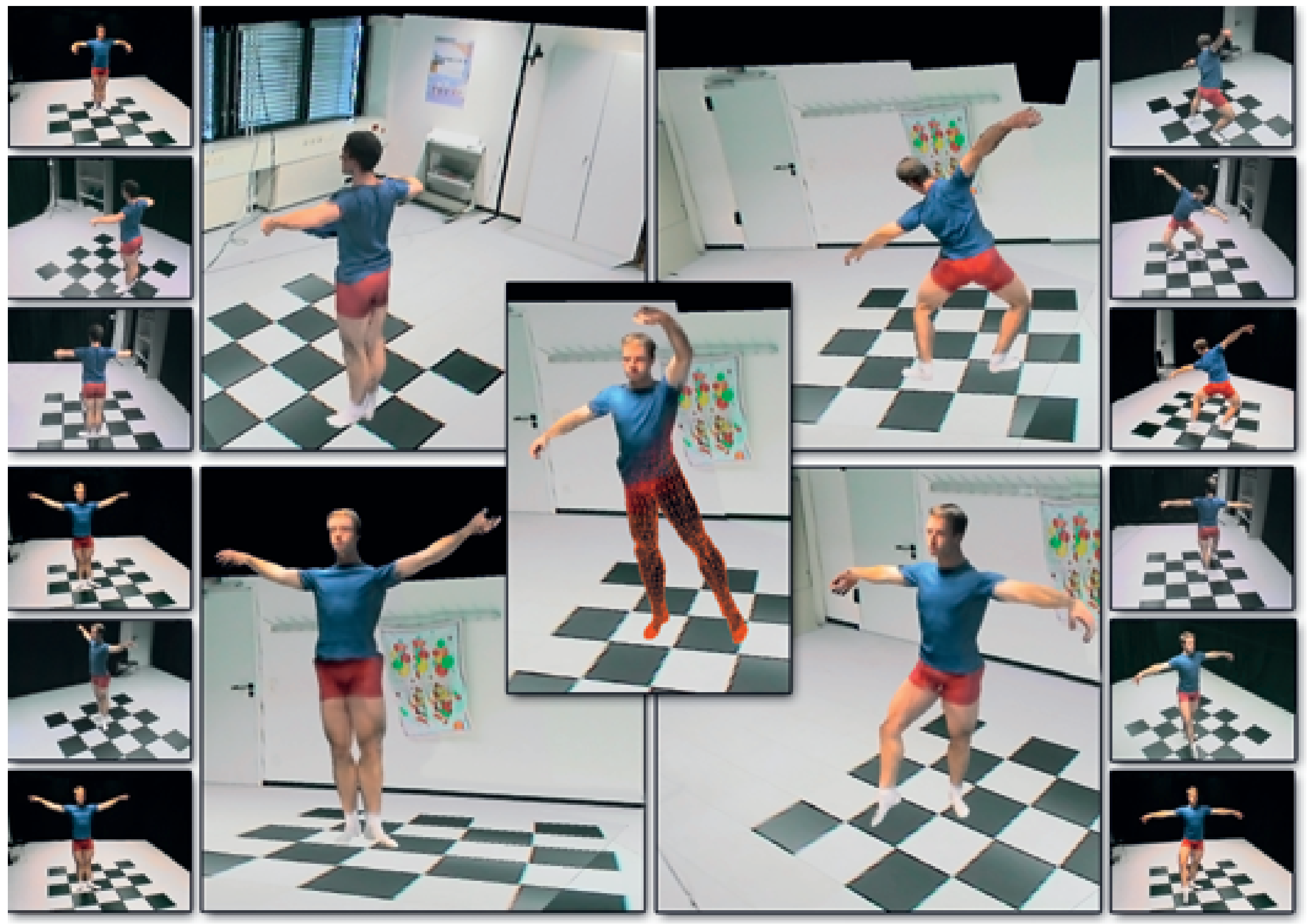
A Perfect Semblance of Reality

Computer graphics specialists dream of computer-generated images with a highly realistic appearance. But our sensitive eyes can discern even slight differences between reality and imitation. **Marcus Magnor** *and his independent research group at the* **Max Planck Institute for Computer Science** *in Saarbrücken are looking for new ways to create more realistic computer graphics. Their specific focus is the reconstruction of three-dimensional scenes from two-dimensional photographs or videos.*



The movements of the dancer are recorded by eight synchronized video cameras, then aligned with a digital human model – temporarily transforming the dancer into a Spiderman of sorts.

Photos: MPI for Computer Science

I would describe myself as a very visual person," says Marcus Magnor when our discussion turns to human vision. He and his team are confronted daily with the problem that we have a tremendously precise ability to identify faces and movement patterns, and to differentiate between an enormous variety of objects and materials. Shapes play a key role in this, but so does texture: we are immediately able to identify skin, hair, leaves and fabric based on their surface appearance. We register even the tiniest details that subtly affect the reflection of light.

Thanks to our keen sense of vision, computer graphics specialists have their work cut out for them – especially if they want their synthetic images to be convincingly realistic. And that is just what Marcus Magnor is aiming for: his basic research involves developing new ideas for image production. The subjects he studies are as varied as the world around us. They range from virtual models of candle flames, trees, sunsets, animated articles of clothing and dancing people to celestial bodies.

"Graphics – Optics – Vision" is the name of a six-member independent research group at the Max Planck Institute for Computer Science in Saarbrücken. Marcus Magnor headed this group until the end of last year. He was so successful, in fact, that he is now continuing his scientific career as a professor for computer graphics at TU Braunschweig. He has left Saarbrücken, but that is the whole idea of the Max Planck Society's support program for junior scientists. At the same time, however, the young professor will remain closely affiliated with the Max Planck Institute for Computer Science. Some of his team members are still there completing their theses or doctoral dissertations.

"Our specialty is creating three-dimensional videos of moving scenes," explains the 33-year-old researcher. By this, he does not mean films that can be viewed with 3-D glasses to create the illusion of three-dimensionality. Magnor's team approaches the third dimension from a different angle. Their "Free Viewpoint Videos" allow the audience to choose any vantage point. Viewers can interactively turn running scenes back and forth on the screen, and even watch them from above, turning the scene into a game board with living pieces.

Such videos require a disproportionately greater amount of work than conventional films that comprise merely a straight sequence of images. For the 3-D videos, the computer must provide the right pictures from every point of view. This requires a full three-dimensional model for each frame, and this model must capture the original movement exactly. The computer then uses the model to synthesize the scene from the desired perspective – a process known in the graphics world as rendering.

Eight Points of View for Three Dimensions

As a basic researcher, Marcus Magnor is interested primarily in the question of – to put it simply – how the 3-D model gets into the computer. Modern-day animators could, of course, reconstruct their virtual models manually. After all, such a labor- and cost-intensive approach has been used to create the odd character for certain animated films and computer games. However, it is of limited use for figures that are supposed to be realistic in their appearance and movements.

Thus, Magnor and his team of researchers took a different approach. Their models are based on photographs and video images of real objects and people. The team in Saarbrücken is developing a largely automated series of steps by which the essential image data can be harvested from the raw footage. Following sophisticated image analysis, the data is then used to generate a virtual model of the scene that very closely resembles the original. Finally, the computer can use the model to render images in real time from any point of view. "Real time" means that the computer carries out the commands with no time lag or jumps in the flow of movement.

For this to work, all of the necessary calculations must be completed sufficiently quickly. This requires intelligent optimization of the synthesis processes, which the team is constantly working to refine. They have set the bar very high for themselves: their 3-D videos are not designed to run on elite supercomputers, but on commercial graphics cards on standard PCs.

A look at a three-dimensional video that the team created from a solo dance scene exemplifies their methods. The footage was shot at the Max Planck institute in an improvised studio that was originally a computer room for students and that is about the size of a large classroom. Black curtains were used to hide the walls and windows. Eight video cameras were distributed around the room and focused on center stage, where the dancer performed on a floor made up of black and white squares. This grid pattern

With "Free Viewpoint Video," the team in Saarbrücken can portray moving humans or objects from any point of view.

aided image analysis, helping the computer identify the silhouette of the dancer and the relative position of the floor.

The eight cameras were set for synchronous filming. "The synchronization was the most expensive aspect of the studio," says Magnor. But it was important, because the program always needed eight simultaneous images to compare during the analysis stage. If the computer had miscued and combined images shot at different moments, the three-dimensional reconstruction would have failed.

The first step in analyzing the images involved the computer identifying the silhouettes of the dancer. These were used to frame the figure in a sort of cage comprising eight contour lines taken from the respective camera angles. To complete the 3-D model, the computer had to fill in the remaining spaces with the appropriate convex surfaces. For this part, the team reached into its bag of tricks and found a prefabricated human model – a digital model of a human body with a simple skeleton, the 17 major joints of the limbs and torso, and even the ability to flex its muscles in a convincingly real way.

The Construction of a Virtual Clone

The computer expanded this homunculus into the contours of the real dancer. After each stage of growth, it synthesized images from the model and compared them to the original video. If there were discrepancies, the computer adjusted the digital human model in the succeeding step, until it ultimately filled in the contours of the real dancer perfectly. Such computing processes that gradually conform a model to an original image of a real object are the specialty of the researchers in Saarbrücken. "I think it is a very good idea to approach analysis as a synthesis of images," says Magnor. "The best part about the resulting graphic is that we quickly arrive at a realistic picture. The great challenge, of course, is how to make the model match the images." The researcher calls this step in the process the inverse problem, and that is what the research group is now focused on solving. "It calls for a bit of real mathematical genius," says Magnor.

Of course, animators and computer game developers have long implemented similar processes. Modern animated characters such as Shrek use motion capture filming, in which actors bring the characters to life. But current technology reproduces movements only at a small number of markers on the bodies of the actors. That is enough for animated characters from the realms of fantasy.

For images that are meant to depict realistic human characters, however, the team in Saarbrücken is at the forefront with their process that precisely models the contours of the performer from every angle. The result is a nearly perfect "digital clone": it is bound to the original video images, and copies the movements of the living dancer down to the subtlest detail, convincing even our critical human eyes.

"We are very good at identifying movements as authentic or obviously recreated," explains Magnor. "This problem was already recognized by Disney animators some sixty or seventy years ago." Their clever solution was to draw characters with comically exaggerated, rubber-like movements. "But that is just what we have to avoid when attempting to render images realistically," stresses Magnor: "Between caricature and realism is a range to which our senses are very critically attuned."

Clothes Make the Man

The virtual dancer now twirls with convincingly realistic flair across the screen of the 3-D video in Saarbrücken, but only as a gray figurine. To complete the perfect virtual clone, skin, hair and clothing are needed. The interplay of light and shadow is also missing. Computer graphics specialists refer to these aspects as "texture". In order to create texture when rendering, the computer first covers the model in a tight-fitting bodystocking made up of triangles. It then fills these triangles with surface characteristics such as "red fabric" or "blonde hair." These are taken from the original video images of the dancer and transplanted onto the gray surface of the model – an elegant solution to another major problem presented by synthetic computer images. Texture is enormously complex, especially when it comes to representing humans. There is a reason that the prominent computer-game heroine of ten years ago, Lara Croft, looked like a defiant sister of Barbie. It is also no coincidence that the producers of the first fully computer-animated film Toy Story chose plastic figures to star in their 1995 feature.

"Real skin is problematic, for instance, because it is semi-transparent," says Magnor. "Light goes in at one place and comes back out at another." Since computer graphics can't copy the complex realities of molecules and light quanta, convincing alternatives must be developed.

Bad Hair Days in Virtual Reality

"Some very interesting approaches are now being used for skin, and much progress has been made in this area over just five years ago," says Marcus Magnor. "Hair and fur, on the other hand, can be endlessly troublesome." There is much too much going on in a head of hair for the strands to be modeled individually. "But I can't simply take hair as a texture and overlay the head with a photo of the hairdo," explains the researcher. Hair is too thick for that to work. Our eyes can discern individual strands even in the fullest head of locks. "Moreover, hairs have a directional structure and thus reflect light more strongly in some directions than in others," he adds.

If green isn't in this season, one click will change the color of the dress to suit any taste – at least on the screen.

On top of this, because it is semi-transparent, hair also tends to direct light much like fiber optics do. In the

The Max Planck researchers in Saarbrücken use a technique from the field of computer tomography to portray foliage realistically.

mass of a head of hair, all of these aspects result in very complex properties in terms of light and shadow. "There is a broad field of research to cover," Marcus Magnor goes on, "before that can be realistically rendered in real time." The methods used in Saarbrücken work well with fibers that are finer than human hair, such as fabrics. Unlike hair, fabric textures can be lifted directly from photographic images. The team has even devoted an entire research project to the subject. "You could call it chameleon clothing," says Magnor, grinning. One click changes the color, or even the pattern, of the article worn by the animated model on screen. Each new outfit is impressively realistic, with even the critical shadows cast by the swaying folds being rendered perfectly.

Heavenly Bodies in a New Light

Where there is shadow, there must be light, and its properties are the leitmotiv of Marcus Magnor's research career. He studied physics at the University of Würzburg and the University of New Mexico, finishing with a master's thesis on experimental optics. Thereafter, he turned his attention to electronic communications in Erlangen, where he focused on the coding of light fields for computer graphics. "That's a data structure in which an object is photographed from several different angles," he explains. Thus he found his research calling, which he continued to pursue even as a postdoc in computer graphics at Stanford University. In 2002, the position offered by the Max Planck Institute for Computer Science enticed him from California back to Germany.

Light plays an important role when modeling astronomical bodies. Marcus Magnor is particularly interested in animating these for educational purposes. "The digital revolution in video projection is offering planetariums completely new possibilities," says Magnor, "such as 3-D flights around astronomical objects, or even right through them." But the projectors need suitable footage in order to provide these 3-D effects, and production of such film material costs disproportionately more in time and effort than traditional flat images of the celestial realm.

Marcus Magnor demonstrates how the production process can be made more efficient using the impressive subjects known as bipolar planetary nebulae. Some stars in the throes of death surround themselves with these gas clouds. "They are generated by the rotation of the dying star, or by a second star orbiting the burnout victim." These rotations can give the gas cloud, which is released by the star as a kind of cosmic last breath, a symmetrical form: the forces generated by the motion work like a tight ring around the star and impel the gas to move out into space in opposing directions. This gives planetary nebulae a delicate hourglass shape.

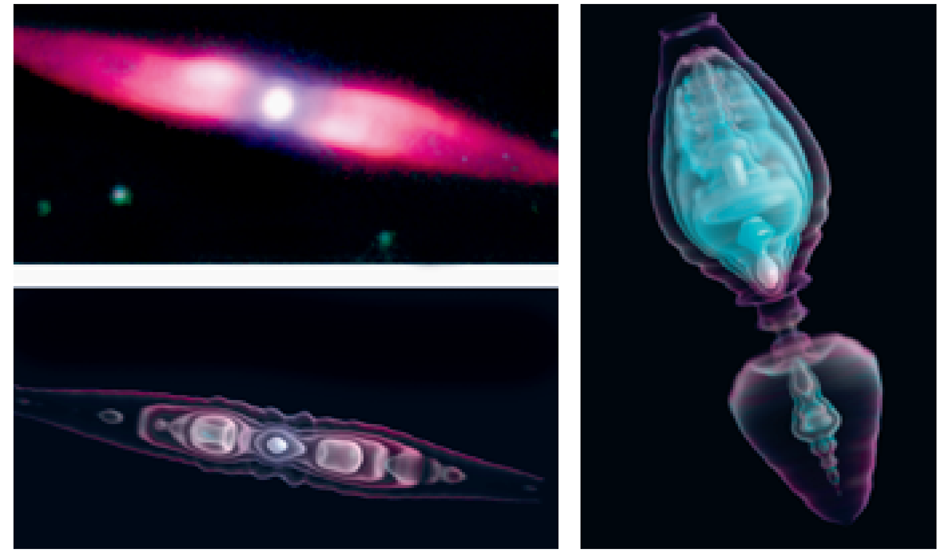
The images of such objects, however, are always projected onto the celestial sphere in only two dimensions. Unlike with the dancer, Magnor and his team can't move around astronomical nebulae and view them from different directions. However, this gap was filled by astronomers' physical models. They have identified some 1,500 planetary nebulae to date, each of which presents a different aspect to the Earth.

Even Computer Games Require Research

"The planetary nebulae emit a very beautiful glow in various colors," explains Marcus Magnor enthusiastically. The variety of colors comes primarily from hydrogen, helium, sulfur and oxygen that glow from within the cloud as if it were a gigantic fluorescent lamp. Magnor and his team analyzed astronomers' images, for example those of nebulae M1-92, and used them to develop three-dimensional models. As with the dance video, their program synthesized images from the models, compared them with the original photo, and aligned them in a step-by-step process. This resulted in striking three-dimensional nebulae models that could be interactively turned and viewed from all angles on the screen.

Marcus Magnor's research group is developing new ideas for 3-D worlds in a broad array of fields, from planetariums and astronomy research to the clothing industry. And, in addition to the film industry, computer games are increasing in importance for the scientists. "We are being driven by this mass market, even though it does not sound very scientific," says the young professor. "But because of its scientific potential, this market has done much to advance computer graphics."

Computer games also led Magnor to look at some objects that are much closer – namely trees. "Many computer games contain outdoor scenes," he explains, "and realistic representation of trees is a problem that remains to be solved." At the heart of it, trees present another "hairy" problem, this time clad in green: there are too many leaves on a tree for them to be modeled individually, and yet the individual leaves are too large to simply ignore. A simple cut-and-paste texture would give the tree a strange, lifeless appearance. And on a real tree, there is a different interplay of light and shadow caused by the leaves whenever the point of view is shifted.



Planetary nebulae, viewed through a telescope at the top left, can be modeled in three dimensions with a computer (bottom left and right), orbited virtually and viewed from any direction.

A Virtual Tree Bears No Leaves

In order to synthesize more realistic trees, the team integrated billboards into their three-dimensional model – small surfaces that fill out the entire volume of the tree. A texture of leaves, taken from photographs of real trees, was then projected onto these. This voluminous distribution of tiny projection screens, however, is only the first step. For the realistic depth effect, the computer scientists needed a second process to intelligently distribute the textures across the billboards, taking into consideration, for instance, that outward-lying leaves shine more brightly than those closer to the trunk.

The team cracked this tough nut with the help of researchers at the INRIA institute in Nice, using a process taken from the field of computer tomography.

The imaging programs used in computer tomography must be able to automatically identify the depth within the body from which the image is taken. The program for the distribution of leaves works in a similar manner: it automatically identifies the billboards that are farther from the outer foliage, and applies darker textures to them.

The result on the screen is quite realistic, despite the fact that the virtual tree has no actual leaves. "When viewing an image with a resolution of 1,000 x 1,000 pixels, which is the setting on a typical computer screen, the eye cannot perceive individual leaves," explains Marcus Magnor.

Of course, researchers dream of models that include even more detail and that approximate reality even more closely: "We are already considering construction of multi-scale models with which you can zoom deeper and deeper into the tree, to view first an individual leaf, then a cell, and then for example the mitochondria," reveals Magnor with a meaningful smile. "That would be something!"

Roland Wengenmayr