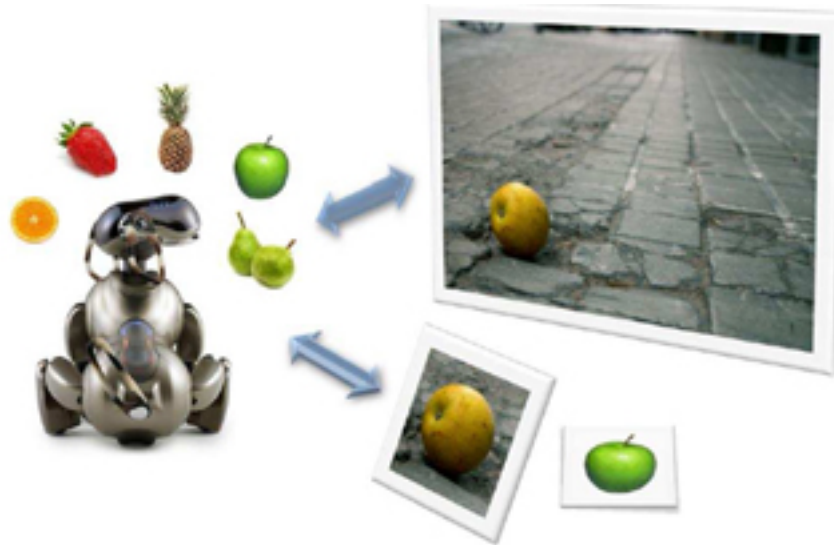


# "Blurred Shape Model": innovador reconeixement automàtic d'objectes

03/2010 - **Telecomunicacions, Electrònica i Informàtica.**

Una de les dificultats que trobem al reconeixement automàtic d'imatges prové de les alteracions que sofreixen els objectes en la seva representació. El fonament dels actuals sistemes intel·ligents consisteix a extraure informació rellevant de l'objecte, com podria ser el seu contorn, i aprendre que certes combinacions d'informació corresponen a objectes determinats. Investigadors del Centre de Visió per Computador, han proposat una nova metodologia per a la descripció d'objectes i el seu aprenentatge: el premiat descriptor "Blurred Shape Model" (BSM) considera la relació entre el nivell de difuminació del contorn i la facilitat per ser après com a un objecte determinat. Els usos d'aquesta nova metodologia van des de l'anàlisi de textos per ser processats, fins a la robòtica, passant per l'anàlisi de contingut d'imatges a internet.



Descripció, detecció i classificació d'objectes en robòtica.

El reconeixement automàtic d'objectes en imatges porta prop de 50 anys sent estudiat per les àrees d'Intel·ligència i Visió Artificial. No obstant això, encara que per a nosaltres reconèixer un objecte en una imatge és una mica trivial, encara no hem aconseguit desenvolupar un procés computacional que sigui capaç d'obtenir uns resultats mínimament semblats al nostre sistema visual.

Entre molts altres motius, el principal problema prové de les alteracions que sofreixen els objectes quan apareixen en imatges. Encara que nosaltres som robusts a la detecció d'objectes encara que les imatges continguin soroll, canvis en la il·luminació, ombres, oclusions, diferents punts de vista, etc., incloure aquest coneixement a un sistema computacional encara és una tasca complexa. Podríem dir que la majoria dels sistemes intel·ligents actuals intenten imitar el nostre comportament visual a partir de dos passos bàsics: descripció i aprenentatge. Primer hem d'extreure la informació rellevant de les imatges (contorns per exemple) i després aprendre que una combinació determinada d'informació correspon a un objecte en particular.

Al treball Blurred Shape Model (BSM), hem proposat una nova metodologia per a la descripció d'objectes així com per al seu aprenentatge. En particular, ens hem centrat en imatges binàries i en escales de grisos per a detectar diferents tipus de símbols, tant en imatges en entorns oberts com en documents manuscrits. En el primer escenari cal tractar la problemàtica dels canvis a l'ambient així com sobre els punts de vista. En el segon escenari, el principal problema és tractar la variació que els símbols sofreixen a causa dels diferents tipus d'escriptura de diferents autors.

La primera proposta d'aquest treball, el descriptor Blurred Shape Model (que va rebre el premi al millor treball en el Iberian Conference on Pattern Recognition and Image Analysis de 2007), representa els objectes a partir de la codificació de les relacions espacials entre les seves parts més representatives. Per exemple, suposem com a punts representatius de l'objecte aquells píxels que formen part dels contorns. Llavors, el BSM defineix una reixeta de regions i per a cada punt de contorn calcula la inversa de les distàncies a la resta de punts de contorn pròxims a ell. D'aquesta forma, es codifiquen relacions espacials i la grandària de la reixeta sobre l'objecte defineix el nivell de difuminació que volem que aprengui el sistema intel·ligent. En general, difuminacions grans fan que s'aprenguin bé els objectes però es podrien arribar a confondre amb uns altres. Per altra banda, difuminacions reduïdes serien massa específiques i no podríem trobar objectes semblants que

hagin sofert algun tipus de deformació. Per aquest motiu, el mètode testeja tots els valors de difuminació possibles sobre unes dades d'aprenentatge per adaptar els paràmetres a les necessitats de cada problema. Seria com ensenyar una imatge darrere d'una altra a un nen d'un any, amb la finalitat d'incloure nova informació a la seva font de coneixement. La Figura 1 mostra un exemple dels models difuminats per als contorns d'un objecte de tipus poma.



Una vegada el sistema detecta els nivells de desenfocament o difuminació dels objectes a aprendre, tècniques d'Intel·ligència Artificial són usades per a aprendre a diferenciar entre els objectes. Per exemple, vam utilitzar un classificador estadístic que aprèn a diferenciar models difuminats de pomes de models difuminats de peres, un altre que separi pomes de plàtans, etc. Finalment, si combinem les respostes d'aquests aprenentatges parcials, obtenim una metodologia multi-classe, que permet detectar de forma automàtica un conjunt d'objectes en imatges. En particular, la metodologia aplicada usa mètodes estadístics que se centren en trobar aquelles parts que millor separen uns objectes d'uns altres. A més, el sistema es beneficia de correcció d'errors mitjançant l'ús d'una classificació que es basa en la metodologia seguida en el camp de teoria de la informació. La figura de l'esquerra mostra un exemple de com un robot programat amb un sistema intel·ligent i amb una font de coneixement de descriptors difuminats de fruita, pot detectar les zones d'interès en una imatge i classificar l'objecte d'entre un conjunt de possibilitats.

Els resultats d'aquesta metodologia sobre dades capturades manualment i de fonts públiques han resultat ser altament satisfactoris. Comparats amb els resultats dels mètodes que competeixen en aquest camp, s'han trobat millores estadísticament significatives. La Figura 2 mostra els resultats comparant diferents descriptors de l'estat de l'art amb el BSM i aplicats sobre dues metodologies d'aprenentatge diferents. Les dades consisteixen en 70 objectes diferents en imatges de la base de dades pública MPEG7. Es pot veure que la millora és considerable.

Descriptor	3NN	Gentle Adaboost
BSM	<b>65.79 (8.03)</b>	<b>77.93 (7.25)</b>
Zernike	43.64 (7.66)	51.29 (5.48)
Zoning	58.64 (10.97)	65.50 (6.64)
CSS	37.01 (10.76)	44.54 (7.11)
SIFT	29.14 (5.68)	32.57 (4.04)

El futur d'aquest treball se centra a incrementar el nombre d'objectes a diferenciar i a trobar característiques més rellevants que siguin codificades dintre del model difuminat, tals com el color o la textura. L'aplicabilitat d'aquesta nova metodologia va des de l'anàlisi automàtica de textos per al seu processat, fins a la robòtica, permetent la navegació i interactivitat de sistemes intel·ligents, així com per a l'anàlisi eficaç de contingut entre tots els milions d'imatges dels quals disposem en internet.

Sergio Escalera

Centre de Visió per Computador

"Blurred Shape Model for Binary and Grey-level Symbol Recognition". Sergio Escalera, Alicia Fornés, Oriol Pujol, Petia Radeva, Gemma Sánchez, and Josep Lladós. Pattern Recognition Letters, doi:10.1016/j.patrec.2009.08.001, 2009.