

# Teaching Computers to See



Original picture

Computer vision: Weizmann technique

Computer vision: other techniques

Computers have difficulty finding objects in a picture. In the pictures appearing in the two columns on the right, each color signifies a different object as perceived by the computer. The Weizmann technique enables the computer to identify the contours of the leopard and polar bear almost perfectly, while other methods cause the computer to perceive the animals as many separate objects (the leopard's mouth, for instance, is interpreted as being part of the log).

**Skeptic:** Billions of nerve cells in the human brain are involved in the process of vision. Since scientists find it difficult to understand even one of those cells, it's inconceivable that computers will ever be able to see.

**Scientist:** All the more reason to start working on it.

If one could make a computer see, reproducing the natural visual process, one would be adding an important dimension to many aspects of life: Computers could lend an "extra eye" in the operating room, alerting doctors to problems; surveillance systems would be greatly improved; and quality control in production lines could be significantly enhanced.

Along the way, we might attain crucial insights into how the brain constructs images.

The problem is the way computers read images. They see them as grids whose tiny squares (pixels) each give the computer only one type of information: color. Color, though, can be an ambiguous guide when it comes to distinguishing between objects. A yellow butterfly

on a yellow flower could be mistaken by the computer for part of the flower. A clown, displaying numerous colors, could be interpreted as many different objects. A zebra, having two alternating colors, could be split into the number of its stripes.

Recently, a team from the Weizmann Institute devised a method that significantly improves



computers' ability to see. The new approach goes from the bottom up – beginning with one pixel. Each pixel in the image is compared to surrounding pixels in terms of color. As groups of pixels emerge, they are compared to one another using

Color can be an  
ambiguous guide when  
it comes to distinguishing  
between objects

a wide range of more complex parameters: texture (the zebra's alternating black and white stripes emerge as a "texture" characterizing one object), shape (groups of pixels, in contrast to individual pixels, will form a shape), average fluctuations in color, and more.

Groups having common parameters are joined. The bigger the groups, the more the parameters can be made refined and complex. The different objects in the image are separated according to the combination of parameters, which are so diverse that the error range

is very small.

It doesn't sound as though any of this can be done in the blink of an eye, but in fact this method is much faster than all other existing methods. The reason: In the first stages, when individual or small groups of pixels are compared, only a few "simple" parameters are employed to compare them. At later stages, many more complex parameters are used to compare large groups of pixels, but by then there are much fewer groups to compare. Thus the complex parameters are also not as time consuming as might be expected.

The technique was developed by then Ph.D. student Eitan Sharon under the supervision of Profs. Achi Brandt and Ronen Basri and in collaboration with Dr. Meirav Galun, all in the Applied Mathematics and Computer Science Department. They are currently working to improve the approach.

Of course, even if a computer is finally able to distinguish easily between objects, other obstacles will have to be overcome before it is able to "see" as we do. For one thing, it will have to be taught

to interpret what it sees and to correctly categorize objects (for instance, to understand that a poodle and a German shepherd both belong in the "dog" category despite their different appearances). Thanks to the complexity of our

A clown, displaying  
numerous colors, could  
be interpreted as  
many different objects

brains, we perform these functions easily. Will scientists find a way to simplify these functions so that they can be incorporated into computers? It may take a while until we find out. ■

---

*Prof. Brandt's research was supported by the Karl E. Gauss Minerva Center for Scientific Computation. He is the incumbent of the Elaine and Bram Goldsmith Professorial Chair of Applied Mathematics.*

---

*Prof. Basri conducts his research in the Moross Laboratory for Vision Research and Robotics.*



(l-r) Prof. Ronen Basri, Dr. Meirav Galun and Prof. Achi Brandt. Computer vision



Dr. Eitan Sharon. Colorful Ph.D. work





משמאל: תמונות רגילות של נמר ודב. מימין: הדרך שבה מחשב מפרש את התמונות כשהוא מחלק אותן לאובייקטים שונים (כל צבע מייצג אובייקט נפרד לפי "הבנתו" של המחשב). במרכז: תמונות הנמר והדב כפי שמפרש אותן מחשב שמיישם את השיטה שפיתחו מדעני המכון

## דברים שרואים

מה ההבדל בין זברה לבין אוסף מקרי של סרטים שחורים ולבנים?

לקחת את ידי בידך ואמרת לי  
דברים שרואים מכאן, לא רואים משם

מלים: יעקב רוטבליט  
לחן: מתי כספי  
ביצוע: יהודית רביץ

בחדרי ניתוח ותזעיק את הרופאים כשבעיות כלשהן יחלו להתהוות; יעילותן של מערכות אבטחה תשופר ללא הכר; חנויות גדולות יוכלו להשתמש במחשבים כדי לזהות כשלי תחזוקה או לעקוב אחר חשודים בגניבה;

מדען: זו סיבה טובה להתחיל לעבוד על זה. אם אפשר היה לגרום למחשב לראות, לחקות את תהליך הראייה הטבעי, היו נוספות לחיינו אפשרויות רבות ומלהיבות. מחשבים יוכלו להוות "עין נוספת" שתפעל

פקן: מיליארדי תאי עצב פועלים יחד במוח בתהליך הראייה. מכיוון שכל תא עצב הוא מערכת מורכבת מאוד בפני עצמה, נראה שלא יהיה זה מציאותי לצפות שמחשבים יוכלו לראות.



ומפעלי תעשייה יפעילו "עיניים ממוחשבות" כדי לזהות תקלות בפסי ייצור. ואם לא די בכל אלה, הרי שמחקר בתחום זה עשוי לשפוך אור על הדרך שבה המוח שלנו מעבד את הקלט העצבי המגיע אליו מהעיניים שרואות את התמונה שלפנינו. הבעיה היא בדרך שבה מחשבים "קוראים"

התהליך החדש של עיבוד המידע החזותי אמנם מורכב למדי, אבל על אף זאת הוא מבוצע במהירות רבה יותר בהשוואה לזמן שנדרש לתהליכים אחרים לעיבוד תמונה, שמידת מורכבותם דומה

דימויים ותמונות. הם רואים אותם כצירוף של ריבועים קטנים ("פיקסלים"), שנושאים רק סוג אחד של מידע: צבע. אבל צבע עשוי להיות מדריך כוזב כשמדובר בהבחנה בין אובייקטים. פרפר צהוב הנמצא על פרח צהוב עשוי להיראות למחשב כחלק מהפרח. ליצן, המציג צבעים רבים, יכול להיראות כמספר אובייקטים שונים. זברה, שלה שני צבעים המתחלפים לסירוגין, יכולה

להיראות למחשב כאוסף של סרטים שאין ביניהם קשר.

כדי לפתח את יכולתם של מחשבים לראות החלו כמה מדענים ממכון ויצמן למדע לפעול באסטרטגיית פעולה המתנהלת "מלמטה כלפי מעלה". באסטרטגיה זו תהליך ניתוח המידע מתחיל מפיקסל בודד. כלומר, בשלב הראשון מבצע המחשב השוואה בין כל פיקסל בתמונה למשנהו, לפי צבע. כשנוצרות קבוצות פיקסלים המאורגנות לפי הצבע, עובר המחשב להשוואה בין הקבוצות, אלא שהפעם הוא מוסיף להשוואת צבע מדדים מורכבים יותר: טקסטורה (הפסים השחורים והלבנים של הזברה מתגלים כ"טקסטורה" הקיימת בתוך אובייקט אחד ויחיד); צורה (כל פיקסל בודד דומה בתכלית לכל פיקסל בודד אחר, אבל קבוצות של פיקסלים יוצרות צורות רבות ושונות); ועוד. קבוצות שמדדים רבים משותפים להן ייחשבו על-ידי המחשב כחלקים שונים של אובייקט אחד, או של קבוצה גדולה שתשווה לקבוצות גדולות אחרות. ככל שהקבוצות גדולות יותר, המדדים להשוואה ביניהן נעשים רבים ומורכבים יותר. ריבוי הגורמים ומורכבותם מקטין מאוד את טווח הטעות האפשרית של המחשב המפרש את התמונה הנצפית. בדרך זו הצליחו מדעני המכון להשיג שיפור ניכר ביכולתו של המחשב להבחין בין אובייקטים שונים. את המחקר ביצע ד"ר איתן שרון, במסגרת עבודת דוקטורט שהכין בהנחייתם של פרופ' אחי ברנד

ופרופ' רונן בצרי, בהשתתפות ד"ר מרב גלון, כולם מהמחלקה למדעי המחשב ומתמטיקה שימושית במכון ויצמן למדע. התהליך החדש של עיבוד המידע החזותי אמנם מורכב למדי, אבל על-אף זאת הוא מבוצע במהירות רבה יותר בהשוואה לזמן שנדרש לתהליכים אחרים לעיבוד

כדי לפתח את יכולתם של מחשבים לראות החלו כמה מדענים ממכון ויצמן למדע לפעול באסטרטגיית פעולה המתנהלת "מלמטה כלפי מעלה". באסטרטגיה זו תהליך ניתוח המידע מתחיל מפיקסל בודד

תמונה, שמידת מורכבותם דומה. גורם היתרון כאן הוא המיעוט היחסי של מדדים המושווים בשלביו הראשונים של התהליך, כאשר המחשב משווה פיקסלים בודדים או קבוצות קטנות. התוצאות המתקבלות בחישובים ראשוניים אלה מסייעות באופן מצטבר להשגת הפתרונות בשלבים המתקדמים והמורכבים יותר של התהליך. I



פרופ' אחי ברנד, ד"ר מרב גלון, ופרופ' רונן בצרי. תמונה שלמה



ד"ר איתן שרון. חלוקה למרכיבים