
8. MIRANT AMB UNS ALTRES ULLS: EL TRACTAMENT D'IMATGES

Marcelo del Giorgio, Xavier Llovet i Antònia Sanz*

8.1. DESCRIPCIÓ DE L'APARELL

L'experiment que ens ocupa és veure, i no només mirar; és interpretar la imatge observada, és extraure informació d'ella. L'instrument del qual ens servirem consta d'un ull i un cervell. L'ull ha d'ésser capaç d'adquirir les imatges; el cervell, d'interpretar-les.

L'instrument que ha d'ésser l'ull és, en essència, una càmera de vídeo, és a dir, una òptica adequada i una col·lecció enorme de sensors de llum ordenats com en una matriu en el pla on es projecta la imatge per l'òptica. Aquests sensors donen un senyal elèctric proporcional a la intensitat de la llum que els arriba. Un convertidor analògicodigital (A/D) digitalitza aquests senyals, és a dir, els quantifica en un cert nombre de nivells, que podran ésser emmagatzemats per un ordinador, «el cervell», el qual podrà finalment extraure la informació desitjada.

Així, en el cervell quedarà, no la imatge real, que és la intensitat de llum provinent d'una escena, o sigui una funció contínua (intensitat) de variables contínues (posició), sinó una imatge digital que és una quantificació (realitzada pel convertidor A/D) d'una presa de mostra espacial d'aquesta intensitat (corresponent a la matriu de sensors), o sigui, una funció discreta de variables discretes.

Com que l'ordinador que emmagatzema la informació imposa un lògica binària, el convertidor A/D es caracteritza pel nombre de bits

* Unitat de Tractament d'Imatges, Serveis Científicotècnics, Universitat de Barcelona.

de sortida (zeros o uns), que són normalment, quatre, vuit o setze.

Per a fer-nos-en una idea, en una càmera en blanc i negre, si el convertidor és de vuit bits, podem obtenir una imatge amb 256 nivells de grisos diferents; si, per contra, té setze bits, n'hi haurà 65.536. Aquí s'ha de recordar que l'ull humà sap distingir de trenta a seixanta grisos diferents. Un instrument típic té una xarxa de 512 files amb 512 sensors cadascuna, i amb un convertidor A/D de vuit bits, la qual cosa fa que cada imatge en blanc i negre amb 256 nivells de grisos ocupi un espai de 262.144 bytes en un ordinador.

Si el que es desitja és veure en color, s'ha de tenir en compte que la llum policromàtica es pot representar com una combinació additiva de tres colors primaris, i es pot definir una àlgebra de colors, concretament un espai vectorial. Llavors, en comptes de tenir un sol sensor per punt de la xarxa, hi ha una tripleta, i cada sensor de la tripleta és sensible a un dels colors base. D'aquesta manera, s'obtenen tres senyals independents a cada punt, i si, per exemple, disposem de convertidors A/D de vuit bits, tindrem, lògicament, vint-i-quatre bits per punt de la imatge i, consegüentment, haurem triplicat l'espai que ocupava una imatge en blanc i negre.

L'ordinador actua com a receptor de tota aquesta informació, la qual pot processar mitjançant mètodes matemàtics, així com visualitzar-la.

8.2. PRECISIÓ DE L'ÀPARELL DE MESURA

Quan parlem d'un aparell de mesurament, sempre tenim un instrument i un observador; del conjunt dependrà la precisió de les mesures.

8.2.1. *La precisió de l'instrument*

En el nostre cas, l'instrument és l'ull, o sigui la càmera i el convertidor A/D. Aquest instrument de mesura té una precisió intrínseca que dependrà de la bondat de l'òptica de la càmera, de la qualitat i la quantitat dels sensors elementals que hi disposem per a prendre mostra de la intensitat de llum, de l'electrònica associada, i del nombre de bits del convertidor que quantifica aquestes mostres. Hi ha, però, altres factors externs a l'instrument que empitjoren la precisió de la captació de la imatge:

— Factors ambientals, com ara voler veure amb poca llum, ja que sempre que es mesura radiació, el comptatge és un procés de Poisson, és a dir, un procés en el qual les fluctuacions són inversament proporcionals a l'arrel quadrada del nombre de fotons que arriben al sensor. Això fa que, per a una escena amb poca llum, s'obtingui una imatge de menor precisió (imatge nevada), la qual es podrà millorar augmentant el temps d'exposició o el flux de llum, que millorarà l'estadística del comptatge.

— Defectes de manipulació, com no enfocar correctament l'escena desitjada, o com moure la càmera, o l'escena, a l'hora d'adquirir la imatge.

— El fet que hi hagi gran distància entre l'instrument i l'escena; pensem, per exemple, en una escena a la Terra i una càmera instal·lada en un satèl·lit a gran altura.

8.2.2. *La precisió de l'observador*

L'observador és l'ordinador, amb l'avantatge de la seva invariabilitat temporal. Això ens condueix a pensar que els errors de mesura d'un objecte de la imatge deguts a l'observador no existeixen, però no és així. Hi ha decisions que ha de prendre que es basen en estimacions estadístiques, les quals originaran, en general, errors aleatoris, però poden també originar errors sistemàtics.

A més, en moltes ocasions les condicions de treball no estan fixades i al seu darrere es troba un operador (nosaltres), que és qui dona els paràmetres que definiran les pautes a seguir pel mètode, la qual cosa s'ha de comptar com un altre factor que modifica la precisió de la mesura.

Centrarem la nostra atenció en un cas particular de l'anàlisi d'imatges per concretar idees. Analitzarem la *suma de l'àrea d'una regió* dins la imatge. Per a aquest fi utilitzarem dos exemples.

8.2.2.1. Primer exemple

Prenem, per exemple, una imatge senzilla que correspon a un aliatge de Cu-Ag-S, vista en un microscopi electrònic de rastreig. Com es pot apreciar en la figura 1, hi ha tres fases ben determinades. Aquest fet es reflecteix en el fet que, si construïm un histograma (població de punts amb cada nivell de gris, on el nivell 0 cor-

respon al negre i el nivell 255 correspon al blanc), veurem tres pics ben diferenciats, com es pot apreciar en la figura 2, que corresponen a les tres fases. Ens proposem mesurar l'àrea ocupada per la fase de Ag, que és la més clara i que correspon al pic de nivell més alt. En primer lloc, haurem de determinar la frontera entre aquesta fase i la resta de la imatge, posant en blanc els píxels que pertanyen al seu pic i en negre tota la resta. És evident que, ara, hem de començar establint on són els límits del pic a l'histograma. Hem donat quatre possibles límits, que els mostrem en les figures 3, 4, 5 i 6. Un cop ressaltada la fase clara podem extraure'n informació, podem determinar els objectes que s'hi troben a través d'una cerca de tots els punts que representen una vora entre el blanc i el negre, tal com es pot veure en la figura 7. Ja amb els objectes extrets, és possible el càlcul de qualsevol paràmetre d'aquests, com són les àrees.

Tal com s'aprecia en les figures 3, 4, 5 i 6, en canviar els límits del pic a l'histograma variem les àrees de les regions blanques. És a partir d'això que ens adonem que som un observador testimoni, que donem a l'ordinador un criteri a seguir. En aquest exemple senzill que hem mostrat, l'error que es comet és de tipus sistemàtic.

Un detall que s'ha de tenir en compte és que, en quantificar l'àrea d'un objecte, ho fem en termes del punts que ocupa. D'aquí a donar-la en unitats comparatives, s'ha de determinar l'àrea en el pla de l'escena que ocupa un sensor elemental, que també originarà un possible error sistemàtic.

8.2.2.2. Segon exemple

Un altre exemple és el que presentem en la figura 8, on podem veure unes partícules que es diferencien respecte al fons, no pel to de gris com abans, el qual caracteritzava una fase, sinó pel patró de variació d'aquests tons, que consisteix en una variació des d'un to molt fosc fins a un to molt clar.

Prenem una funció adequada per ajustar el fons de la imatge. Perquè l'ajust sigui només amb els punts del fons, utilitzarem un procés que ajusta la funció en forma iterativa. En un primer pas, aquest procés iteratiu dona el mateix pes a cada punt; en els passos següents baixa el pes dels que tenen una diferència més gran amb l'ajust i, per contra, puja el dels que tenen una diferència menor. Tot això fins que la variació de la fluctuació mitjana sigui molt petita o que el nombre

de passos sigui molt gran. Amb això s'aconsegueix que l'ajust es faci principalment amb els punts que s'assemblen més a la funció proposada. Amb aquests requisits assolits podem transformar la imatge original en una nova imatge on els tons de gris de cada punt siguin proporcionals a la diferència entre la intensitat de la imatge original i la prevista per la funció ajustada, en unitats de desviacions estàndard (σ), que podríem anomenar *imatge de fluctuacions* (figura 9).

Per a distingir un objecte en aquesta imatge de fluctuacions, tenim en compte que, si un píxel és molt clar o molt fosc (per exemple, s'allunya del fons més de tres o quatre vegades la desviació estàndard, o sigui, més de tres o quatre unitats sigma), té poques probabilitats de pertànyer al fons, i moltes d'ésser un objecte. Aquests píxels, per tant, es consideren com a llavors de possibles objectes. Al voltant d'aquestes llavors, pot donar-se el cas que trobem píxels amb tons menys extrems (per exemple, fins a dues unitats sigma) que, per la seva proximitat a una llavor, tenen moltes probabilitats de formar part d'un objecte i no del fons.

Seguint aquest criteri, fem una nova imatge en què cada píxel prendrà un de tres nivells: si pertany a un objecte clar, el posem en blanc; si pertany a un de fosc, en negre, i si no, el posem en gris mitjà. Aquest procés s'anomena *ternarització*. Seguidament, les regions blanques i negres molt properes es connectaran i formaran un sol objecte, ja que el patró que els caracteritzava consistia en una variació del to de fosc a clar.

L'operador fixa els valors dels llindars abans esmentats, que determinen el resultat d'aquest mètode. Analitzem la variació del resultat que s'obté d'acord amb la decisió que pren l'operador.

Si per al llindar que fixa la presència d'una llavor escollim un valor de moltes unitats σ , la quantitat d'objectes que es troben és menor que si n'escollim un de més baix, ja que ens quedem només amb els objectes més brillants i els més foscos. Pel que fa als llindars que fixen quins píxels s'han d'agrupar al voltant de les llavors (llindar de nucleació), si escollim un valor de més unitats σ , s'obté com a resultat objectes menys extensos.

Podem veure els efectes esmentats sobre la imatge de la figura 9.

En la figura 10 hem pres un valor per a les llavors fosques i les clares de 15σ i per al llindar dels píxels nucleats de $7,5 \sigma$; en canvi, en la figura 11 hem canviat el valor de les llavors fosques a $12,5 \sigma$. Com podem apreciar, l'última imatge detecta més objectes foscos.

Per contra, en la figura 12 i la figura 13 hem posat per als llindars

de les llavors els mateixos valors que la figura 10, però hem canviat els llindars de nucleació. En la figura 12 hem posat 10σ , i en la figura 13, $12,5\sigma$. En aquest últim cas, els objectes són menys extensos.

Després d'aquest procés, i atès que els objectes queden totalment caracteritzats pel seu to blanc i negre, i el fons, per un to gris mitjà, estem en condicions d'aplicar el procediment del primer exemple sobre l'histograma, fent que el fons passi a negre i la resta, a blanc, cosa que ens permet obtenir imatges com la de la figura 14, i extraure finalment les vores dels objectes, com es veu en la figura 15, amb les quals calculem la seva àrea.

Per tal d'arribar a aquest punt, hem passat dues etapes, que han generat una degradació de la informació. En la primera, o sigui, quan hem passat a una imatge de fluctuacions, hem introduït un error estadístic, ja que el procés d'ajust utilitzat és intrínsecament aleatori. En la segona etapa, quan escollim els diferents llindars a fi de detectar els objectes i la seva extensió, hem generat errors de tipus sistemàtic, com resulta del fet que a mesura que el llindar de nucleació creix, l'extensió dels objectes decreix.

8.3. CONCLUSIONS

A partir d'aquests exemples podem observar de quina manera es poden simular alguns processos elementals del *veure*. Hem simulat la mesura de les dimensions d'objectes, la qual cosa implica el reconeixement dels seus límits. Aquest problema involucra el de l'existència d'un valor *exacte*, que és la causa dels errors sistemàtics que mencionàvem. De fet, és el mateix que ens passa quan volem mesurar la longitud d'un objecte, ja que la transició de l'objecte al fons és gradual; no hi ha un tall net.

Hem vist de quina manera podem utilitzar eines estadístiques per a emular el que fa *trivialment* el nostre cervell a l'hora d'identificar objectes. Perquè aquests algorismes tinguin una bona precisió, potser millor que l'obtinguda per nosaltres, és necessària una gran quantitat d'informació provinent de l'escena. Per contra, analitzar estadísticament el gran volum de dades corresponent, fa que el temps emprat sigui molt llarg, i segons quin sigui l'objectiu perseguit pot ésser que això no sigui operatiu. En tot cas, a mesura que els ordinadors millorin les seves prestacions i es desenvolupin algorismes més eficients, es podrà aconseguir que ho sigui.



FIGURA 1. Imatge química d'un aliatge de Cu-Ag-S vista amb un microscopi electrònic.

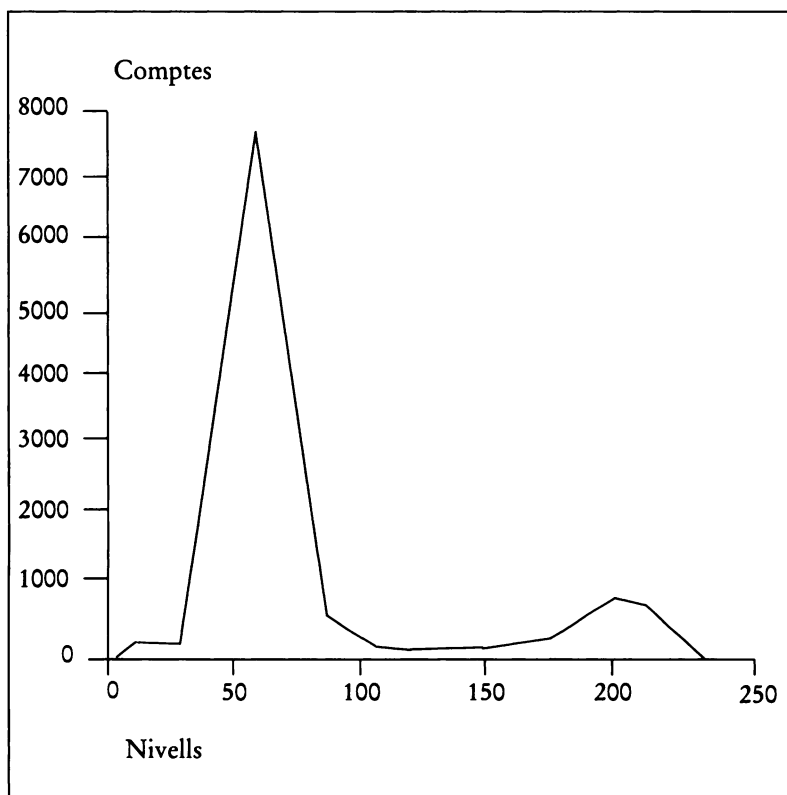


FIGURA 2. Histograma de la imatge de la figura 1.



FIGURA 3. Binarització de la imatge de la figura 1, amb els límits entre 80 i 255.

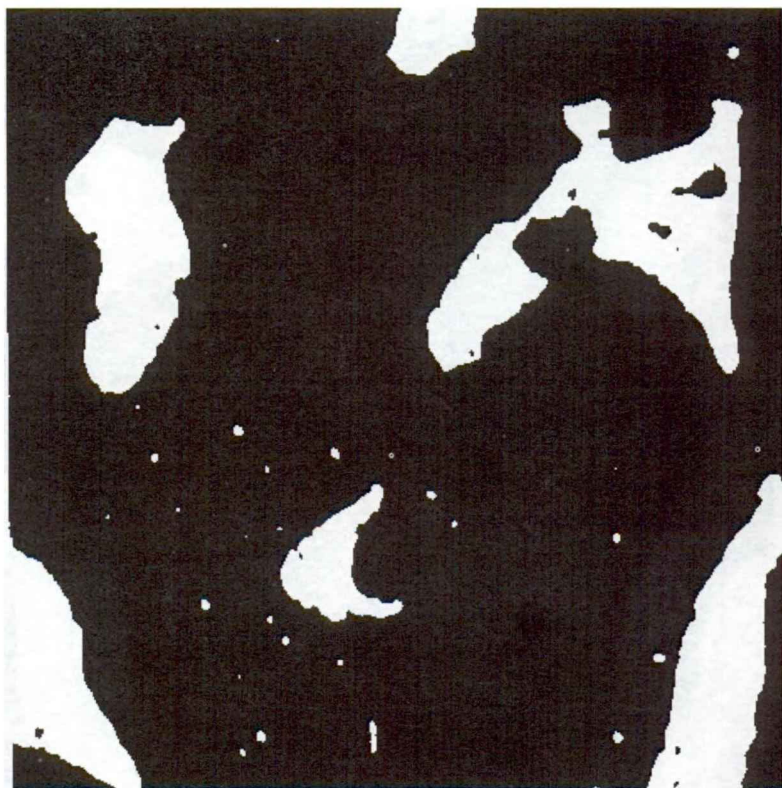


FIGURA 4. Binarització de la imatge de la figura 1, amb els límits entre 115 i 255.

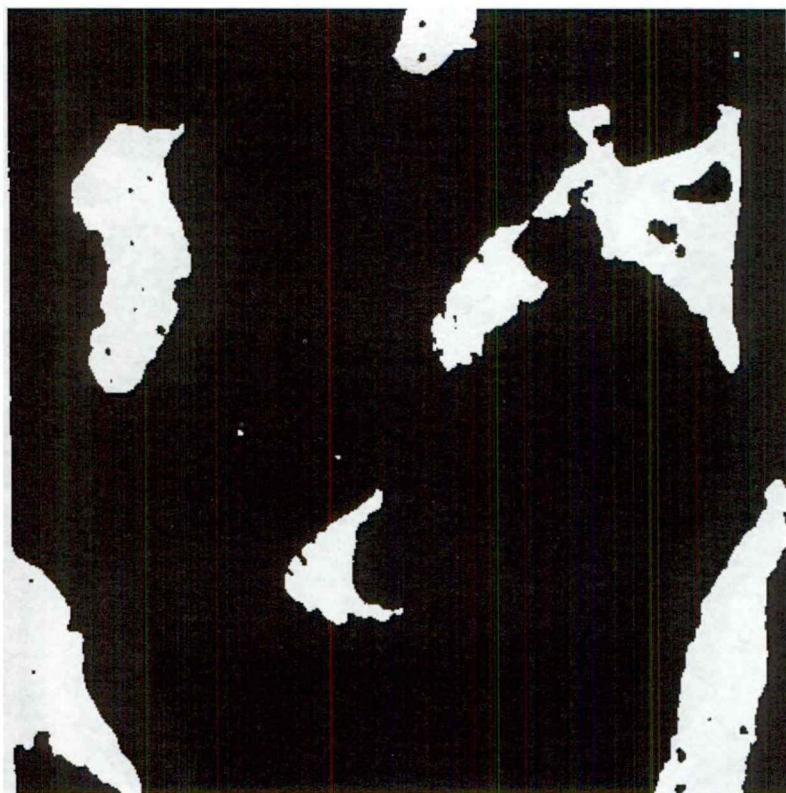


FIGURA 5. Binarització de la imatge de la figura 1, amb els límits entre 146 i 255.

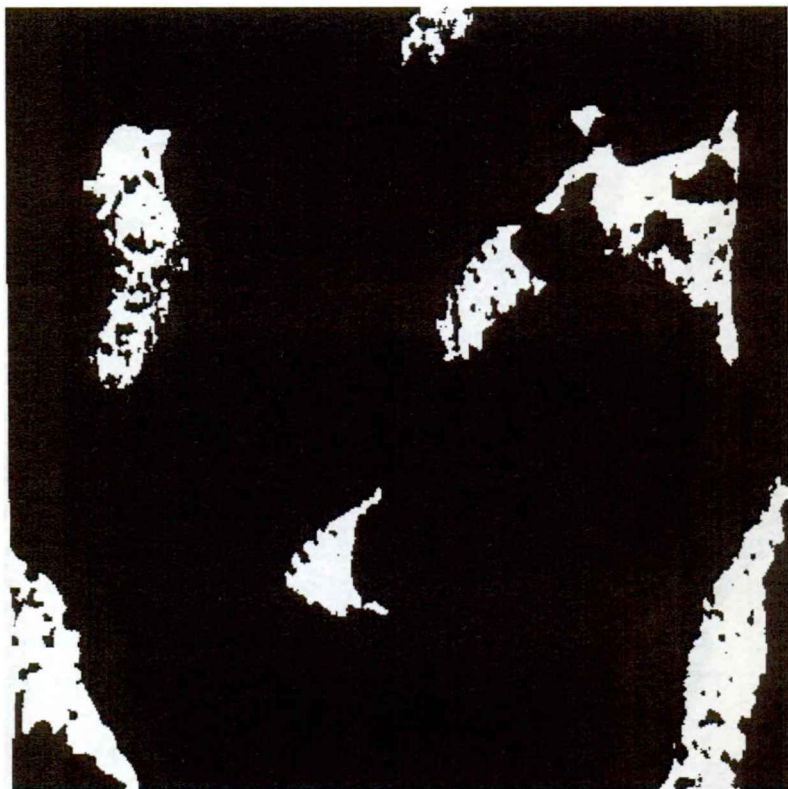


FIGURA 6. Binarització de la imatge de la figura 1, amb els límits entre 180 i 255.

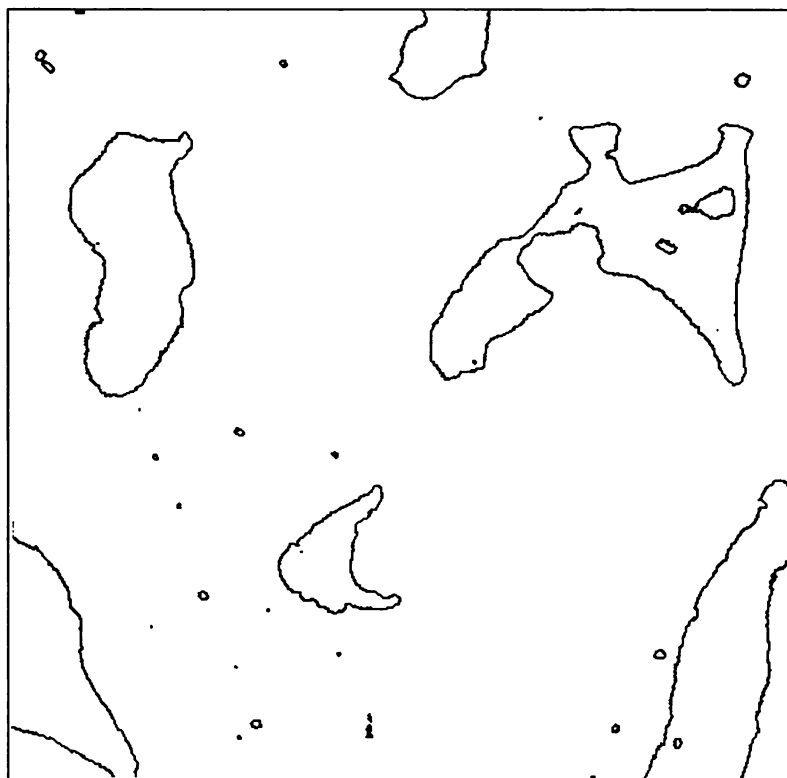


FIGURA 7. Objectes de la imatge de la figura 1 a partir de la binarització de la figura 4.

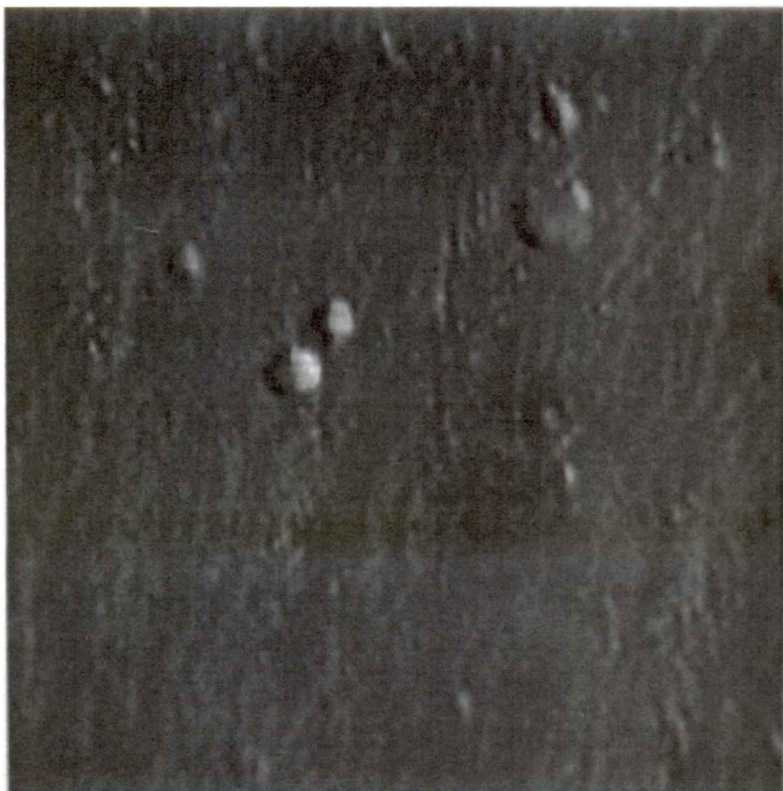


FIGURA 8. Imatge topogràfica d'un aliatge de Cu-Ag-S vista amb un microscopi electrònic.

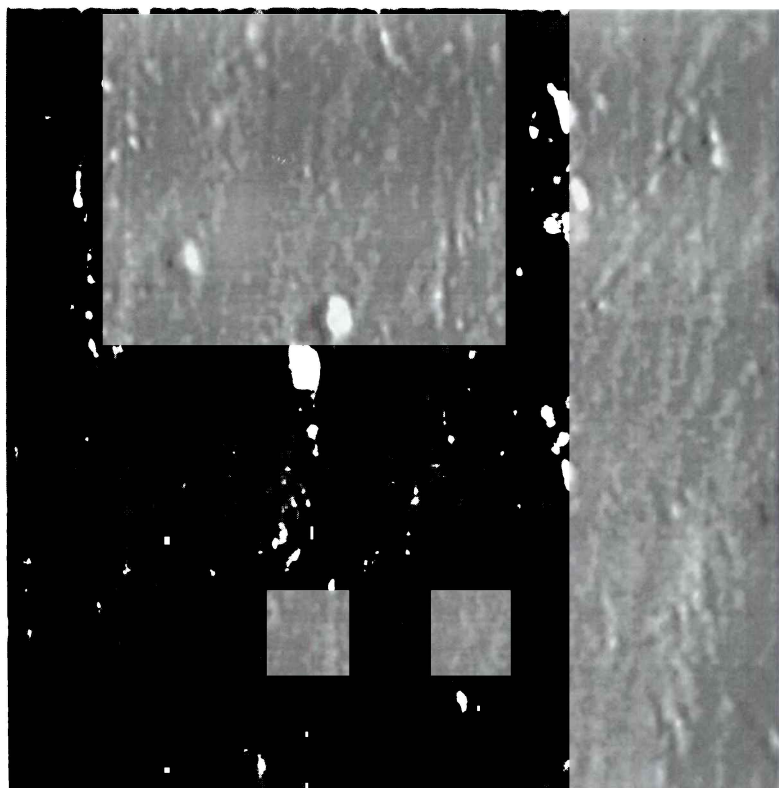


FIGURA 9. Imatge de fluctuacions corresponent a la de la figura 8.

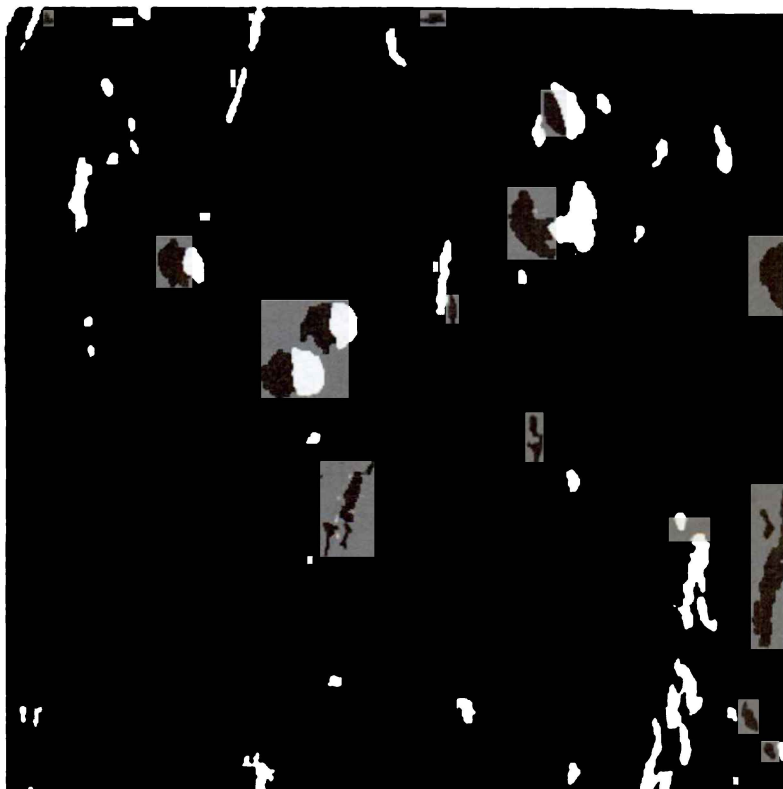


FIGURA 10. Ternarització de la imatge de la figura 9 amb un valor del llindar de llavor (fosc o clar) de 15σ i un valor del llindar de nucleació de $7,5 \sigma$.

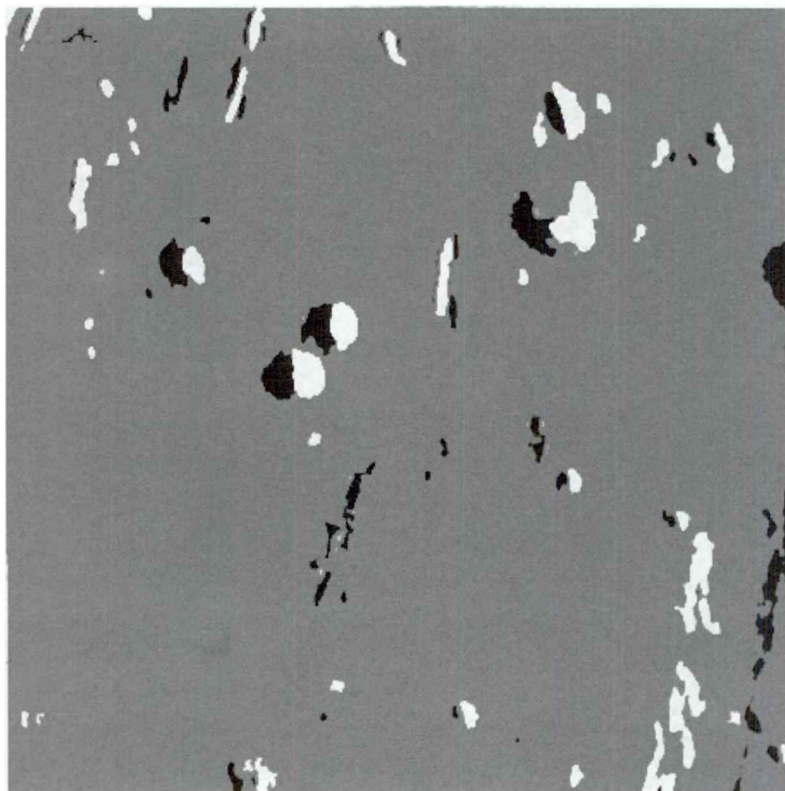


FIGURA 11. Ternarització de la imatge de la figura 9 amb un valor del llindar de llavor fosc de $12,5 \sigma$ i de llavor clar de 15σ , i un valor del llindar de nucleació de $7,5 \sigma$.

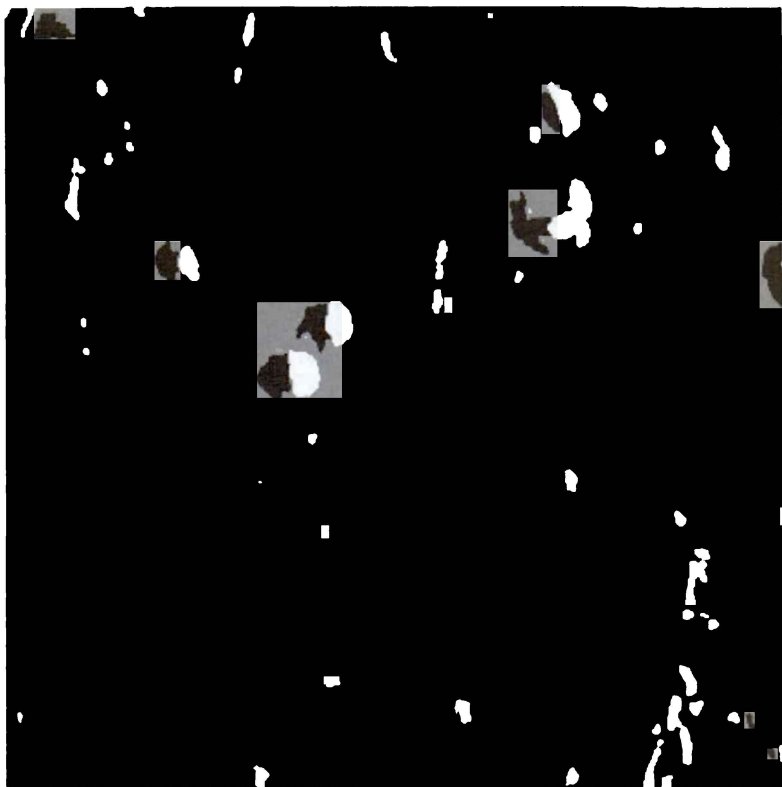


FIGURA 12. Ternarització de la imatge de la figura 9 amb un valor del llindar de llavor (fosc i clar) de 15σ i un valor del llindar de nucleació de 10σ .

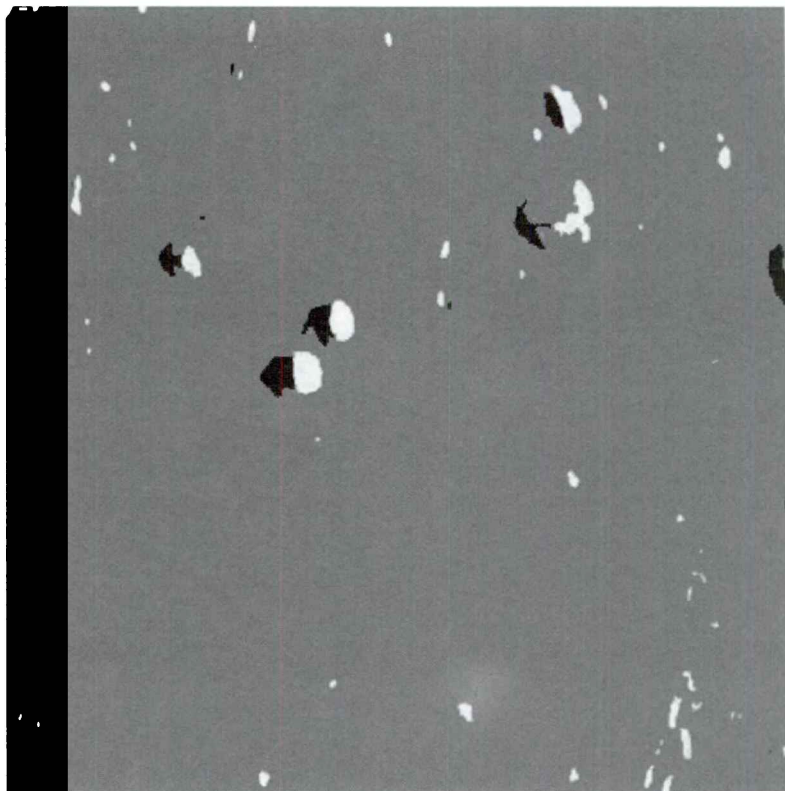


FIGURA 13. Ternarització de la imatge de la figura 9 amb un valor del llindar llavor (fosc i clar) de 15σ i un valor del llindar de nucleació de $12,5 \sigma$.

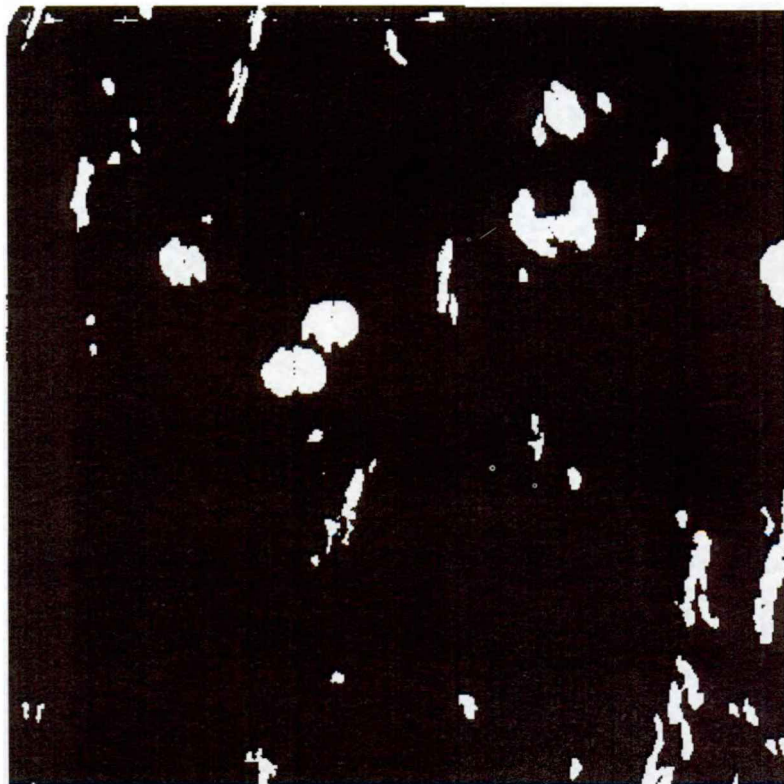


FIGURA 14. Binarització de la imatge de la figura 10.

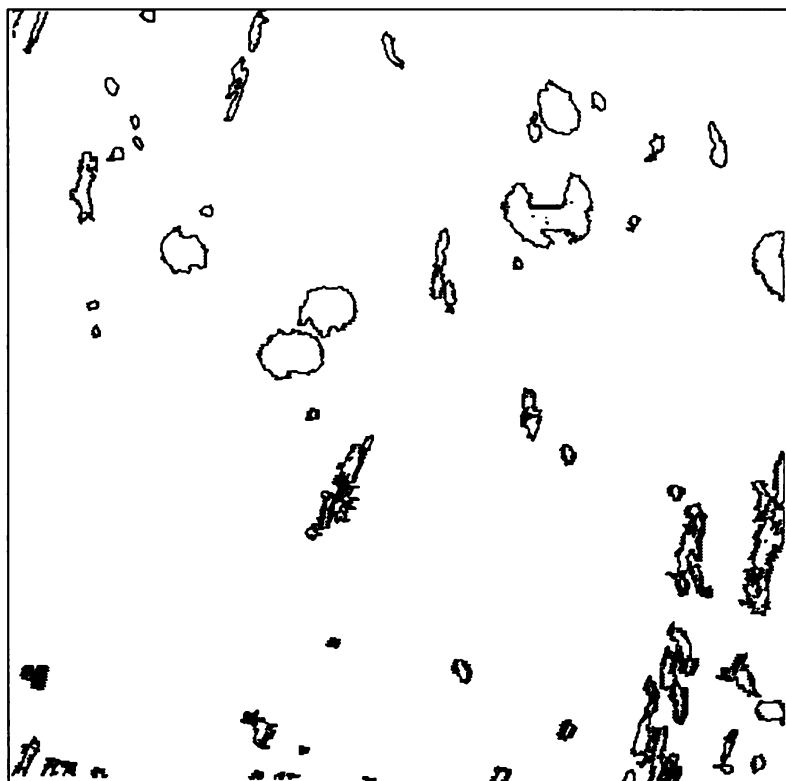


FIGURA 15. Vores dels objectes de la imatge de la figura 8 a partir de la binarització de la figura 14.