

# Interpretació de dibuixos lineals mitjançant tècniques d'isomorfisme entre grafs

Josep Lladós

Unitat de Processament d'Imatges i Intel·ligència Artificial  
Departament d'Informàtica - Edifici C - Universitat Autònoma de Barcelona.  
08193 Bellaterra (Barcelona)  
e-mail: josep@upisun1.uab.es

## Resum

L'anàlisi de documents té com a objectiu la interpretació automàtica de documents impresos sobre paper, amb la finalitat d'obtenir una descripció simbòlica d'aquests, que permeti el seu emmagatzemament i posterior tractament computacional. Les tècniques basades en grafs relacionals d'atributs permeten representar de manera compacta la informació continguda en dibuixos lineals i, mitjançant mecanismes d'isomorfisme entre grafs, reconèixer-hi certes estructures i, d'aquesta manera, interpretar el document. En aquest treball es dona una visió general de les tècniques de grafs aplicades al reconeixement visual d'objectes en problemes d'anàlisi de documents. Aquestes tècniques s'illustran amb un exemple de reconeixement de plànols dibuixats a mà alçada. Finalment es proposa la utilització de tècniques de Hough com a mecanisme per a accelerar el procés de reconeixement aplicant un cert coneixement sobre el domini en el que es treballa.

## 1 Introducció

Tradicionalment, l'home ha emmagatzemat i tramès la informació utilitzant documents sobre paper. L'aparició de la informàtica ha possibilitat la introducció d'un nou concepte de document: el *document electrònic*. Actualment, els sistemes de disseny assistit per ordinador (CAD) constitueixen una eina de gran ajut en la creació i modificació d'aquests documents. Però existeix el problema invers: convertir documents sobre paper a un format integrable dins d'un sistema CAD. El camp de l'*anàlisi de documents* treballa amb aquest objectiu mitjançant la utilització de tècniques de processament d'imatges i de reconeixement de formes aplicades sobre digitalitzacions de documents sobre paper. En aquest treball ens centrarem en el reconeixement d'elements gràfics del document, entenent per elements gràfics aquelles parts de la imatge formades per conjunts de línies i regions omplertes amb una certa textura. L'objectiu serà interpretar el dibuix i reconèixer les entitats d'alt nivell que l'integren, comptant amb el suport d'un cert coneixement sobre el domini en el que es treballa. Una vegada interpretat el document, serà possible editar-lo en un sistema CAD per a poder-lo modificar o integrar en altres documents.

Un sistema d'anàlisi i interpretació de documents gràfics es pot descomposar en tres nivells (figura 1). El *nivell lèxic* extreu les primitives de baix nivell que constitueixen el dibuix lineal (línies rectes, arcs de circumferència, punts d'intersecció, punts extrems i punts d'inflexió) i les seves propietats geomètriques i topològiques. El *nivell sintàctic* estableix les relacions estructurals entre les primitives de baix nivell i proporciona una representació simbòlica del document. L'objectiu del *nivell semàntic* és interpretar el document. Això voldrà dir reconèixer els elements constituents d'aquest amb l'ajut d'una base de coneixement de-

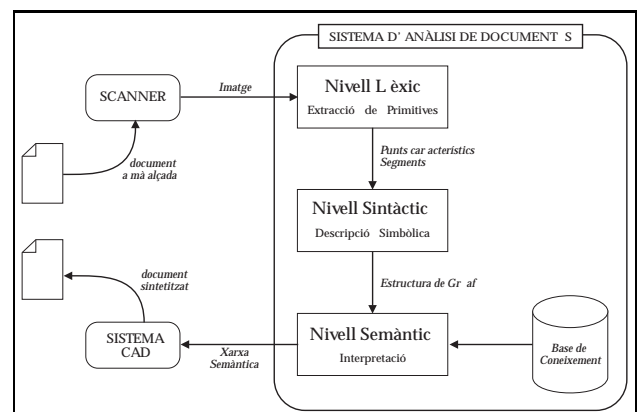


Figura 1: Nivells d'un sistema d'anàlisi de documents gràfics.

penent del domini al que pertanyi el document. Aquesta base de coneixement estarà formada per models de formes a reconèixer, restriccions geomètriques i topològiques que s'han de satisfer, etc.

En aquest treball ens centrem en el nivell semàntic que consisteix, com s'ha esmentat, en el reconeixement d'estructures complexes. Els *grafs relacionals d'atributs* [2] permeten representar de manera eficient i compacta aquestes estructures i converteixen el problema del reconeixement en un problema de correspondència entre grafs. Diferents problemes de reconeixement de dibuixos lineals han estat plantejats des d'aquest punt de vista [3][6][7] per a diferents àmbits tècnics com la cartografia, el disseny electrònic, mecànica, etc. La interpretació final del document s'emmagatzema en termes d'una xarxa semàntica de la qual els grafs esmentats en formen part.

El present treball està organitzat de la següent manera: a la secció 2 es descriuen els mètodes de correspondència

entre grafs, a la secció 3 es proposen mètodes basats en tècniques de Hough per a accelerar el procés de reconeixement. La secció 4 ilustra els mètodes exposats amb una aplicació dins de l'àmbit de l'arquitectura. Finalment, a la secció 5 es presenten les conclusions i les línies de continuació dins del treball de tesi en el que s'emmarquen els processos descrits.

## 2 Representació de dibuixos lineals amb estructures de grafs

### 2.1 Definicions i notació

Un *graf relacional d'atributs*  $G$  és una 4-pla  $(V, E, L_V, L_E)$  on  $V$  és un conjunt finit d'elements, anomenats nodes;  $E$  és un conjunt finit de parells d'elements de  $V$ , anomenats arcs;  $L_V$  i  $L_E$  són dues funcions, anomenades funcions d'etiquetatge, definides com  $L_V : V \rightarrow \Sigma \times A^k$  i  $L_E : E \rightarrow \Sigma \times A^k$ , on  $\Sigma$  és un conjunt finit d'etiquetes simbòliques i  $A$  és un conjunt finit de valors d'atributs, és a dir, a cada node (arc) del graf se li assigna una etiqueta i un conjunt de valors d'atributs que el caracteritzen. Un graf  $G' = (V', E', L'_V, L'_E)$  és un *subgraf* de  $G$  (denotat per  $G' \subseteq G$ ) si  $V' \subseteq V$  i  $E' \subseteq E$ . Denotarem per  $|V|$  el número de nodes de  $V$  i per  $|E|$  el número d'arcs d' $E$ . Utilitzarem, per simplicitat,  $G = (V, E)$  en lloc de  $G = (V, E, L_V, L_E)$  sempre i quan no hi hagi confusió.

Una funció  $f : V \rightarrow V'$  és un *isomorfisme entre grafs* de  $G = (V, E, L_V, L_E)$  a  $G' = (V', E', L'_V, L'_E)$  si  $f$  és bijectiva i es compleix:

1.  $L_V(v) = L'_V(f(v))$  per tot  $v \in V$ .
2.  $L_E(e) = L'_E(e')$  per qualsevol arc  $e = (v_i, v_j) \in E$  i  $e' = (f(v_i), f(v_j)) \in E'$ .
3. Per tot arc  $e' = (v'_i, v'_j) \in E'$  existeix un arc  $e = (f^{-1}(v'_i), f^{-1}(v'_j)) \in E$  tal que  $L_E(e) = L'_E(e')$ .

### 2.2 Isomorfisme entre subgrafs com a procés de reconeixement

Un graf d'atributs permet emmagatzemar de manera eficient i compacta un dibuix lineal, fent correspondre els punts característics del document amb els nodes del graf i les línies del document amb els arcs. D'aquesta manera, reconèixer estructures complexes dins del document consistirà en aplicar un isomorfisme entre subgrafs, concretament, buscar el graf que representa el model a reconèixer com a subgraf del graf que representa el document d'entrada. A aquests dos grafs els anomenarem respectivament *graf model* ( $G_M$ ) i *graf d'entrada* ( $G_I$ ).

Hi ha dos inconvenients que cal assenyalar per a l'isomorfisme entre subgrafs. El primer és la seva complexitat computacional (pertany a la classe de problemes NP-complets). El segon inconvenient és que moltes vegades es treballa amb informació que conté soroll, la qual cosa provoca que una correspondència exacta entre grafs no sigui viable i s'hagi de recórrer a tècniques de correspondència inexacta. Aquest segon problema és especialment sensible en interpretació de documents que han estat fets a mà alçada com en l'exemple que es mostra a la secció 4.

### 2.2.1 Correspondència exacta entre grafs

La correspondència exacta entre grafs (*EGM*) pot ser resolta mitjançant algorismes clàssics de cerca amb *backtracking* dins d'un arbre. L'arbre de cerca comença amb el primer node  $v_M^1 \in V_M$ . Aquest node pot correspondre potencialment a qualsevol node  $v_I^j \in V_I$ . Cadascuna d'aquestes assignacions potencials és un node de nivell 1 de l'arbre. L'algorisme tria un d'aquests nodes i fa l'assignació  $f(v_M^1) = v_I^j$ . A continuació es selecciona el següent node  $v_M^2 \in V_M$  i es construeixen les hipòtesis de segon nivell. En aquest nivell, ja no es té en compte aquelles assignacions que violen les condicions necessàries per a l'isomorfisme. El procés continua fins al nivell  $|V|$  de l'arbre. Un camí des de l'arrel fins a una fulla exitosa de l'arbre defineix un isomorfisme.

El cost exponencial del mètode exposat fa necessària la utilització de criteris heurístics d'esporgat. Una de les tècniques més conegudes per a fer aquest esporgat és la *relaxació discreta* que redueix el problema de l'isomorfisme entre subgrafs al problema de l'*etiquetatge consistent* [4]. El problema d'etiquetatge es pot enunciar de la següent manera:  $O$  és un conjunt d'objectes a identificar,  $L$  és un conjunt d'etiquetes que representen hipòtesis sobre la identificació dels objectes i  $R$  és un conjunt de restriccions entre els parells objecte-etiqueta. L'objectiu és el d'obtenir un conjunt  $H$  d'hipòtesis que assignin una etiqueta a cada objecte satisfent les restriccions. En el cas de l'isomorfisme entre grafs, els objectes a etiquetar són els nodes model, les etiquetes es construeixen a partir dels nodes del graf d'entrada i el graf model es pot entendre com un graf de restriccions. Aquestes restriccions estan basades en els arcs del graf i representen condicions geomètriques i topològiques que han de complir els nodes en connectar-se. Des d'aquest punt de vista, a [4] el problema d'etiquetatge consistent es descomposa en tres passos:

- *Consistència de nodes*: cada node model s'etiqueta amb tots els nodes del graf d'entrada que poden correspondre-li. L'etiquetatge es basa en criteris topològics (número d'arcs que hi conflueixen) i criteris geomètrics (configuració del node, és a dir, disposició dels arcs que hi conflueixen).
- *Consistència d'arcs*: Es verifica la consistència entre els etiquetatges de dos nodes veïns i s'eliminen tots aquells que són inconsistents. L'algorisme *AC4* [9], basat en tècniques de relaxació discreta, és un dels més utilitzats per aquest propòsit.
- *Consistència de camins*: es busca un conjunt d'etiquetes globalment vàlides per als nodes model. Aquest conjunt d'etiquetes es consideren com la solució global. Aquest procés es pot fer mitjançant un algorisme de *backtracking with forward checking* que consisteix en tenir en compte, per a cada node de l'arbre de cerca, els parells node-etiqueta futurs que són inconsistents amb l'actual i, per tant, no obrir branques de l'arbre que portin cap a aquest etiquetatge inconsistent.

## 2.2.2 Correspondència inexacta entre grafs

Hi ha casos en els quals pot haver-hi distorsions en els grafs (manca de nodes o arcs, divisió d'aquests, etc.). Aquestes distorsions poden ser degudes a diferents factors: deficient adquisició de la imatge, dibuix a mà alçada, poca robustesa del procés de vectorització, obtenció del graf model mitjançant una generalització de diferents instàncies (aprenentatge inductiu), etc. Això vol dir que determinats nodes o arcs del model poden no tenir correspondència sobre el graf d'entrada o que els seus atributs poden aparèixer sensiblement diferents. Tanmateix, aquesta distorsió no ha de ser inconvenient per a que el model no sigui reconegut. Això farà necessària la utilització de tècniques de correspondència inexacta entre grafs (IGM). El problema d'IGM es pot entendre com el de trobar el subgraf  $G_S \subseteq G_I$  que més s'assembla al model  $G_M$  en termes d'una distància entre ells inferior a un cert valor  $d$ . La *distància d'edició* per a *strings* [1] pot ser generalitzada i utilitzada com a distància entre grafs. Per a calcular aquesta distància es proposen les següents *operacions d'edició* que serveixen per a transformar un graf  $G = (V, E)$  en un altre:

1. *Rectiquetatge d'un node o un arc.* Consisteix en canviar la seva etiqueta  $l$  i el valor dels seus atributs  $a_1, \dots, a_k$ . Es denota per  $((v, l, a_1, \dots, a_k) \rightarrow ((v, l', a'_1, \dots, a'_k)))$ .
2. *Esborrat d'un node o un arc.* Es denota per  $v \rightarrow \Lambda$ . Cal tenir en compte que l'esborrat d'un node implica l'esborrat de tots els arcs que hi conflueixen.
3. *Inserció d'un node o un arc.* Denotat per  $\Lambda \rightarrow v$ . Només es pot inserir un arc si existeixen els vertexs adjacents.

Es defineix una funció de cost  $c : (a \rightarrow b) \rightarrow \mathbf{R}^+$  com la funció que assigna un cost a cadascuna de les operacions d'edició esmentades. El cost d'una seqüència  $S = s_1, \dots, s_n$  d'operacions d'edició serà per tant  $c(S) = \sum_{i=1}^n c(s_i)$ . La distància entre dos grafs  $G$  i  $G'$  es defineix com  $dist(G, G') = \min\{c(S), S \text{ és una seqüència d'operacions d'edició que transforma } G \text{ en } G'\}$ . Aquesta distància pot ser computada mitjançant tècniques de *programació dinàmica*.

Com es veurà en l'exemple presentat més endavant, les tècniques de relaxació també permeten dur a terme un IGM sempre i quan es tinguin en compte les possibles distorsions del graf dins del conjunt de restriccions  $R$ .

## 3 Acceleració del reconeixement utilitzant tècniques de Hough

Sovint, els dibuixos lineals que ens ocupen contenen, a més dels símbols a reconèixer, certes regions omplertes amb una textura estructurada. En aquesta secció ens centrem en el cas més simple de textura que és el constituït per línies rectes paral·leles i equiespaiades. En aquest cas, detectar i eliminar les línies que formen part d'aquest omplert comporta l'eliminació del graf d'entrada de diversos nodes i arcs. El cost de la correspondència entre grafs és, en el cas pitjor,  $O(|V_I|^{|\mathcal{M}|})$ . Si, partint del coneixement que es

té sobre el tipus de documents a tractar, es pot eliminar una sèrie de nodes i arcs que no pertanyen a cap model a reconèixer degut a que formen part de les regions esmentades, podrem accelerar el reconeixement que, com hem vist, és inherentment exponencial. La detecció de línies paral·leles dins d'un dibuix lineal es pot fer utilitzant la transformació de Hough exposada a continuació.

## 3.1 Transformació de Hough bàsica

La transformació bàsica de Hough per a detecció de línies rectes es basa en el següent: donat un punt de l'espai euclidi  $(x, y)$ , l'equació  $\rho = x \cos \theta + y \sin \theta$  representa el conjunt de rectes que passen per ell. El mètode fa un canvi de coordenades calculant, per cada punt de la imatge original, el conjunt de rectes que passen per ell, representant-les a l'espai  $(\theta, \rho)$ . Així, un punt de l'espai  $(\theta, \rho)$  representa una recta a l'espai  $(x, y)$ . El procés consisteix en recórrer tots els punts de l'espai  $(x, y)$  (imatge inicial) i, per cadascun d'ells, incrementar els respectius  $(\theta, \rho)$  corresponents a les rectes que passen per aquest punt. Al final del procés, els punts  $(\theta, \rho)$  "més votats", corresponen a les rectes a les quals pertanyen més punts, per tant, fent la transformació inversa dels màxims de l'espai  $(\theta, \rho)$  obtenim les rectes a l'espai  $(x, y)$ . Partint d'aquesta idea bàsica, s'ha proposat diferents extensions que tenen per objectiu detectar objectes d'una imatge que poden ser descrits donant valors als paràmetres d'una equació (veure recull a [5]).

## 3.2 Transformació de Hough per a la detecció de línies paral·leles

La detecció de línies paral·leles dins d'un dibuix lineal vectoritzat mitjançant un graf d'atributs es redueix a la detecció d'arcs paral·lels del graf. Suposem que, per a qualsevol arc  $e$  que aproxima una línia recta, la funció d'etiquetatge li fa correspondre com a atributs les coordenades polars d'aquesta recta  $(L_E(e) = (recta, \theta_e, \rho_e))$ . En aquest cas, cada arc del graf acumula un vot a l'espai  $(\theta, \rho)$ . Com es pot observar a la figura 2, els arcs paral·lels del graf generen un conjunt de punts aliniats verticalment i equiespaiats a l'espai  $(\theta, \rho)$ . La distància  $(\Delta\rho)$  entre aquests punts correspon a la distància entre els arcs paral·lels i la coordenada  $\theta$  sobre la que estan aliniats correspon a la inclinació d'aquests arcs. Per tant, detectar regions formades per arcs paral·lels dins d'un graf es redueix a buscar, dins l'espai  $(\theta, \rho)$ , configuracions com la que s'acaba de descriure.

## 4 Exemple d'aplicació

A continuació es mostren uns resultats d'aplicació de les tècniques esmentades per a la interpretació de plànols d'arquitectura fets a mà alçada com el de la figura 4(a). El sistema, descrit més detalladament a [8], té com a objectiu la interpretació dels documents esmentats com a eina alternativa d'entrada a un sistema CAD. D'aquesta manera, l'arquitecte podria introduir de manera senzilla el document dibuixant el plànol sobre paper i, posteriorment, fer-ne els retocs finals sobre el CAD una vegada hagués estat interpretat.

Interpretar aquests plànols implica reconèixer certs e-

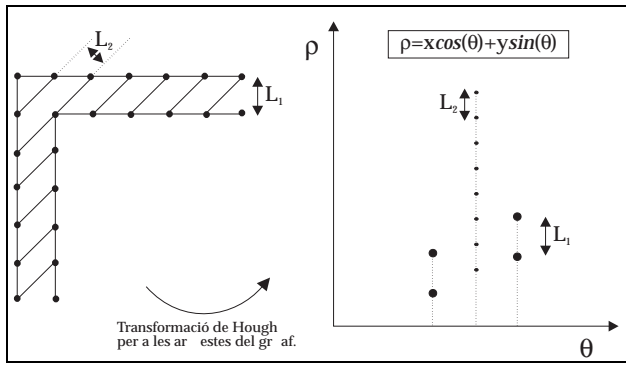


Figura 2: Transformació de Hough a partir del graf d'entrada.

lements com portes, taules, parets, etc. Per això, es vectoritza el dibuix mitjançant un graf d'atributs (figura 4(b)) i s'aplica un isomorfisme entre subgrafs respecte els models a reconèixer. Per a fer aquest isomorfisme s'ha utilitzat l'algorisme *AC4* abans comentat. Donada la introducció manual del document, hi ha certs errors d'imprecisió que cal tenir en compte. Les restriccions imposades per al procés de relaxació han de ser prou flexibles per a reconèixer els models malgrat aquests errors, és a dir, s'ha dut a terme un *IGM*. Així, per exemple, a un arc del graf model se li pot assignar un conjunt d'arcs del graf d'entrada que aproximem la seva forma. Això es pot observar a la figura 4(c) on es mostren els resultats de l'isomorfisme entre subgrafs respecte als models de la figura 3(a)(b).

En aquest tipus de plànols s'hi pot aplicar el criteri d'esporgat descrit abans per a eliminar arcs del graf que són paral·lels entre si i que pertanyen a la mateixa regió. Concretament, com es pot observar, les regions corresponents a parets presenten aquestes característiques. Si s'aplica el procés basat en la transformació de Hough que s'ha descrit es poden detectar les parets i eliminar del graf d'entrada els arcs que les formen. Això incrementa notablement el rendiment del procés d'isomorfisme doncs hi ha menys nodes i arcs a comparar. El resultat final consistent en la generació del document a partir de la informació obtinguda després del procés d'interpretació es mostra a la figura 4(d).

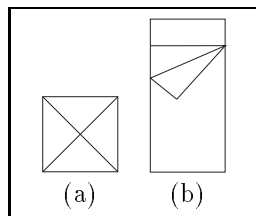


Figura 3: Alguns grafs model a reconèixer.

## 5 Conclusions

Donada la semblança entre un dibuix lineal i una estructura de graf, les tècniques existents per a buscar un isomorfisme entre grafs són de gran utilitat per a interpretar aquest tipus de documents. Aquesta interpretació es pot entendre com un reconeixement d'estructures comple-

xes i de la seva relació estructural. D'aquesta manera, es pot donar una representació simbòlica del document mitjançant, per exemple, una xarxa semàntica. En aquest treball s'ha descrit les tècniques més representatives per a realitzar un isomorfisme entre grafs (exacte o inexacte) i s'ha proposat criteris heurístics basats en les característiques que ofereixen els dibuixos lineals utilitzats en dominis d'enginyeria. Finalment, es mostren alguns resultats en els que s'ha aplicat les tècniques descrites que resumeixen el treball de tesi realitzat fins al moment.

Actualment, el treball de tesi en el que es basa aquest treball està enfocat vers l'aprenentatge inductiu de models. Concretament, aprendre una estructura de graf a partir de diferents instàncies fetes a mà alçada del dibuix lineal que es vol aprendre. Donat que un aprenentatge d'aquest tipus suposa assignar un grau d'incertesa als nodes i arcs obtinguts, s'ha escollit una estructura de *graf aleatori* [1]. Els grafs aleatoris són grafs en els que els atributs dels seus nodes i arcs són variables aleatòries de manera que un graf d'atributs pot ser considerat com una realització del graf aleatori.

## Referències

- [1] H. Bunke and A. Sanfeliu. *Syntactic and Structural Pattern Recognition. Theory and Applications*. World Scientific Publishing Company, 1990.
- [2] Shi-Kuo Chang. *Principles of pictorial information systems design*. Prentice-Hall, 1990.
- [3] A.H. Habacha. Structural recognition of disturbed symbols using discrete relaxation. In *1st. ICDAR*, pages 170–178, September–October 1991. Saint Malo, France.
- [4] Thomas C. Henderson. *Discrete Relaxation Techniques*. Oxford University Press, 1990.
- [5] J. Illingworth and J. Kitler. A survey of the hough transform. *Computer Vision, Graphics and Image Processing*, (44):87–116, 1988.
- [6] P. Kuner and B. Ueberreiter. Pattern recognition by graph matching. combinatorial versus continuous optimization. *International Journal of Pattern Recognition and Artificial Intelligence*, 2(3):527–542, September 1988.
- [7] S.W. Lee, J.H. Kim, and F.C.A. Groen. Translation-, rotation-, and scale-invariant recognition of hand-drawn symbols in schematic diagrams. *International Journal of Pattern Recognition and Artificial Intelligence*, 4(1):1–25, 1990.
- [8] Josep Lladós and Enric Martí. Structural recognition of hand drawn floor plans. In *VI Spanish Symposium on Pattern Recognition and Image Analysis*, pages 27–34, April 1995. Cordoba.
- [9] R. Mohr and T.C. Henderson. Arc and path consistency revisited. *Artificial Intelligence*, (28):225–233, 1986.

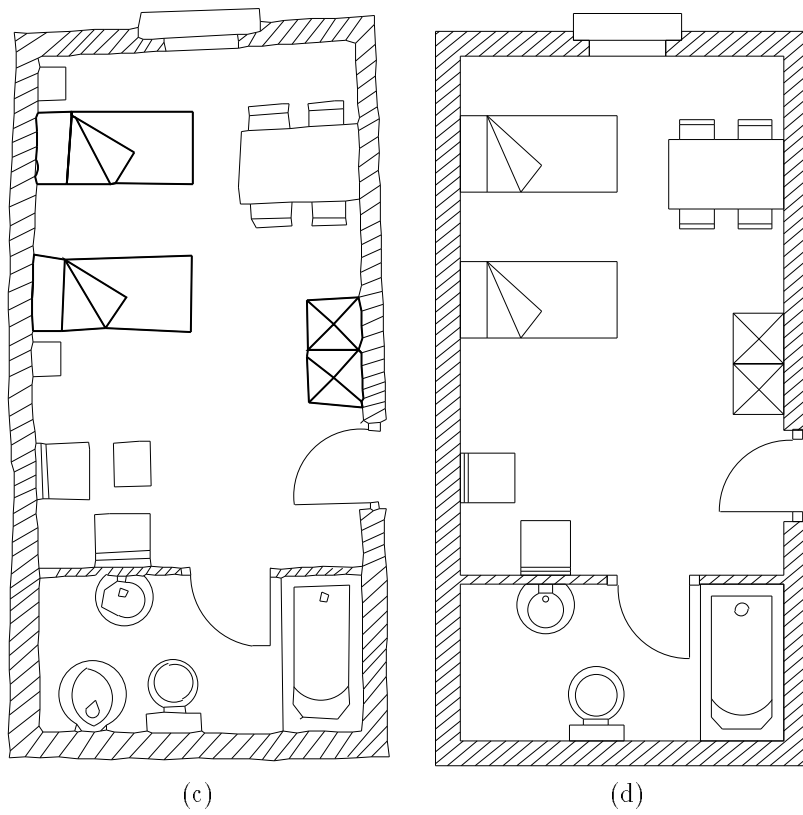
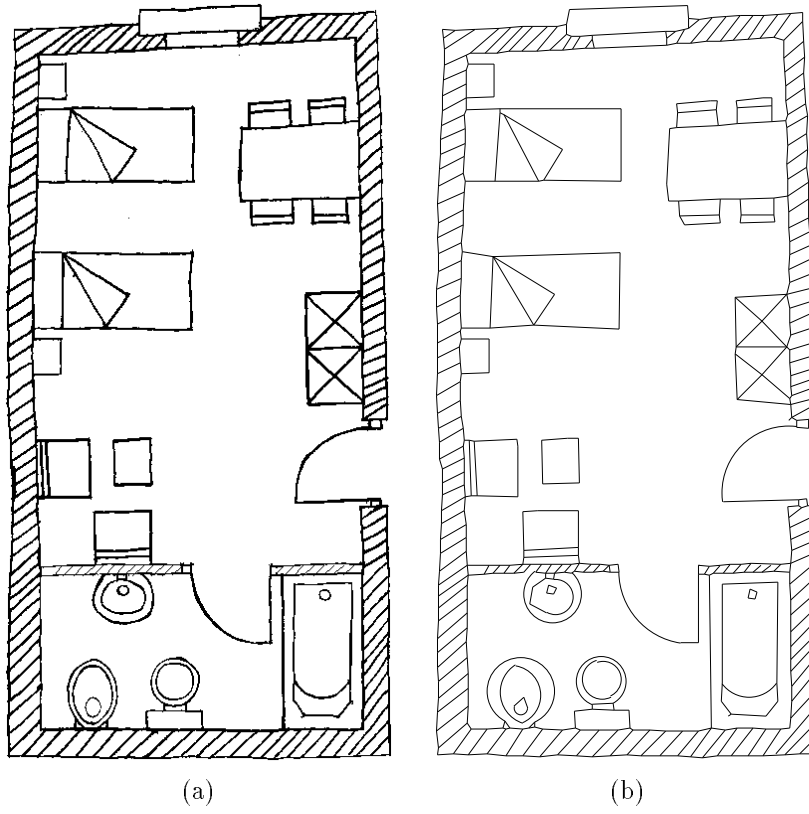


Figura 4: Imatge (a), grafs (b) resultat de l'isomorfisme (c), document sintetitzat després de ser interpretat (d).