

Mira, màquina

Al Centre de Visió per Computador s'investiga perquè les màquines siguin capaces de veure i entendre imatges. Aquesta intel·ligència òptica té moltes aplicacions presents i futures, des de cotxes capaços de circular sols fins a programes informàtics que puguin detectar una patologia o identificar la perpetració d'un delicte a partir d'una imatge.

Reportatge de Àstrid Bierge

Una de les facultats del cervell humà que més reptes planteja en el camp de la intel·ligència artificial és la capacitat que té de reconèixer patrons. El nostre cervell és un as relacionant elements similars, com una vorera de Manhattan i una vorera de Bangkok o un número cinc imprès i un número cinc manuscrit. Encara que no siguin exactament iguals, associem els dos patrons i entenem que corresponen a un mateix valor. Aparentment pot semblar una tasca senzilla, però el cert és que als programadors informàtics els costa Déu i ajuda dissenyar programari capaç d'imitar aquesta habilitat, que requereix un gran nivell d'intel·ligència. Les instruccions que cal donar a l'ordinador, els algorismes, són molt complexes. És per això que, navegant per internet, sovint se'ns fa omplir un test Captcha. Si sabem identificar els caràcters deformats, probablement és que som humans.

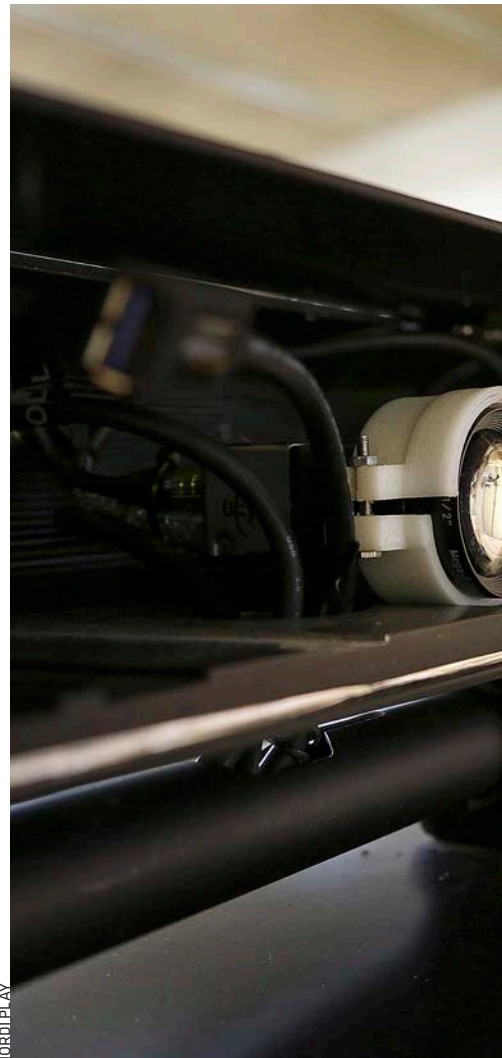
Alguns programes informàtics ja saben reconèixer òpticament certs patrons. Per exemple, qualsevol càmera ja sap identificar automàticament on hi ha una cara i l'emmarca amb aquell petit quadrat virtual. A més, fa molts anys que les màquines industrials poden veure

peces o objectes i determinar si tenen cap deformitat.

El camí que falta recórrer, però, encara és llarg. El Centre de Visió per Computador (CVC) és un centre Cerca que investiga precisament en aquesta direcció. La visió per computador és un camp de la intel·ligència artificial que té com a objectiu programar els ordinadors perquè *entenguin* una escena o les característiques d'una imatge. Dit d'una altra manera, es tracta de fer que les màquines hi vegin.

Fernando Vilariño, sots-director del CVC, en dóna un exemple molt gràfic: "Una de les nostres línies de recerca consisteix a ensenyar a un ordinador a descriure escenes. És a dir, que miri una imatge i sigui capaç de determinar, per exemple, que és una escena campestre on hi ha dues persones que probablement estan enamorades". I això com es fa? Doncs creant algorismes perquè l'ordinador pugui detectar en una imatge elements identificadors. En aquest cas, caldria haver-li ensenyat que una imatge amb molt de verd correspon al camp i que dues cares molt juntes denoten amor.

Aquesta tecnologia seria molt útil per a la videovigilància. Si una càmera capta dins d'un garatge una persona amb acti-



JORDI PLAY

tuds sospitoses o una baralla al carrer, un programa informàtic podria identificar el perill i donar el senyal d'alarma.

La visió per computador és un camp transversal i multidisciplinari que té aplicacions en àmbits molt diversos. Imaginem un arxiu de milers de pàgines que contenen registres matrimonials manuscrits. Si un ordinador fos capaç d'analitzar les imatges i identificar les diferents dades que apareixen a cada registre -noms propis, professions, adreces, etc.-, es podrien omplir fitxes sense la necessitat de picar-les manualment i, per tant, es podrien crear bases de dades de manera automàtica. Això encara no es pot fer, però al CVC s'hi estan acostant. En un projecte fet en col·laboració amb el Centre d'Estudis Demogràfics, s'han digitalitzat tots els llibres de matrimonis dels últims cinc segles que hi ha a l'arxiu de la Catedral de Barcelona. El CVC ha



Un grup de recerca del CVC investiga en el disseny de software que, a partir d'imatges captades per càmeres, permeti els vehicles entendre el seu entorn i circular de manera autònoma.

El grup de Modelatge Interactiu i Augmentat del CVC està especialitzat en el tractament i l'anàlisi d'imatges biomèdiques. Dues investigadores del grup, Debora Gil i Agnès Borràs, ens parlen d'un projecte per ajudar els metges a planificar les biòpsies de pulmó. Si una radiografia detecta nòduls al pulmó, cal agafar-ne una mostra per veure si són benignes o malignes. Una tècnica molt poc invasiva per aconseguir aquesta mostra és la navegació endoscòpica. El metge introdueix per la boca del pacient un broncoscopi –un tub amb una càmera a la punta– i va empenyent-lo de manera manual per l'arbre bronquial del pacient. Encara que el metge vegi per on està anant, no és gens fàcil desplaçar-se per aquest terreny. Hi ha moltes desviacions i els canals són estrets. A més, cada persona té les seves pròpies ramificacions, el camí sempre és diferent.

A partir de l'escàner dels bronquis del pacient, abans de la intervenció el metge planifica la ruta per arribar al seu destí. “Es fa un mapa mental: la primera a la dreta, la segona a l'esquerra, etc.”, diu Gil. El problema és que la ruta no sempre és viable. El broncoscopi té un funcionament mecànic i hi ha llocs on no pot arribar, no sempre es podrà anar retorçant per l'entortolligat arbre bronquial tal com marca el pla del metge. Potser hi ha un canal que és massa estret o que té un angle massa forçat.

Aquí és on entra l'equip del CVC, que ha ideat una manera perquè els metges puguin provar el camí abans de recórrer-lo. En primer lloc, han dissenyat un programari capaç de *mirar* l'escàner del pacient i a partir d'aquí crear un model digital tridimensional dels seus bronquis. En segon lloc, han dissenyat un videojoc perquè el metge pugui practicar navegant amb el broncoscopi a través dels bronquis digitals del seu pacient. El controla amb un *joystick*! Aquesta simulació –que segueix les lleis físiques de la realitat– li permet saber si podrà o no arribar al nòdul i planificar la millor ruta. La Secció d'Endoscòpia Respiratòria de l'Hospital Universitari de Bellvitge ja ha incorporat aquest sistema. →

dissenyat un programa gràcies al qual l'usuari pot seleccionar una paraula d'un dels documents i trobar els altres punts de l'arxiu on surt aquesta paraula. Cal tenir en compte que el text és manuscrit, aquí rau la dificultat perquè els caràcters escrits a mà sempre són diferents. És imprescindible que l'ordinador reconegui patrons òptics. Aquesta tecnologia pot ser molt útil com a motor de cerca d'informació per a qualsevol tipus d'arxiu que hagi estat prèviament escanejat.

Un tercer ull clínic

Actualment ja hi ha programes informàtics que poden identificar patologies a partir d'una imatge mèdica. És una eina que facilita molt la feina del metge a l'hora de fer un diagnòstic. El CVC ha treballat amb tots els hospitals de l'àrea de Barcelona i ha fet patents amb diverses grans empreses del sector, com

Boston Scientific o Medtronic & Covidien. La darrera, una empresa d'Israel, domina el mercat mundial de la càpsula endoscòpica, una píndola amb una càmera que grava el seu recorregut pels intestins. El vídeo, però, dura cinc hores i no és gaire divertit de veure. “A més, un intestí malalt està alterat i per tant tot costa més de veure”, diu Vilariño, expert en aquest camp de la visió per computació. Un projecte conjunt entre aquesta empresa, el CVC, l'Hospital Vall d'Hebron i la UB, va idear una solució. Es va dissenyar un programa que, a partir del vídeo enregistrat per la càpsula, sap comptar les contraccions intestinals i identificar si hi ha molta o poca quantitat de líquids intestinals, dos factors que són molt significatius en la diagnosi de malalties. Amb un informe que mostri aquests valors, els metges s'estalvien el visionat complet del vídeo.

Amb aquest sistema d'interacció persona-document desenvolupat pel CVC, una càmera reconeix un document manuscrit, un programa informàtic comprèn alguns aspectes de la imatge i un projector afegeix informació virtual.

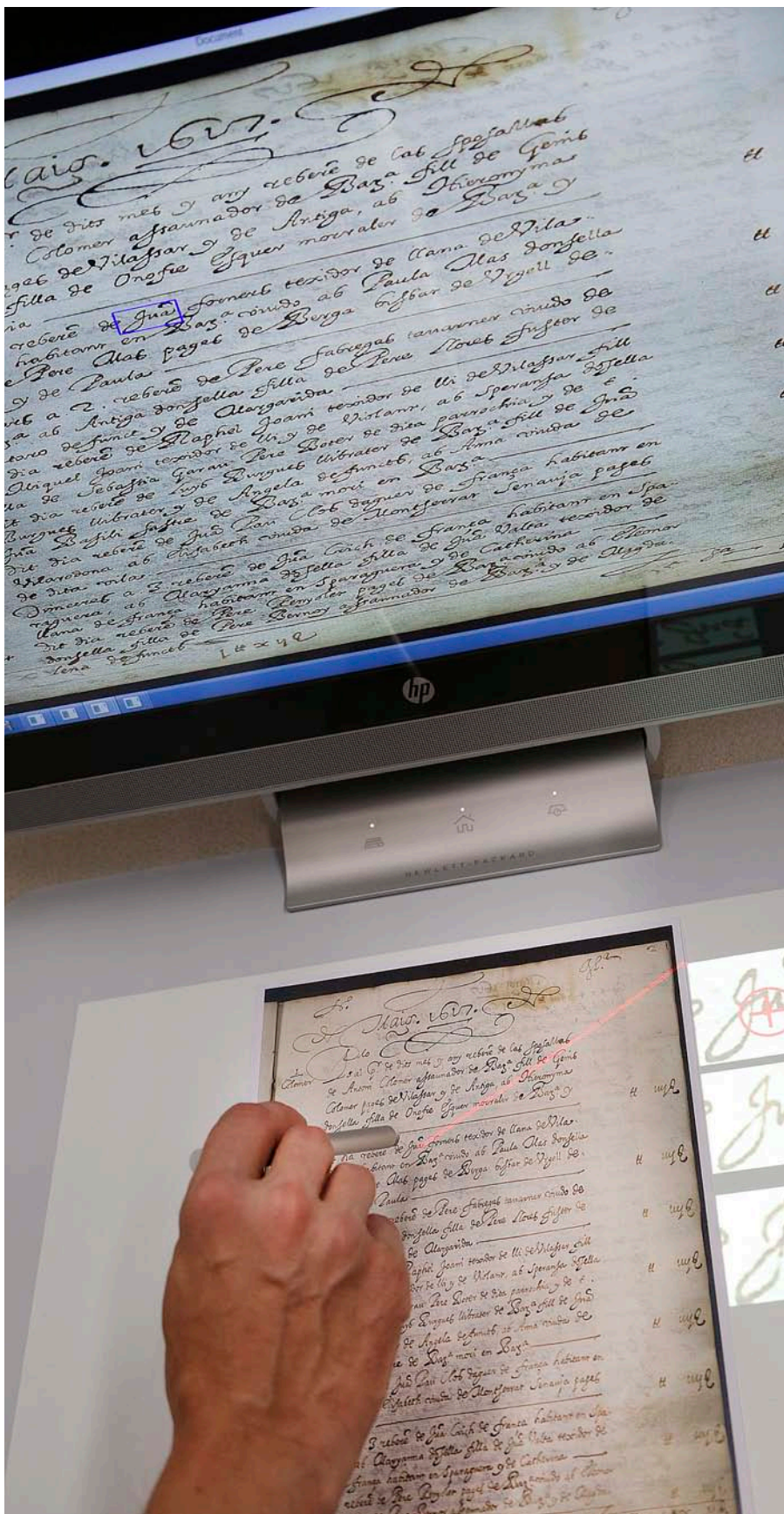
→ **Cultura 2.0**

En un espai de la Biblioteca Miquel Batllori de Sant Cugat s'hi exploren maneres completament noves de gaudir de la cultura a través de tecnologia desenvolupada al CVC. És el que s'anomena un *living lab*, un nou concepte de recerca, desenvolupament i innovació que involucra els usuaris en totes les fases del procés. És com un banc de proves reals on els ciutadans, els investigadors i els productors experimenten i cocreen innovacions. Al *Library living lab* -el nom oficial de l'espai- hi ha tota una sèrie d'equipaments tecnològics per desenvolupar molts tipus d'activitats: un conjunt d'ordinadors, pantalles tàctils, una impressora 3D, una multipantalla a la paret, etc.

També hi tenen dos Kinects, que són sensors de moviment. I no cal agafar un comandament, com amb la Wii, saben captar la posició de les articulacions de les persones. Per exemple, en el taller pilot "Jo sóc el meu dibuix", organitzat pel *Library living lab*, un grup de nens va poder donar vida als seus dibuixos amb els moviments del seu cos. Primer van idear i dibuixar entre tots un conte. Llavors se'n van digitalitzar els dibuixos i es van anar projectant a la multipantalla. Quan un nen es posava en el camp de percepció dels sensors Kinect i es movia, el personatge digitalitzat imitava els seus gestos. Cal aclarir que l'únic que fan els sensors és captar el moviment. Perquè els dibuixos es moguin com els nens, cal crear un programa informàtic capaç de comprendre les imatges generades pels sensors, i aquí és on entra la visió per computador. Si les Kinect són els ulls del sistema, el programari n'és el cervell.

Una de les *spin-off* del CVC utilitza els sensors Kinect per monitoritzar les persones grans amb risc de caure. Un programari capaç d'interpretar els moviments captats pels sensors podria veure que una persona ha caigut de manera sobtada i que no s'ha pogut aixecar.

Un dels experiments científics del *Living lab* de Sant Cugat, liderat per Marta Expósito, del CVC, pretén ensenyar el



JORDI PLEY



Al taller "Sóc el meu dibuix", organitzat al *Library living lab* de la Biblioteca Miquel Batllori de Sant Cugat, els nens van poder animar els seus dibuixos amb els moviments del cos. L'activitat es va dur a terme amb l'artista Enric Socias.

LIBRARY LIVING LAB

bon gust als ordinadors. Però, com se li explica a una màquina què és bonic i què és lleig? Cal començar definint el concepte de bellesa que tenen els humans. En aquest experiment es van mostrar una sèrie d'imatges a un grup de persones perquè en valoressin la bellesa de l'u al deu. Amb aquests valors es poden intentar definir alguns dels paràmetres associats a les qualitats estètiques. Aquesta tecnologia podria servir, per exemple, perquè un cercador mostrés primer les imatges més boniques.

Vilariño pronostica maneres molt més interactives d'experimentar la cultura, algunes de les quals ja s'estan estudiant al CVC. Per exemple, a través d'un sistema de seguiment del moviment dels ulls –aquest maquinari ja existeix– es podrien crear obres d'art digitals que reaccionessin i canviessin en funció de com són observades. També es podria fer amb pel·lícules i obres de teatre: "En funció de què està cridant més l'atenció de la majoria dels espectadors, alguns elements podrien canviar, com ara la il·luminació, els plans i fins i tot la trama".

El cotxe mans lliures

Un dels camps de recerca més coneguts i complexos de la visió per computador

és el dels cotxes autònoms. Al CVC hi tenen un prototipus per poder provar el programari que dissenyen.

"Fer que un cotxe es mogui completament sol és molt complicat, s'ha de dotar el vehicle de diversos nivells de comprensió", explica Antonio López, cap del grup de Sistemes Avançats d'Assistència en Conducció (ADAS) del CVC. Anant de baix a dalt, de primer cal que el cotxe sàpiga rebre i complir les ordres bàsiques que li envia l'ordinador. Si l'ordinador li diu que freni, acceleri o giri tants graus, el cotxe ha d'entendre-ho i obeir. En segon lloc, cal ensenyar a l'ordinador a decidir quan i com ha de frenar, accelerar o girar. Per exemple, se li ha d'ensenyar que, si al davant té un vianant, cal que freni. Naturalment, per poder frenar primer cal que el vegi, i aquesta és la tercera habilitat que ha de tenir el sistema, la de la percepció. Ha de saber *veure* l'entorn.

L'últim nivell de comprensió que ha de tenir el programari d'un cotxe autònom ja pertany al terreny de la intel·ligència artificial més sofisticada. Per exemple, cal que aprengui a tenir en compte les normes de circulació. També ha de saber prendre decisions en situacions extremes: si un vianant apareix de

sobte i el cotxe calcula que no té temps de frenar, què ha de fer? Cal haver-li donat una sèrie d'instruccions perquè ho pugui decidir. Aquests algorismes són molt difícils de dissenyar, és molt complicat crear un manual d'instruccions prou vast i precís perquè el cotxe pugui funcionar completament sol.

El problema encara no està solucionat, hi ha un buit en el disseny del programari que cal anar omplint per mitjà de la recerca. És veritat que alguns cotxes comercials semiautònoms apaquen sols, es mantenen al carril sense que el conductor hagi de tocar el volant o frenen abans de xocar, però encara necessiten la col·laboració humana. El Germán Ros, que fa el doctorat en aquest grup del CVC, explica que cal millorar en el terreny del que s'anomena *l'scene understanding*, o comprensió de l'escena: "Els sistemes d'assistència a la conducció dels cotxes semiautònoms no tracten de *comprendre* l'escena, només resolen problemes concrets. Tenen una combinació de sensors –radars, sonars i càmeres– que realitzen algunes mesures. En funció dels valors que obtenen, actuen seguint unes regles predefinides. Unes regles aproximades i senzilles".

Perquè un vehicle pugui ser veritablement autònom, diu, ha de ser prou intel·ligent per comprendre tots els objectes que té al voltant, des de les persones fins als senyals de trànsit, ha de saber a quina distància són, a quina velocitat i en quina direcció es mouen, etc. I ho han de saber amb una precisió altíssima, en tots els casos de circulació possibles, en qualsevol condició climàtica i a tots els països del món. El grup ADAS del CVC és capdavanter a nivell mundial en la investigació de *l'scene understanding*, el gran repte dels vehicles autònoms i un dels grans reptes de la intel·ligència artificial. Mirar, interpretar i actuar. Qui sap si algun dia aconseguirem ensenyar aquest procés tan complex, tan humà, a les nostres màquines. ●