Reducing video files to one-tenth their initial size enables speedy analysis of laparoscopic procedures

June 19 2017, by Larry Hardesty

Researchers at MIT and Massachusetts General Hospital have designed a new system that can efficiently search through hundreds of hours of video of laparoscopic surgeries to identify relevant events and visual features. Credit: MIT News

Laparoscopy is a surgical technique in which a fiber-optic camera is inserted into a patient's abdominal cavity to provide a video feed that guides the surgeon through a minimally invasive procedure.

Laparoscopic surgeries can take hours, and the video generated by the camera—the laparoscope—is often recorded. Those recordings contain a wealth of information that could be useful for training both medical providers and computer systems that would aid with surgery, but because reviewing them is so time consuming, they mostly sit idle.

Researchers at MIT and Massachusetts General Hospital hope to change that, with a new system that can efficiently search through hundreds of hours of video for events and visual features that correspond to a few training examples.

In work they presented at the International Conference on Robotics and Automation this month, the researchers trained their system to recognize different stages of an operation, such as biopsy, tissue removal, stapling, and wound cleansing.

But the system could be applied to any analytical question that doctors deem worthwhile. It could, for instance, be trained to predict when particular medical instruments—such as additional staple cartridges—should be prepared for the surgeon's use, or it could sound an alert if a surgeon encounters rare, aberrant anatomy.

"Surgeons are thrilled by all the features that our work enables," says Daniela Rus, an Andrew and Erna Viterbi Professor of Electrical Engineering and Computer Science and senior author on the paper. "They are thrilled to have the surgical tapes automatically segmented and indexed, because now those tapes can be used for training. If we want to learn about phase two of a surgery, we know exactly where to go to look for that segment. We don't have to watch every minute before that. The

other thing that is extraordinarily exciting to the surgeons is that in the future, we should be able to monitor the progression of the operation in real-time."

Joining Rus on the paper are first author Mikhail Volkov, who was a postdoc in Rus' group when the work was done and is now a quantitative analyst at SMBC Nikko Securities in Tokyo; Guy Rosman, another postdoc in Rus' group; and Daniel Hashimoto and Ozanan Meireles of Massachusetts General Hospital (MGH).

Representative frames

The new paper builds on previous work from Rus' group on "coresets," or subsets of much larger data sets that preserve their salient statistical characteristics. In the past, Rus' group has used coresets to perform tasks such as deducing the topics of Wikipedia articles or recording the routes traversed by GPS-connected cars.

In this case, the coreset consists of a couple hundred or so short segments of video—just a few frames each. Each segment is selected because it offers a good approximation of the dozens or even hundreds of frames surrounding it. The coreset thus winnows a video file down to only about one-tenth its initial size, while still preserving most of its vital information.

For this research, MGH surgeons identified seven distinct stages in a procedure for removing part of the stomach, and the researchers tagged the beginnings of each stage in eight laparoscopic videos. Those videos were used to train a machine-learning system, which was in turn applied to the coresets of four laparoscopic videos it hadn't previously seen. For each short video snippet in the coresets, the system was able to assign it to the correct stage of surgery with 93 percent accuracy.

"We wanted to see how this system works for relatively small training sets," Rosman explains. "If you're in a specific hospital, and you're interested in a specific surgery type, or even more important, a specific variant of a surgery—all the surgeries where this or that happened—you may not have a lot of examples."

Selection criteria

The general procedure that the researchers used to extract the coresets is one they've [previously described](#), but coreset selection always hinges on specific properties of the data it's being applied to. The data included in the coreset—here, frames of video—must approximate the data being left out, and the degree of approximation is measured differently for different types of data.

Machine learning can be thought of as a problem of approximation, however. In this case, the system had to learn to identify similarities between frames of video in separate laparoscopic feeds that denoted the same phases of a surgical procedure. The metric of similarity that it arrived at also served to assess the similarity of [video](#) frames that were included in the coreset, to those that were omitted.

"Interventional medicine—surgery in particular—really comes down to human performance in many ways," says Gregory Hager, a professor of computer science at Johns Hopkins University who investigates medical applications of computer and robotic technologies. "As in many other areas of human endeavor, like sports, the quality of the human performance determines the quality of the outcome that you achieve, but we don't know a lot about, if you will, the analytics of what creates a good surgeon. Work like what Daniela is doing and our work really goes to the question of: Can we start to quantify what the process in surgery is, and then within that process, can we develop measures where we can relate human performance to the quality of care that a patient receives?"

"Right now, efficiency"—of the kind provided by coresets—"is probably not that important, because we're dealing with small numbers of these things," Hager adds. "But you could imagine that, if you started to record every surgery that's performed—we're talking tens of millions of procedures in the U.S. alone—now it starts to be interesting to think about efficiency."

More information: Machine Learning and Coresets for Automated Real-Time Video Segmentation of Laparoscopic and Robot-Assisted Surgery. people.csail.mit.edu/rosman/pa … /icra_17_medical.pdf

Provided by Massachusetts Institute of Technology