Mask-bot: A robot with a human face

November 7 2011

Dr. Takaaki Kuratate and his robot communication interface Mask-Bot Credit: Uli Benz / Technical University of Munich

Robotics researchers in Munich, Germany, have joined forces with Japanese scientists to develop an ingenious technical solution that gives robots a human face. By using a projector to beam the 3D image of a face onto the back of a plastic mask, and a computer to control voice and facial expressions, the researchers have succeeded in creating Mask-bot, a startlingly human-like plastic head.

Mask-bot can already reproduce simple dialog. When Dr. Takaaki Kuratate says "rainbow", for example, Mask-bot flutters its eyelids and responds with an astoundingly elaborate sentence on the subject: "When the sunlight strikes raindrops in the air, they act like a prism and form a rainbow". And when it talks, Mask-bot also moves its head a little and raises its eyebrows to create a knowledgeable impression.

What at first looks deceptively like a real talking person is actually the prototype of a new [robot](#) face that a team at the Institute for Cognitive Systems (ICS) at TU München has developed in collaboration with a group in Japan. "Mask-bot will influence the way in which we humans communicate with robots in the future," predicts Prof. Gordon Cheng, head of the ICS team. The researchers developed several innovations in the course of creating Mask-bot.

The projection of any number of realistic 3D [faces](#) is one of these. Although other groups have also developed three-dimensional heads, these display a more cartoon-like style. Mask-bot, however, can display realistic three-dimensional heads on a transparent plastic mask, and can change the face on-demand. A projector positioned behind the mask accurately beams a [human face](#) onto the back of the mask, creating very realistic features that can be seen from various angles, including the side.

Many comparable systems project faces onto the front of a mask – following the same concept as cinema projection. "Walt Disney was a pioneer in this field back in the 1960s," explains Kuratate. "He made the installations in his Haunted Mansion by projecting the faces of grimacing actors onto busts." Whereas Walt Disney projected images from the front, the makers of Mask-bot use on-board rear projection to ensure a seamless face-to-face interaction.

This means that there is only a twelve centimeter gap between the high-compression, x0.25 fish-eye lens with a macro adapter and the face mask. The CoTeSys team therefore had to ensure that an entire face could actually be beamed onto the mask at this short distance. Mask-bot is also bright enough to function in daylight thanks to a particularly strong and small projector and a coating of luminous paint sprayed on the inside of the plastic mask. "You don't have to keep Mask-bot behind closed curtains," laughs Kuratate.

This part of the new system could soon be deployed in video conferences. "Usually, participants are shown on screen. With Mask-bot, however, you can create a realistic replica of a person that actually sits and speaks with you at the conference table. You can use a generic mask for male and female, or you can provide a custom-made mask for each person," explains Takaaki Kuratate.

In order to be used as a robot face, Mask-bot must be able to function without requiring a video image of the person speaking. A new program already enables the system to convert a normal two-dimensional photograph into a correctly proportioned projection for a three-dimensional mask. Further algorithms provide the facial expressions and voice.

To replicate facial expressions, Takaaki Kuratate developed a talking head animation engine – a system in which a computer filters an extensive series of face motion data from people collected by a motion capture system and selects the facial expressions that best match a specific sound, called a phoneme, when it is being spoken. The computer extracts a set of facial coordinates from each of these expressions, which it can then assign to any new face, thus bringing it to life. Emotion synthesis software delivers the visible emotional nuances that indicate, for example, when someone is happy, sad or angry.

Mask-bot can realistically reproduce content typed via a keyboard – in English, Japanese and soon German. A powerful text-to-speech system converts text to audio signals, producing a female or male voice, which can then set to quiet or loud, happy or sad, all at the touch of a button.

But Mask-bot is not yet able to understand much of the spoken word. It can currently only listen and make appropriate responses as part of a fixed programming sequence. However, this is not something that is restricted to Mask-bot alone. A huge amount of research worldwide is

still being channeled into the development of programs that enable robots to listen and understand humans in real time and make appropriate responses. "Our established international collaboration is therefore more important than ever," explains Kuratate.

The CoTeSys group has chosen a different route to researchers who use dozens of small motors to manipulate individual facial parts and reproduce facial expressions. Mask-bot will be able to combine expression and voice much faster than slower mechanical faces. The Munich researchers are already working on the next generation. Mask-bot 2 will see the mask, [projector](#) and computer control system all contained inside a mobile robot. Whereas the first prototype cost just under EUR 3,000, costs for the successor model should come in at around EUR 400. Besides video conferencing, "These systems could soon be used as companions for older people who spend a lot of time on their own," continues Kuratate.

Provided by Technische Universitaet Muenchen