

# Engineer and vision scientist nets Emmy for tool to predict perceived video quality

October 5 2015

---



An example of JPEG masking that shows the contrast masking effect in action. Bovik applied insights from visual psychology and neuroscience to design tools that allow broadcasters and streaming video sites to compress and distribute video so its distortion is minimized. In the above image, the distortion is the same everywhere, but is less visible in busy regions (in the grass and flowers or the girl's hair) and more visible in smooth regions (on the girl's face). This does not occur because the grass hides the distortion, but instead, is a by-product of a visual process that renders the distortion invisible or nearly so. SSIM is designed to predict these 'content-dependent' effects. Credit: The University of Texas at Austin

Sometimes it's not so much what you see as what you don't see that matters.

Alan Bovik led a team of researchers that, in the early 2000s, invented a tool to predict how the average person will perceive the quality of an image or a [video](#). The tool allows broadcasters and streaming video sites to compress and distribute video with minimal distortion.

So, every time you don't see jagged images, blurring, glitches or the dreaded spinning wheel of death on TV or the Internet, that's likely thanks in part to the Structural SIMilarity index that Bovik and his research team created.

For his efforts, Bovik and three of his colleagues will each receive television's highest honor, an individual Primetime Emmy Award for Outstanding Achievement in Engineering Development from the Television Academy (formerly the Academy of Television Arts and Sciences).

Bovik—the Ernest J. Cockrell Endowed Chair in Engineering at The University of Texas at Austin—will receive his Emmy statuette at the Primetime Emmy Engineering Awards ceremony at Loew's Hollywood Hotel on Oct. 28, 2015. Two of Bovik's former PhD students, Zhou Wang (now at the University of Waterloo, Canada) and Hamid Sheikh (now at Samsung, Dallas) will also receive individual Emmy statuettes for this work, along with Eero Simoncelli, their collaborator at NYU.

"An Engineering Emmy Award is bestowed upon an individual, company or organization for developments in engineering that are either so extensive an improvement on existing methods, or so innovative in nature, that they materially affect the transmission, recording or reception of television," the Television Academy wrote.

Bovik was initially surprised about the Emmy—in fact, he found out about the award while reading his email in the morning, on his way to brush his teeth. Eventually, his surprise gave way to appreciation.

"I'm very happy and grateful that people are recognizing our work," Bovik said. "I so appreciate NSF support of our research and our students, who are so much of the inspiration and perspiration behind research success. If you have good ideas, this kind of support can change your life."

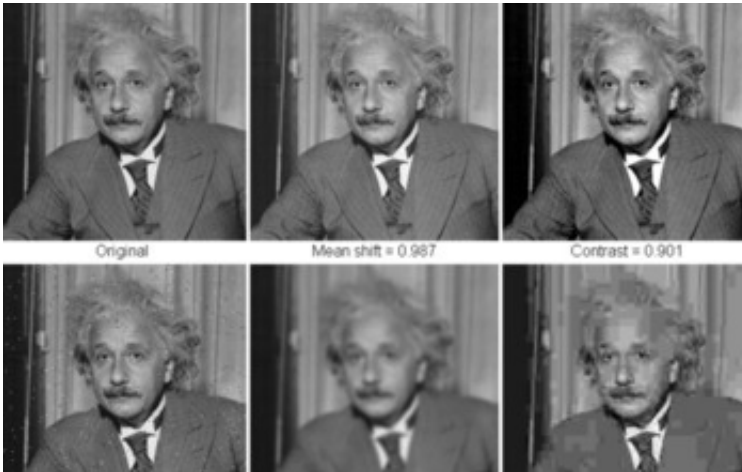
## **Compression issues**

Due to limits in bandwidth and the ever-expanding number of streaming videos, broadcasters need to compress their video before sending it over the airwaves or the Internet. This compression can cause a range of problems, from blurring to blackouts, which make watching a broadcast unpleasant or distracting.

"What you would like to be able to do is have a reliable way of determine acceptable levels of perceived quality—as a human viewer would report it," Bovik said. "This is particularly important to cable and satellite video content carriers in regards to how much they compress the television signal."

"It's not like a certain compression level will give the same quality for every video. You actually have to do something content dependent, and being able to predict the perceived quality of a compressed video while accounting for content makes a big difference."

The Structural SIMilarity (SSIM) index is a method for measuring the similarity between a compressed image or video and the uncompressed original, in terms of human perception. Today, SSIM is part of the technological toolkit used by most major broadcasters and cable and satellite companies around the world, including AT&T, Comcast, NBC, FOX and PBS.



In the following image, all distorted images have roughly the same mean squared error values with respect to the original image, but very different quality. SSIM gives a much better indication of image quality. Credit: Zhou Wang, Alan C. Bovik, Hamid R. Sheikh and Eero P. Simoncelli

Technology manufacturers such as Cisco, Motorola (Arris), Intel and Texas Instruments rely on SSIM to ensure the broadcast, networking and TV equipment they produce maintains the best possible [video quality](#).

When Bovik and his team started working on the issue of predicting how people would perceive video quality, they faced some difficult challenges. At the time, state-of-the-art prediction algorithms either did not account for how humans perceive distorted images, or were too computationally intensive to be practical. Drawing from neuroscience models of low-level vision, the SSIM model disruptively solved both problems by providing the world's best solution to the video quality prediction problem at a low computational cost.

"The first breakthrough was a very simple model developed by an advanced graduate student, Zhou Wang and myself called Universal Quality Index. We later developed the final SSIM model with another

grad student (Hamid Sheikh) and our NYU collaborator, Professor Eero Simoncelli," Bovik said.

The students' efforts were critical, Bovik said, particularly Wang's development of the efficient SSIM model and Sheikh's NSF-supported large-scale human study of picture quality.

"NSF support of graduate students is a major reason for the success of many research projects," Bovik said.

## **From science to standards**

Early in Bovik's academic career—back in the mid-1980s—he began collaborating with visual psychologists and neuroscientists to better understand perception. In the process, he became an accomplished vision scientist.

With support from NSF, he applied discipline-crossing theories about how humans see and perceive objects and motion, and used them to accurately assess how naturalistic or distorted an image or video would appear to a human viewer.

Among the insights from visual perception theory that the SSIM model incorporates are:

- Contrast masking, where the texture or "busyness" causes image distortions to become less perceptible.
- Luminance masking, whereby distortion is less visible in brighter regions.
- And structural similarity, the idea that visually important structures, such as edges and details, are modified or destroyed by compression, which can also give rise to new "false" structures that are perceived as distortion.

Humans perceive visual distortions, such as compression, blur or noise, remarkably consistently with each other. When Bovik's team tested their SSIM index on the world's largest databases, including as many as 25,000 human judgments of quality, they found it successfully predicted human quality assessments to a larger degree than any previous algorithms. This was later confirmed on even larger datasets containing more than 250,000 human judgments of picture distortion.

"When distortions such as blur, noise, compression artifacts and channel errors appear, it's something that your brain responds to very quickly," Bovik said. "It's instantly annoying to some degree and that degree of annoyance is amazingly consistent among humans."

The SSIM model has proved indispensable to broadcasters and has become a de facto industry standard, widely commercialized around the globe in a variety of products.

Beyond its adoption as a standard tool in broadcast and post-production houses throughout the television industry, the SSIM index is part of the global ITU standard H.264 video coding reference software—one of the most commonly used formats for the recording, compression and distribution of video content. This allows developers everywhere to "SSIM-optimize" their encoder implementations and rate-control protocols.

"This research is an excellent example of how a deep understanding of basic science can lead to technological advances that have positive benefits across the globe," said Lynne Parker, division director for Information and Intelligent Systems at NSF. "Dr. Bovik's research team has translated a scientific understanding of human perception into vast improvements in our global, video-based communications technologies. Breakthrough advances such as the SSIM index illustrate the broad impact that is possible as a result of investment in fundamental,

interdisciplinary research."

## **Other applications**

Bovik and his students did not stop with SSIM. He has gone on to create many newer technologies for image processing and video quality assessment with NSF funding. These include MOVIE (the MOtion-based Video Integrity Evaluation) index for video quality assessment, the DIIVINE, BRISQUE, BLIINDS and NIQE no-reference image and video quality models and the LIVE Image and Video Quality Databases.

Recently, again with funding from NSF, he created algorithms that can assess the quality of images without an original reference image—an even trickier task. This research was inspired in part by work by a famous result in visual neuroscience that suggests that images of the real world, if they are of good quality, obey certain statistical laws.

"These laws can be used to explain the success of the SSIM model and of other algorithms that we've created, and we have also used them to design the new breed of 'blind' models that do not need a reference signal," Bovik said. "The human visual system has evolved in response to those statistics. By observing those statistics, we can predict how neurons will respond to the early vision system and vice-versa: by examining the brain, we can get insights into natural scene statistics. If you have a photograph or video and it doesn't obey those scene statistics, then it's likely been distorted."

Bovik and his team are applying this idea to a range of digital image quality issues. Their solutions range from algorithms can give digital cameras the ability to determine whether photographs are objectively "good," to systems that can significantly improve the accuracy of facial recognition, to advanced image capture methods using infrared and other wavelengths.

Provided by National Science Foundation

Citation: Engineer and vision scientist nets Emmy for tool to predict perceived video quality (2015, October 5) retrieved 12 September 2024 from <https://phys.org/news/2015-10-vision-scientist-nets-emmy-tool.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.