

# Researchers design a type of Turing test for computer vision

March 10 2015, by Bob Yirka

---



Credit: MARK I photo album

A small team of researchers has developed a possible means for creating a Turing test for computer vision. In their paper published in *Proceedings*

*of the National Academy of Sciences*, Donald Geman, Neil Hallonquist and Laurent Younes with Johns Hopkins University and Stuart Geman with Brown University describe their idea of using a query engine to create a series of yes/no questions to show how well another computer system is able to understand content appearing in a photograph.

Image recognition has improved markedly in recent years, anyone with a cell phone knows that, little boxes appear around faces, etc., helping to focus on what is important. But, as the authors in this new study note, recognizing parts of pictures is not the same thing as understanding what is being shown. A picture of people smiling while on vacation offers very different information than a group of people surveying damage after a tornado, for example. Thus far, not much progress has been made in making computers understand what is going on with a picture—the researchers with this work hope to change that by offering a new kind of Turing test.

A Turing test, is of course a way to test a computer system on its ability to mimic human thought—a person sits down at a computer and queries the system and then after a while decides if the answers coming back are being given by the computer or if they are being fed to it by another person out of view. The system passes the test when a human cannot tell the difference. Geman et al envision a similar system for computer vision, only instead of verbal information, the test seeks to measure how well a [computer system](#) can mimic the human ability to pick out information in photographs or perhaps video, or things going on in real life.

The researchers envision a binary query engine—software that generates questions that can be answered by only yes or no, e.g. is there a person in an identified part of the picture? The query engine is then to be connected to [image recognition](#) software (IRS) to see how well it scores. In practice, the query engine would first be run with a human being to

ascertain the "correct" answers or to get rid of those that are ambiguous. As the question/answer session continues, the query engine records the answers given by the IRS and is also given the correct answer by a human being. In this way the query engine is able to slowly customize its questions based on the information it has received. In the end the IRS receives a score based on its ability to understand what is going on in the test picture.

Developing such a test is important for developing intelligent image recognition systems, without them, people that create them are forced to proceed based only on assumptions.

**More information:** Visual Turing test for computer vision systems, *PNAS*, Donald Geman, [DOI: 10.1073/pnas.1422953112](https://doi.org/10.1073/pnas.1422953112)

## **Abstract**

Today, computer vision systems are tested by their accuracy in detecting and localizing instances of objects. As an alternative, and motivated by the ability of humans to provide far richer descriptions and even tell a story about an image, we construct a "visual Turing test": an operator-assisted device that produces a stochastic sequence of binary questions from a given test image. The query engine proposes a question; the operator either provides the correct answer or rejects the question as ambiguous; the engine proposes the next question ("just-in-time truthing"). The test is then administered to the computer-vision system, one question at a time. After the system's answer is recorded, the system is provided the correct answer and the next question. Parsing is trivial and deterministic; the system being tested requires no natural language processing. The query engine employs statistical constraints, learned from a training set, to produce questions with essentially unpredictable answers—the answer to a question, given the history of questions and their correct answers, is nearly equally likely to be positive or negative. In this sense, the test is only about vision. The system is designed to

produce streams of questions that follow natural story lines, from the instantiation of a unique object, through an exploration of its properties, and on to its relationships with other uniquely instantiated objects.

[Press release](#)

© 2015 Tech Xplore

Citation: Researchers design a type of Turing test for computer vision (2015, March 10)  
retrieved 9 April 2024 from <https://techxplore.com/news/2015-03-turing-vision.html>

<p>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.</p>
--