multimedia, and Microsoft Research Asia is

well known for its leading position and breakthrough research results in the field of multimedia. When you go to any multimedia conference, academic conference, and especially the ACM multimedia conference, you see a predominant presence of papers from Microsoft research groups. Last year at the ACM multimedia conference the second week of November in New York there were 55 presentations, and several of them come from Microsoft Research Asia.

**UBIQUITY:** What is it that people outside your specialty would find it hardest to understand about what you're doing?

**ZHANG:** Well, we're lucky, because it's actually easier to understand multimedia than to understand networking, because multimedia is something you can feel. After all, you turn a dumb PC into a multimedia PC so you watch movies and you edit videos, and you share photos on the PC, and edit movies and search on multimedia content. And now you do video streaming over wireless, over to our mobile phone. So people really do not have much difficulty understanding the multimedia, and they find it fascinating, because they can see, they can feel, they can listen. And in recent years, when bandwidth has become so cheap, what we do with multimedia is becoming more popular every day.

**UBIQUITY:** But isn't it popular like magicians are popular? People are amazed at what you do -- but don't some of them wonder how in the world you do it?

**ZHANG:** Yes. That's right, that's right. I found that when I was giving research seminars in universities about doing automatic video editing (and by the way I'm a guest professor in quite a few universities in China and also ones in Singapore). When I show people the computer program in which I automatically edit 10 hours of raw video footage down to a 10-minute highlight synchronized with them with the music, people don't believe me. They say, "How could you do that?" I tease them by saying, "If anyone here in the audience is a film major, one day you may lose your job because of this automatic editing technique." Of course they're very much amazed and ask me, "How did you do this, how did you do that?" I tell them, "Well, in the beginning I didn't believe I could do that either, because I always think that video editing or movie editing is an artistic process, but when we're talking about home videos what's need is a quick, fast way of doing basic editing."

**UBIQUITY:** So what's the trick to this magic?

**ZHANG:** What I tell them is that if you do motion analysis and object analysis you'll get many clues about what's important. For example, as you track people who are in the video you can assume that if the camera is always focused on a single it that person must be important. And if you see the camera zooming into an object, that object must be important. And if you see a very fast panning, whatever is in the pan is not important. Then, when they stop panning, that's also important. And so on. If you see a very slow panning, then you know the cameraman is basically doing a panoramic view. So by using all those clues you can do a decent job of editing, and when I explain it like then people then are able to say, "Oh, yes, that's understandable, that's doable."

**UBIQUITY:** Then is the secret of your explanation to walk them through the process itself?

**ZHANG:** Yes. When you're dealing with something like video editing, which is so practical and so artistic, you have to use some out-of-the-box thinking, and be creative enough to fit this very vertical framework to something that is very realistic, very practical. And it leaves people feeling amazed, the way they look when I use a new Microsoft smart phone to make real-time hook into some portals and watch the news. People are just amazed at that, and the quality is not that bad at all. You know, you really don't need much higher bandwidth, you've got a decent quality and you have error correction and scalable media compression to handle fading in the wireless channel.

**UBIQUITY:** Is this kind of research you're doing on automatic editing being done in many other places?

**ZHANG:** Not many. We are among the earliest.

**UBIQUITY:** Have you dealt with the evaluation issue and found a way of deciding whether or not it's good editing?

**ZHANG:** That's a very good question. Everything depends on who is the user you are targeting. If you're targeting movie editing professionals, then we are totally outclassed, but we're talking about people who have a camcorder and take pictures of their newborn babies or their travels, but they're not professional videographers, and a lot of times their hand is shaking or the angle is not right or the lighting is not good enough or they shoot something they shouldn't shoot. Or they forget to turn the camera off and shoot a lot of garbage there. These are the kinds of people we can help with automatic editing. By editing a home video clip you will be happy, and you will not feel embarrassed to share with others. My own experience is probably typical. I bought my first video camera when my older child was born, and the only person who can watch it without editing is my mom, because she lives 12 miles away and because she wants to see absolutely everything my grandson does -- everything! But when she tried to show the video to my brother, he of course was not interested in watching a child bawling and eating etc. -- he's not interested in watching raw video footage. It's not really edited. So to solve that kind of problem, you don't have to do 100 percent perfect editing job, right? And that tends to be an advantage for us engineers. We engineers always try to measure our successes and failures, but in this situation we now we have a soft measure, which we could call the "Wow" measure. Since we place a very high value on what end users think, we measure our success by having users look at our edited video and then look at what we edited out. And they usually say, "Oh, wow, that's amazing." One fellow said that only if you know how boring and long the original video is can you appreciate how amazing this system is.

**UBIQUITY:** Could you imagine this being used for news photography or for sports photography?

**ZHANG:** News would be hard. Sports are easier because sport is a sequence of events with a specific grammar, by which I mean the specific rules of the game.

**UBIQUITY:** What you've done is really quite fascinating. Do you consider this the main product that you're interested in?

**ZHANG:** No, it's just one example that I'm allowed to talk about. There are many other products we're working on now that will not be released for two years or even

three years. Automatic video editing is something we released a year or so ago, when ATC was still in incubation form.

**UBIQUITY:** During the course of your career, what has surprised you most?

**ZHANG:** Yes, that's a very good question. One thing that surprised me was that China in such a very short time became the largest mobile phone market in the world, which shows that when you have the right technology, right applications, and good matches with people's needs and needs of a particular culture, then you have something that will be phenomenally successful.

**UBIQUITY:** What do you have in mind when you say culture?

**ZHANG:** For example, in the U.S., you usually don't want to be bothered incessantly with phone calls, you don't want to have a phone follow you all the time and everywhere you go. But in China it's totally the opposite, and so that's one thing that surprised me. Another surprise is that I never expected that when the telecom bubble burst in early 2000 and 2001 it would come back so quickly, but it has. You can almost see another bubble forming.

**UBIQUITY:** So what do you see exciting on the horizon?

**ZHANG:** I continue thinking that wireless is and mobility are very exciting, and I definitely think mobile multimedia is very exciting. Wireless multimedia is going to be really interesting and exciting. I think broadband do amazing things for the people of Asia.

[END]

Hong-Jiang Zhang's address is [hjzhang@microsoft.com](mailto:hjzhang@microsoft.com)

*Source: Volume 6, Issue 7, (March 1 - March 8, 2005)*

<http://www.acm.org/ubiquity>