

THESE

présentée

DEVANT L'INSTITUT NATIONAL DES SCIENCES APPLIQUEES DE LYON

pour obtenir

LE GRADE DE DOCTEUR

FORMATION DOCTORALE : DEA d'Informatique de Lyon

SPECIALITE : Informatique

PAR

Jean-Yves RAMEL

***INTERPRETATION AUTOMATIQUE DE DESSINS
METHODES D'ANALYSE ET DE RECONNAISSANCE
APPLICATION AUX PLANS CINEMATIQUES***

Soutenue le 26 novembre 1996 devant la Commission d'Examen

Jury MM.

Jean-Pierre Cocquerez

Président, Rapporteur

Guy Lorette

Rapporteur

Jean-Marc Brun

Hubert Emptoz

Karl Tombre

Nicole Vincent

Directeur de thèse

Janvier 1998

Institut National des Sciences Appliquées de Lyon

Directeur : J. Rochat

Professeurs

S.	AUDISIO	PHYSICOCHIMIE INDUSTRIELLE
J.C.	BABOUX	GEMPPM*
B.	BALLAND	PHYSIQUE DE LA MATIERE
D.	BARBIER	PHYSIQUE DE LA MATIERE
G.	BAYADA	MODELISATION MATHEMATIQUE ET CALCUL SCIENTIFIQUE
C.	BERGER (Mlle)	PHYSIQUE DE LA MATIERE
M.	BETEMPS	AUTOMATIQUE INDUSTRIELLE
J.M.	BLANCHARD	LAEPSI**
C.	BOISSON	VIBRATIONS ACOUSTIQUES
M.	BOIVIN	MECANIQUE DES SOLIDES
H.	BOTTA	EQUIPE DEVELOPPEMENT URBAIN
G.	BOULAYE	INFORMATIQUE
J.	BRAU	CENTRE DE THERMIQUE
M.	BRISAUD	GENIE ELECTRIQUE ET FERROELECTRICITE
M.	BRUNET	MECANIQUE DES SOLIDES
J.C.	BUREAU	THERMODYNAMIQUE APPLIQUEE
J.Y.	CAVAILLE	GEMPPM*
J.P.	CHANTE	COMPOSANTS DE PUISSANCE ET APPLICATIONS
B.	CHOCAT	UNITE DE RECHERCHE EN GENIE CIVIL
B.	CLAUDEL	LAEPSI**
M.	COUSIN	UNITE DE RECHERCHE EN GENIE CIVIL
M.	DIOT	THERMODYNAMIQUE APPLIQUEE
A.	DOUTHEAU	CHIMIE ORGANIQUE
R.	DUFOUR	MECANIQUE DES STRUCTURES
J.C.	DUPUY	PHYSIQUE DE LA MATIERE
H.	EMPTOZ	RECONNAISSANCE DES FORMES ET VISION
C.	ESNOUF	GEMPPM*
L.	EYRAUD (Prof. Emérite)	GENIE ELECTRIQUE ET FERROELECTRICITE
G.	FANTOZZI	GEMPPM*
M.	FAYET	MECANIQUE DES SOLIDES
J.	FAVREL	GROUPE DE RECHERCHE EN PRODUCTIQUE ET INFORMATIQUE DES SYSTEMES MANUFACTURIERS
G.	FERRARIS-BESSO	MECANIQUE DES STRUCTURES
Y.	FETIVEAU	GENIE ELECTRIQUE ET FERROELECTRICITE
L.	FLAMAND	MECANIQUE DES CONTACTS
P.	FLEISCHMANN	GEMPPM*
A.	FLORY	INGENIERIE DES SYSTEMES D'INFORMATION
R.	FOUGERES	GEMPPM*
F.	FOUQUET	GEMPPM*
L.	FRECON	INFORMATIQUE
R.	GAUTHIER	PHYSIQUE DE LA MATIERE
M.	GERY	CENTRE DE THERMIQUE
G.	GIENEZ	CREATIS***
P.	GOBIN (Prof. émérite)	GEMPPM*
P.	GONNARD	GENIE ELECTRIQUE ET FERROELECTRICITE
M.	CONTRAND	COMPOSANTS DE PUISSANCE ET APPLICATIONS
R.	GOUTTE (Prof. Emérite)	CREATIS***
G.	GRANGE	GENIE ELECTRIQUE ET FERROELECTRICITE
G.	GUENIN	GEMPPM*
M.	GUICHARDANT	BIOCHIMIE ET PARMACOLOGIE
G.	GUILLOT	PHYSIQUE DE LA MATIERE
A.	GUNET	GROUPE DE RECHERCHE EN PRODUCTIQUE ET INFORMATIQUE DES SYSTEMES MANUFACTURIERS
J.L.	GUYADER	VIBRATIONS ACOUSTIQUES
J.P.	GUYOMAR	GENIE ELECTRIQUE ET FERROELECTRICITE
J.M.	JOLION	RECONNAISSANCE DES FORMES ET VISION
J.F.	JULLIEN	UNITE DE RECHERCHE EN GENIE CIVIL
A.	JUTARD	AUTOMATIQUE INDUSTRIELLE
R.	KASTNER	UNITE DE RECHERCHE EN GENIE CIVIL
H.	KLEIMANN	GENIE ELECTRIQUE ET FERROELECTRICITE
J.	KOULOUMDJIAN	INGENIERIE DES SYSTEMES D'INFORMATION
M.	LAGARDE	BIOCHIMIE ET PARMACOLOGIE
M.	LALANNE	MECANIQUE DES STRUCTURES

A.	LUBRECHT	MECANIQUE DES CONTACTS
Y.	MARTINEZ	INGENIERIE DES SYSTEMES D'INFORMATION
H.	MAZILLE	PHYSICOCHIMIE INDUSTRIELLE
P.	MERLE	GEMPPM*
J.	MERLIN	GEMPPM*
J.P.	MILLET	PHYSICOCHIMIE INDUSTRIELLE
M.	MIRAMOND	UNITE DE RECHERCHE EN GENIE CIVIL
N.	MONGEREAU (Prof. Emérite)	UNITE DE RECHERCHE EN GENIE CIVIL
R.	MOREL	MECANIQUE DES FLUIDES
P.	MOSZKOWICZ	LAEPSI**
P.	NARDON	BIOLOGIE APPLIQUEE
A.	NAVARRO	LAEPSI**
A.	NOURI (Mme)	MODELISATION MATHEMATIQUE ET CALCUL SCIENTIFIQUE
M.	OTTERBEIN	LAEPSI**
J.P.	PASCAULT	MATERIAUX MACROMOLECULAIRES
G.	PAVIC	VIBRATIONS ACOUSTIQUES
J.	PERA	UNITE DE RECHERCHE EN GENIE CIVIL
G.	PERRACHON	THERMODYNAMIQUE APPLIQUEE
J.	PEREZ (Prof. Emérite)	GEMPPM*
P.	PINARD	PHYSIQUE DE LA MATIERE
J.M.	PINON	INGENIERIE DES SYSTEMES D'INFORMATION
D.	PLAY	CONCEPTION ET ANALYSE DES SYSTEMES MECANQUES
J.	POUSIN	MODELISATION MATHEMATIQUE ET CALCUL SCIENTIFIQUE
P.	PREVOT	GROUPE DE RECHERCHE EN APPRENTISSAGE, COOPERATION ET INTERFACES MULTIMODALES
R.	PROST	CREATIS***
M.	RAYNAUD	CENTRE DE THERMIQUE
J.M.	REYNOUARD	UNITE DE RECHERCHE EN GENIE CIVIL
E.	RIEUTORD (Prof. Emérite)	MECANIQUE DES FLUIDES
J.	ROBERT-BAUDOUY (Mme)	GENETIQUE MOLECULAIRE DES MICROORGANISMES
D.	ROUBY	GEMPPM*
P.	RUBEL	INGENIERIE DES SYSTEMES D'INFORMATION
C.	RUMELHART	MECANIQUE DES SOLIDES
J.F.	SACADURA	CENTRE DE THERMIQUE
H.	SAUTEREAU	MATERIAUX MACROMOLECULAIRES
S.	SCARVARDA	AUTOMATIQUE INDUSTRIELLE
D.	THOMASSET	AUTOMATIQUE INDUSTRIELLE
M.	TROCCAZ	GENIE ELECTRIQUE ET FERROELECTRICITE
R.	UNTERREINER	CREATIS***
J.	VERON	LAEPSI**
G.	VIGIER	GEMPPM*
A.	VINCENT	GEMPPM*
P.	VUILLERMOZ	PHYSIQUE DE LA MATIERE
Directeurs de recherche C.N.R.S.		
Y.	BERTHIER	MECANIQUE DES CONTACTS
P.	CLAUDY	THERMODYNAMIQUE APPLIQUEE
N.	COTTE-PATTAT (Mme)	GENETIQUE MOLECULAIRE DES MICROORGANISMES
P.	FRANCIOSI	GEMPPM
J.F.	GERARD	MATERIAUX MACROMOLECULAIRES
M.A.	MANDRAND (Mme)	GENETIQUE MOLECULAIRE DES MICROORGANISMES
J.F.	QUINSON	GEMPPM
A.	ROCHE	MATERIAUX MACROMOLECULAIRES
Directeurs de recherche I.N.R.A.		
G.	BONNOT	BIOLOGIE APPLIQUEE
G.	FEBVAY	BIOLOGIE APPLIQUEE
S.	GRENIER	BIOLOGIE APPLIQUEE
Y.	MENEZO	BIOLOGIE APPLIQUEE
Directeurs de recherche I.N.S.E.R.M.		
A.F.	PRINGENT (Mme)	BIOCHIMIE ET PHARMACOLOGIE
I.	MAGNIN (Mme)	CREATIS***

GEMPPM* : Groupe d'étude métallurgie physique et physique des matériaux

LAEPSI** : Laboratoire d'analyse environnementale des procédés et systèmes industriels

CREATIS*** : Centre de recherche et d'applications en traitement de l'image et du signal

ECOLE DOCTORALES

➤ MATERIAUX DE LYON

INSAL – ECL -UCB. Lyon1 – Univ. De Chambéry – ENS

Responsable : Professeur A. HOAREAU, UCBL (Tél : 04.72.44.85.66)

Formations doctorales associées :

- Génie des Matériaux (Pr. R. FOUGERES, Tél : 04. 72. 43. 81. 49)
- Matière condensée surfaces et interfaces (Pr. G. GUILLOT, Tél : 04.72.43.81.61)
- Matériaux polymères et composites (Pr. H. SAUTEREAU, Tél : 04.72.43.81.78)

➤ MECANIQUE, ENERGETIQUE, GENIE CIVIL, ACOUSTIQUE (MEGA)°

Responsable : Professeur J. BATAILLE, ECL (Tél : 04.72.43.8079)

Formations doctorales associées :

- Acoustique (Pr. J.L. GUYADER, Tél : 04.72.43.80.80)
- Génie Civil : Sols, matériaux, structures, physique du bâtiment
(Pr. P. LAREAL, Tél : 04.72.43.82.16)
- Mécanique (Pr. G. DALMAZ, Tél : 04.72.43.83.03)
- Thermique et Energétique (Pr. M. LALLEMAND, Tél : 04.72.43.81.54)

➤ ELECTRONIQUE, ELECTROTECHNIQUE, AUTOMATIQUE (EEA)

INSAL - ECL – UCB. Lyon1 – Univ. de Saint-Etienne

Responsable : Professeur G. GIMENEZ, INSAL (Tél : 04.72.43.83.32)

Formations doctorales associées :

- Acoustique (Pr. J.L. GUYADER, Tél : 04.72.43.80.80)
- Automatique Industrielle (Pr. SCAVARDA, Tél : 04.72.43.83.41)
- Dispositifs de l'électronique intégrée (Pr. P. PINARD, Tél : 04.72.43.80.79)
- Génie biologique et médical (Pr. I MAGNIN, Tél : 04.72.43.85.63)
- Génie électrique (Pr. J.P. CHANTE, Tél : 04.72.43.87.26)
- Signal, Image, Parole (Pr. G. GIMENEZ, Tél : 04.72.43.83.32)

➤ ECOLE DOCTORALE INTERDISCIPLINAIRE SCIENCES-SANTE (EDISS)

INSAL – UCB Lyon1 – Univ. de Saint-Etienne – Univ. Aix-Marseille2

Responsable : Professeur A. COZZONE, CNRS-Lyon (Tél 04.72.72.26.75)

Formations doctorales associées :

- Biochimie (Pr. M. LAGARDE, Tél : 04.72.43.82.40)
- Génie biologique et médical (Pr. I. MAGNIN, Tél : 04.72.43.85.63)

AUTRES FORMATIONS DOCTORALES

➤ ANALYSE ET MODELISATION DES SYSTEMES BIOLOGIQUE

Responsable : Professeur S. GRENIER, INSAL
Tél : 04.72.43.83.56

➤ CHIMIE INORGANIQUE

Responsable : Professeur P. GONNARD, INSAL
Tél : 04.72.43.81.58

➤ CONCEPTION EN BATIMENT ET TECHNIQUE URBAINES

Responsable : Professeur M. MIRAMOND, INSAL
Tél : 04.72.43.82.09

➤ DEA INFORMATIQUE DE LYON

Responsable : Professeur J.M. JOLION, INSAL
Tél : 04.72.43.87.59

➤ PRODUCTIQUE : ORGANISATION ECONOMIQUE ET GENIE INFORMATIQUE POUR L'ENTREPRISE

Responsable : Professeur J. FAVREL, INSAL
Tél : 04.72.43.83.63

➤ SCIENCES ET TECHNIQUES DU DECHET

Responsable : Professeur P. MOSZKOWICZ, INSAL
Tél : 04.72.43.83.45

Avant-Propos

Avant toute chose, je tiens à remercier mon directeur de thèse, Mme Nicole VINCENT, maintenant Professeur à l'université Francois Rabelais de Tours, pour la confiance qu'elle m'a toujours accordée, pour le temps précieux qu'elle m'a consacré, pour sa patience et son soutien constant à tous les niveaux, au cours de ces années de travail. Qu'elle veuille bien trouver ici l'assurance de ma gratitude, de ma plus profonde estime et de mon amitié.

Merci à Mr Hubert EMPTOZ, Professeur à l'INSA et directeur du laboratoire RFV, pour les conseils qu'il m'a prodigués et pour la confiance qu'il a su m'accorder tout au long de la préparation de ce doctorat.

Je remercie Mr Jean-Pierre COCQUEREZ, Professeur à l'Université de Cergy-Pontoise, pour l'intérêt qu'il a porté à mon travail et pour le plaisir qu'il me fait en acceptant d'en être le rapporteur. Je lui en suis extrêmement reconnaissant. Je le remercie d'avoir bien voulu présider ce jury.

Merci également à Mr Guy LORETTE, Professeur à l'université de Rennes 1, pour avoir accepté d'examiner cette thèse. Je lui sais gré du temps qu'il m'a consacré avec gentillesse afin d'améliorer la version finale de ce document.

Je remercie Mr Jean-Marc BRUN, Professeur à l'université Claude Bernard de Lyon, sans qui ce travail n'aurait pas eu lieu, d'avoir en plus accepté de le juger.

Je remercie Mr Karl TOMBRE, chargé de recherche à l'INRIA-Lorraine, de l'intérêt qu'il a manifesté pour mon travail, des conseils qu'il m'a donnés pour le mener à bien au travers de ses publications, et du plaisir qu'il me fait finalement en acceptant de le juger.

Mes plus vifs remerciements vont également à tous les membres de l'équipe RFV que j'ai côtoyés durant ces dernières années ; travailler avec eux fut un réel plaisir. Un grand merci à tous ceux avec qui j'ai eu des échanges privilégiés, que ce soit sur le plan scientifique, culturel, sportif ou tout simplement humain. Merci particulièrement à Laurence et Viviane pour m'avoir supporté (et me supporter encore) dans leur bureau. Merci à Claire pour sa disponibilité, son aide et notamment pour sa minutieuse relecture de mon manuscrit. Sa conscience professionnelle fut pour moi un exemple (et le reste). Merci aux amis en somme !

Enfin, j'adresserai mon dernier remerciement, mais non le moindre, à ma famille (Sandrine, Thomas, Véronique, mes parents, ...) pour son indispensable soutien et pour tout le reste ...

Sommaire

<u>Introduction</u>	11
1. Le document : support de la communication écrite	12
2. L'interprétation automatique de documents	14
3. Philosophie de notre méthode	15
4. Organisation de ce mémoire	17
<u>1. Le point sur la situation actuelle</u>	19
1.1. Les différents types de documents techniques	20
1.1.1. Les projections orthogonales	21
1.1.2. Les schémas de principe	22
1.1.3. Les plans	24
1.1.4. Les organigrammes et autres dessins	24
1.2. Evolution des techniques d'analyse	26
1.2.1. Analyse du document sur la base de sa structure physique	26
1.2.2. Les acquisitions : méthodes d'analyse	27
1.2.3. Les problèmes qui subsistent	35
1.3. Vers de nouvelles méthodes	37
1.3.1. Interprétation sur la base de la structure logique	37
1.3.2. Méthodes syntaxiques et méthodes sémantiques	39
1.3.3. Les problèmes qui subsistent	42

1.4. Intelligence artificielle distribuée.....	43
1.4.1. Organisation des données et des traitements.....	44
1.4.2. Modèles de communication entre agents	45
1.4.3. Structure de contrôle	45
1.4.4. Le modèle Blackboard	46
1.5. Conclusion.....	48
<u>2. Perception globale d'un document</u>	51
2.1. Bases conceptuelles.....	52
2.1.1. Perception humaine	52
2.1.2. Les documents techniques.....	53
2.1.3. Modélisation.....	54
2.2. Le modèle de description	56
2.2.1. Construction des vecteurs.....	56
2.2.2. Composantes connexes	60
2.2.3. Construction des quadrilatères	64
2.3. Apport du modèle de description	71
2.3.1. Une vision globale.....	71
2.3.1.1. Directions privilégiées.....	72
2.3.1.2. Epaisseurs des traits	73
2.3.1.3. Type de forme	75
2.3.2 Partage de l'information.....	85

<u>3. Lire et comprendre un document</u>	89
3.1. La méthodologie.....	90
3.2. Les processus d'analyse locale.....	92
3.2.1. Extraction du texte	92
3.2.2. Extraction des lignes	97
3.2.3. Extraction des courbes	104
3.2.4. Extraction des éléments mécaniques.....	114
3.2.4.1. Construction du graphe structurel de l'image	115
3.2.4.2. Reconnaissance des éléments mécaniques.....	118
3.2.5. Reconstruction 3D du mécanisme.....	126
3.3. Coopération et cycles perceptifs	133
3.4. Résultats	138
3.4.1. Un exemple complet.....	138
3.4.2. Images tests	144
<u>Conclusion</u>	147
<u>Références bibliographiques</u>	153

Introduction

1. Le document : support de la communication écrite

Le document, qu'il soit électronique ou non, est présent dans notre environnement professionnel, social, et culturel. Il revêt de multiples facettes : il y est objet de production, vecteur de communication, outil d'aide à la décision, ... Les avancées technologiques, nous ont catapulté dans un monde où l'information est devenue une quantité mesurable et où la définition du mot «document» est en constante évolution, incorporant tous les aspects de la communication écrite. L'arrivée des systèmes de gestion électronique de documents (GED) qui permettent de créer, de manipuler, de stocker, de retrouver et d'archiver des documents (des informations) a d'abord laissé penser que le support papier allait disparaître. Aujourd'hui, il subsiste de nombreux types d'informations pour lesquels l'intégration n'a pu être encore effectuée complètement. Ces systèmes, loin d'être universels, ont une autonomie encore incomplète par rapport au support papier. Même si le développement de ces systèmes reste très récent, il semble que le papier est, et restera encore très longtemps, un composant important dans notre vie de tous les jours. Ceci entraîne un besoin de convertir les documents papier en données numériques utilisables par la machine, besoin qui, à mon avis, ne cessera d'augmenter.

La gestion électronique de documents doit inclure, non seulement le stockage et la restitution, mais aussi un niveau de compréhension de l'information fournie par le média, sans quoi l'information ne peut être effectivement utilisée. Il ne s'agit plus alors, pour les systèmes de conversion automatique (lecture automatique de dessins, systèmes OCR, ...), d'effectuer l'analyse de la structure physique du document mais de comprendre le contenu et même d'intégrer des informations sur le processus qui est à l'origine du document.

Les documents techniques, au sens où nous l'entendons par la suite, sont des documents essentiellement graphiques, produits par un dessinateur, à la main ou par un système CAO. Les documents techniques sont le moyen de communication le plus efficace entre les différents acteurs de la conception et de la fabrication d'un produit, du bureau d'étude à l'atelier. Ils obéissent à des normes souvent strictes (quoique parfois non respectées !) mais doivent pouvoir être compris par tous les acteurs de cette chaîne. L'avènement des systèmes de GED et de CFAO rend nécessaire la représentation de ces informations sous forme électronique. L'interprétation de documents techniques consiste donc en partie à convertir des documents disponibles sur support papier en une représentation numérique de haut niveau.

Le travail présenté dans cette thèse s'inscrit dans le domaine de l'interprétation automatique de dessins techniques. Plus précisément, il porte sur la récupération de schémas cinématiques mécaniques réalisés sur papier (Figure 1) par des concepteurs en mécanique. Ces schémas représentent le fonctionnement d'un mécanisme mécanique composé de plusieurs pièces. En outre des informations implicites peuvent être contenues dans des configurations géométriques spécifiques comme un parallélisme. Ces dessins comportent toujours un ensemble de symboles normalisés représentant différentes liaisons mécaniques : engrenages, pivots, ... Les liens entre ces différentes formes correspondent aux pièces qui entrent dans la construction du mécanisme décrit par le dessin. Ces représentations peuvent être réalisées en 2 ou en 3 dimensions mais traduisent toujours l'aspect final du mécanisme. Lorsque le schéma est en 3 dimensions, de la même manière qu'en vision par ordinateur, on devra interpréter une scène 3D projetée sur un plan 2D.

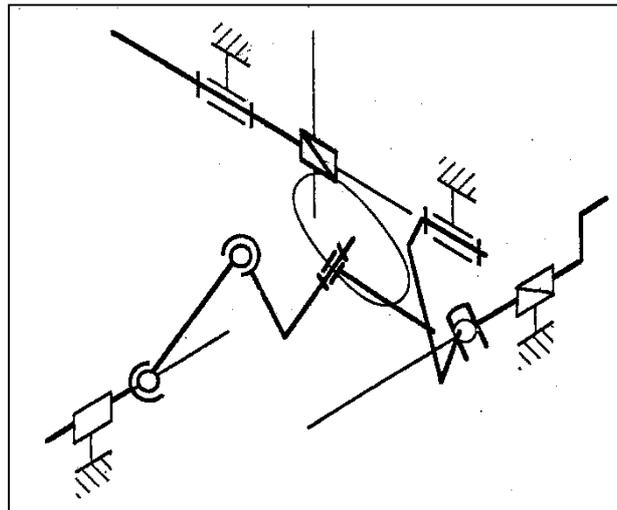


figure 1 : Exemple de dessin cinématique

L'étude que nous désirons réaliser a pour objectif de déterminer, à partir de la géométrie d'une pièce dessinée, les différentes formes et/ou entités qui la constituent, telle qu'elle a été dessinée par le concepteur. Cette transformation nécessite le passage du dessin à un modèle géométrique, puis le passage du modèle géométrique à une liste de formes mécaniques caractéristiques qui pourront ensuite servir d'entrées pour des logiciels plus spécialisés en conception et simulation.

Dans les systèmes de CAO modernes, de nombreuses informations fonctionnelles et sémantiques doivent être attachées aux primitives graphiques d'une représentation. Un document technique ne contient pas seulement de l'information géométrique mais aussi de

l'information sur les fonctionnalités des pièces représentées, des annotations, des références à un catalogue, ... Pour qu'un système de conversion automatique puisse être utilisé, il doit prendre en compte ces évolutions.

2. L'interprétation automatique de documents

L'objectif de l'interprétation automatique de documents est précisément de faciliter l'exploitation des dessins en s'appuyant sur un ensemble d'étapes permettant la transformation du schéma papier en un stockage numérique qui puisse être interprété automatiquement. Il s'agit alors d'extraire l'information graphique et alphanumérique du document pour reconstituer la sémantique associée aux différents objets du dessin et à leurs relations.

Les premières tentatives de création de systèmes d'analyse de documents se voulaient générales ; elles n'introduisaient que très tard dans la chaîne des traitements, les connaissances a priori sur la nature du document traité. Ce type de méthode s'est révélé limité quant aux possibilités, en sortie des modules d'interprétation, de présenter un niveau d'abstraction suffisamment élevé. Notre but sera plus de comprendre le message apporté par le document que de reconnaître les formes géométriques qui le constituent.

La notion de *système à base de connaissances* fut alors introduite. Elle était adaptée au problème puisqu'elle permettait d'incorporer de nombreuses connaissances de natures très différentes. Ces informations pouvaient être de bas niveau (caractéristiques structurelles ou géométriques des objets de l'image) ou de très haut niveau (*contraintes de cohérence* sémantique qu'un objet doit respecter).

Cependant, les expérimentations que nous avons effectuées en utilisant des méthodes classiques, ainsi que les lectures que nous avons faites, ont confirmé l'hypothèse que les traitements de bas niveau entraînent encore très souvent des pertes d'information. Ils restent donc à améliorer. Les systèmes utilisant un contexte global pour l'interprétation dès les premières étapes du traitement sont encore très rares.

Malgré quelques succès relatifs, les systèmes commerciaux sont loin de répondre aux besoins exprimés. Cela explique que plusieurs entreprises, après avoir beaucoup investi dans la vectorisation, retournent à des offres de plus bas niveau : systèmes semi-automatiques offrant différents outils automatisés mais laissant un opérateur humain aux commandes pour l'essentiel de la reprise intelligente du document.

L'intérêt des systèmes d'interprétation de documents ne s'arrête pas au domaine de la mécanique. De nombreuses recherches ont été effectuées sur différents types de schémas, schémas électriques comme les réseaux de distribution, schémas mécaniques comme les projections orthogonales, schémas géographiques comme les cartes et les plans cadastraux, partitions musicales, ... Le but est, dans tous les cas, de supprimer la fastidieuse phase de saisie manuelle des dessins sur un système informatique.

Toutes ces remarques semblent montrer qu'il n'existe pas de solution générique à l'interprétation automatique de documents, mais que dans certains cas précis, il est possible d'obtenir suffisamment de connaissances contextuelles pour permettre une analyse robuste et une interprétation automatisée.

3. Philosophie de notre méthode

Traditionnellement, en reconnaissance de formes, une distinction est faite entre deux phases successives et uniques : la phase d'extraction d'indices et la phase de reconnaissance. Ceci s'explique par le fait que les techniques d'extraction d'indices sont souvent générales et applicables à différents types de documents tandis que les méthodes de reconnaissance sont plus spécifiques puisqu'elles utilisent des données du domaine d'étude.

Dans le système visuel humain, cette séparation n'est pas aussi franche et figée, c'est pourquoi nous nous sommes efforcé, d'une part, de ne pas séparer aussi clairement ces 2 phases, et d'autre part, de rendre le processus d'interprétation plus souple. Bien que le fonctionnement du système visuel humain soit loin d'être encore complètement compris, il est possible d'en tirer de précieux enseignements. Nous sommes tous capables de reconnaître le type d'un document sans en connaître le contenu, de discriminer le texte du graphique, le manuscrit de l'imprimé, de nous placer très rapidement au niveau de détail approprié à la compréhension et ceci pratiquement sans effort. Ce type de remarques montre au moins que l'interprétation de ce qui nous entoure ne se réduit pas seulement à une simple mise en correspondance d'images : «Pattern Matching».

Même si on ne parvient pas encore à modéliser l'acte de lire, pour le transcrire sous forme de règles par exemple, celui-ci n'est plus l'opération mystérieuse qui permet à un «honnête» lecteur de maîtriser 10 mots à la seconde. On sait maintenant que le lecteur enchaîne sans

cesse des activités complexes telles que la **prise d'indices** (graphiques, lexicaux, syntaxiques, sémantiques), l'**élaboration d'hypothèses** (permettant d'anticiper), la réalisation de **corrections éventuelles**; ces activités variant suivant les objectifs.

Pour les schémas qui nous intéressent, avec l'aide d'experts mécaniciens, nous avons remarqué qu'il est possible de déduire, du dessin lui-même, une partie des connaissances nécessaires à l'interprétation. Ainsi, nous avons choisi d'inclure une étape préalable d'**extraction des caractéristiques globales** pour acquérir les éléments indispensables à la réalisation d'une analyse locale plus adaptée et donc plus précise.

L'homme semble posséder une habileté très forte à **extraire les évidences** en faisant **appel aux vastes connaissances qu'il a acquises**. Lorsqu'il lit, des échanges s'effectuent constamment entre l'information visuelle (image du document) et l'information non visuelle (arrière-plan conceptuel) : ce que l'on sait déjà sur le document. L'interface entre les deux est la mémoire immédiate (de faible capacité), sorte de file d'attente des prises d'indices, la dernière entrée chasse impitoyablement la première de la file.

Pour notre cas, nous avons choisi de tirer parti des stratégies à base d'**objets** permettant, notamment, une meilleure distribution des connaissances. Plus précisément, après l'intégration des systèmes experts dans le domaine de l'interprétation automatique, nous pensons que l'utilisation de systèmes **multi-spécialistes** peut permettre de nombreuses avancées.

L'utilisation conjointe de plusieurs «**spécialistes**» et d'un **modèle de représentation** des documents pour le partage de l'information permet, au cours de la reconnaissance et de la reconstruction des objets, d'utiliser de façon duale **vision globale** et **vision locale**. Un mécanisme de "**cycles perceptifs**" permet aux spécialistes de coopérer pour extraire seulement les **évidences** et de limiter donc les erreurs d'analyse. En fin de cycle, la complétude de la compréhension du document est vérifiée et un nouveau cycle d'activation des spécialistes permet si nécessaire d'avancer encore dans l'interprétation.

4. Organisation de ce mémoire

Dans un premier chapitre, nous décrivons l'état actuel des recherches en interprétation automatique de documents et mettons en évidence aussi bien les avancées réalisées depuis quelques années que les lacunes qui restent à combler.

Au cours de nos recherches, nous avons essayé de rapprocher les systèmes d'interprétation artificiels, du système visuel humain. En effet, si le problème de reconnaissance géométrique d'un document est un problème typique de vision par ordinateur ou d'analyse de scène, par contre, la prise en compte de l'information sémantique véhiculée par le document nécessite une analyse plus linguistique de ce dernier. Un document n'est pas seulement une scène mais aussi un langage et il faut en tenir compte lors de son **interprétation**. Deux aspects distincts sont alors mis en évidence dans ce processus :

- le premier, qui fait l'objet du second chapitre, se réfère à ce que nous appelons la **perception globale des documents**. Nous décrivons, dans ce chapitre, les problèmes relevant principalement de l'analyse d'image, c'est à dire les mécanismes qui nous ont permis de mettre en évidence les informations fondamentales concernant l'aspect global du document afin de construire le **modèle de représentation** à même de décrire tous les types de dessins techniques. Cette première structuration permet de réduire la quantité de données à traiter et constitue une base sur laquelle s'appuie le reste de l'analyse.

- le second concerne la compréhension du document (plus particulièrement les schémas cinématiques). Dans le chapitre 3 intitulé lire et comprendre les documents, nous montrons qu'il est possible de faire "**coopérer**" **des processus** pour fusionner les informations fournies par le modèle de représentation avec celles relevant du **langage** utilisé dans le document et qui dépendent beaucoup plus de l'application visée. Le but est de faire **évoluer le modèle de représentation** jusqu'à ce qu'il traduise totalement le message contenu dans le document. Nous décrivons le fonctionnement de chaque spécialiste, son utilisation des données préalablement obtenues, et montrons qu'il est possible de faire coopérer ces spécialistes par l'intermédiaire de cycles perceptifs.

S'inspirer de la démarche humaine lors de la lecture et de la compréhension d'un document afin de perfectionner les systèmes d'interprétation automatique est le but que nous nous

sommes fixé. Bien que les chercheurs soient satisfaits de reproduire ne serait-ce qu'une infime partie des facultés du système visuel humain, bien que de nombreux succès aient été réalisés, de nombreuses questions demeurent sans réponse. Nos recherches se concentrent autour de ces problèmes et nous espérons répondre à certaines de ces interrogations au cours des différents chapitres de ce mémoire.

Chapitre 1

Le point sur la situation actuelle

Il nous paraît nécessaire, afin de mieux situer notre apport à la question de l'analyse de documents techniques, particulièrement d'actualité aujourd'hui, de définir d'abord les documents techniques et de montrer ensuite quel fut le cheminement de l'activité de recherche depuis les débuts, il y a quelques décennies, jusqu'aux dernières avancées décisives qui ont été réalisées il y a peu. La question étant par ailleurs encore loin d'être complètement résolue, nous nous attacherons, tout au long de cette partie, à mettre en lumière les problèmes rencontrés.

1.1. Les différents types de documents techniques

Un document a normalement pour but de transmettre un ensemble d'informations à des destinataires. Traditionnellement le support de communication de ces informations est le papier. Les documents peuvent présenter différents aspects : formulaires, documents composites, dessins techniques, ... Chaque type possède une structure qui lui est propre, choisie de manière à ce que les informations qu'il véhicule soient aisément compréhensibles et réutilisables par les destinataires.

Parmi les documents, les dessins techniques forment une catégorie particulière; ils sont constitués de lignes, de régions pleines, de régions hachurées, de texte, ... Ils contiennent toujours une quantité importante d'informations et peuvent être d'une grande complexité [Belaid92a] et leur domaine d'application est très étendu. On s'accorde à distinguer parmi eux quatre grands types, qui ont fait l'objet de nombreux travaux de recherche. Tombre les classe en 4 catégories [Tombre92]:

- *Les projections orthogonales*
- *Les schémas de principe*
- *Les plans*
- *Les organigrammes et autres dessins*

1.1.1. Les projections orthogonales

Ce sont des vues planes d'objets réels représentés depuis différents points d'observation (figure 1.1). L'image comprend la plupart du temps un ensemble de vues 2D, avec des symboles et une nomenclature (liste complète des pièces ou parties de l'objet représenté). Une partie des lignes représente le mécanisme, l'autre partie fait référence à la cotation. Les traits forts représentent généralement la projection sur un plan, des contours d'une section ou des vues de l'objet. Les lignes en traits interrompus représentent les contours cachés; les lignes mixtes les axes de symétrie, les parties hachurées indiquent la présence de matière. Les symboles (caractères, flèches, ...), les traits fins et les références à la nomenclature constituent la cotation. La cotation fournit les informations complémentaires, indispensables à la compréhension totale du dessin [Collin91]. Toutes ces conventions ou règles sont définies dans différentes normes : DIN, BSN, NDS, ... et témoignent de la rigueur avec laquelle ces dessins sont exécutés.

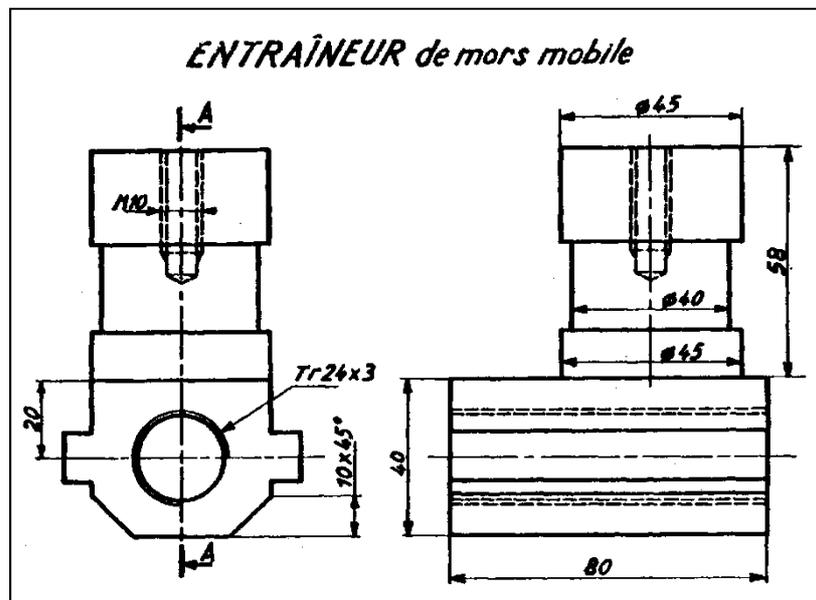


Figure 1.1 : Exemple de projections orthogonales

Le but de l'analyse est, dans ce cas, la conversion du dessin (souvent celui d'une pièce mécanique) en une description utilisable par un système de CAO. Les informations à extraire de l'image sont donc les entités qui figurent dans la base de données de tels systèmes ou même des informations d'un niveau supérieur (groupement d'objets élémentaires) représentant une structure élaborée (boîte de vitesse, ...), entité que l'on retrouve, cette fois, dans les bibliothèques graphiques du système.

Une fois toutes les vues planaires analysées et reconstituées, l'ultime étape serait la reconstruction 3D de l'objet par combinaison des vues 2D obtenues. Ce traitement requiert une reconnaissance quasiment parfaite des vues 2D, ce qui est encore rarement le cas avec les systèmes de reconnaissance actuels.

1.1.2. Les schémas de principe

Ils concernent les plans de câblages électriques, électroniques, ou téléphoniques, ... (figure 1.2) ou encore des schémas cinématiques (figure 1.3). L'information véhiculée, évoque assez rarement l'aspect externe d'un objet réel; il s'agit plutôt de la description du principe de fonctionnement d'un dispositif à partir de formes logiques (les symboles), de liens, de connexions entre les symboles, et de texte.

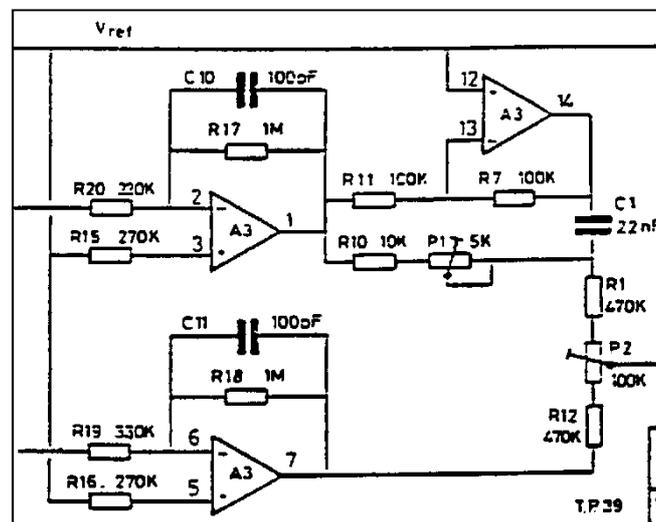


Figure 1.2 : Exemple de schéma de câblage

C'est sur des dessins de cette dernière classe que nous décrirons, en la justifiant, l'approche que nous proposons. Il s'agit de *schémas de fonctionnement de processus mécaniques ou de schémas cinématiques*. Même si ces derniers répondent complètement à la définition donnée ci-dessus, il est à noter que ces dessins reflètent, en outre, la structure externe de l'objet grâce à une schématisation, parfois dans l'espace (en 3 dimensions), du mécanisme (figure 1.3).

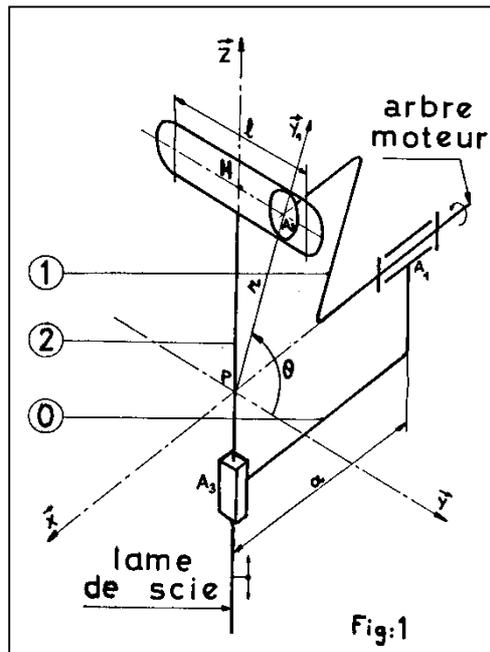


Figure 1.3 : Exemple de schéma cinématique

Les formes à extraire des schémas de principe sont, généralement :

- un ensemble de symboles normalisés (dépendant de l'application et constitués d'une variété de segments, de boucles, ou de régions pleines, ...),
- des liens de connexion entre ces symboles,
- du texte.

Le nombre des symboles est souvent très important et leurs formes comportent, en outre, de nombreuses similitudes [Okasaki88].

Dans le cas des schémas cinématiques, l'analyse porte à la fois sur la position relative des objets et sur leur reconnaissance.

De même que pour les dessins de projections orthogonales, le but de l'analyse est la conversion du dessin afin de le rendre utilisable par un système de CAO qui permettra l'archivage, les modifications, ou même les simulations de fonctionnement du processus. Ce dernier type d'application est caractérisé par la présence, dans ces schémas, d'informations sémantiques. Ces informations sont éventuellement utilisables durant les différentes phases du traitement.

1.1.3. Les plans

Il s'agit de vues cartographiques, aériennes ou satellites, de plans cadastraux, ... (figure 1.4). Ce sont des représentations schématiques de vues d'ensemble, plus ou moins détaillées et présentant différents types de renseignements (selon l'application visée). De tels documents sont généralement constitués d'une superposition d'informations de caractère géographique, géologique, météorologique, ... Un symbolisme ou des codes de couleurs, d'épaisseurs des traits, de hachures, ... sont utilisés pour différencier chaque type de données (réseaux routiers, rivières, immeubles, canalisations, courbes de niveau, ...). Il s'y ajoute parfois des annotations (le texte) mentionnant le nom des villes, des fleuves, ou qualifiant certains attributs d'objets.

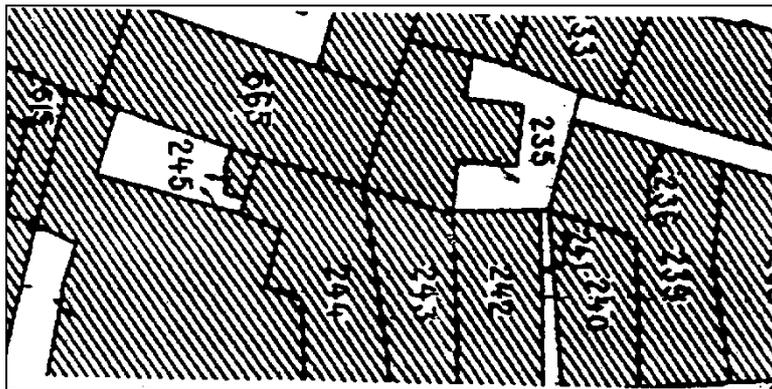


Figure 1.4 : Exemple d'extrait de plan cadastral

Généralement, le but recherché est l'extraction d'une partie des informations contenues dans ces images (par exemple la détection des immeubles), c'est-à-dire la détection de certaines zones ou formes (souvent construites à partir de segments de droites ou de courbes), noyées parmi toutes les autres données que l'on peut alors considérer à ce stade comme du bruit. Les données récupérées seront fournies à un système de CAO cadastral ou à un système d'informations géographiques (S.I.G.).

1.1.4. Les organigrammes et autres dessins

Cette catégorie englobe les schémas représentant une organisation (organigrammes de programmes, modélisation de données, représentation d'une hiérarchie d'entreprises, ...) et les documents de type tableau (figure 1.5).

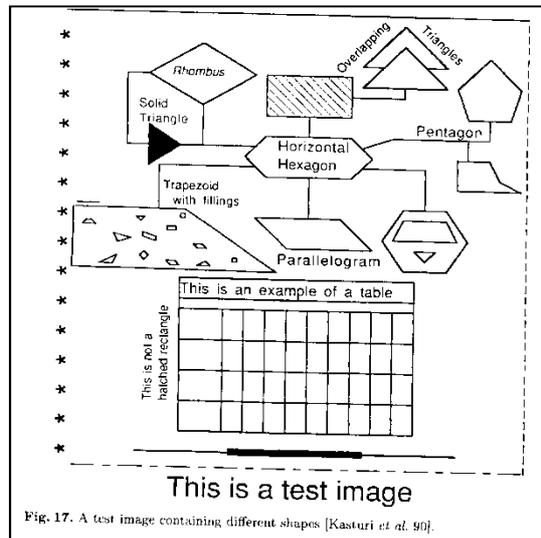


Fig. 17. A test image containing different shapes [Kasturi et al. 90].

Figure 1.5 : Exemple d'organigramme.

Les organigrammes sont caractérisés par la présence de nombreuses zones de texte. Les formes à reconnaître sont généralement des polygones, lignes, cercles, et autres objets géométriques élémentaires.

Les objectifs sont donc l'extraction du texte et l'extraction d'attributs relatifs aux objets élémentaires afin de pouvoir fournir une description des relations spatiales existant entre les composants (inclusion, lien, intersection, ...) [Kasturi90]. Il s'agit parfois de dessins faits à main levée qui constituent, après remise en forme, l'entrée d'un système d'édition de graphiques (dessins ou tableaux). Nous rangerons dans cette catégorie, les dessins d'objets vus en perspective bien que ce type de document puisse être considéré comme un dispositif d'entrée, aussi bien vers un éditeur graphique, que vers un système de CAO.

Il est à noter que tous ces documents techniques peuvent se présenter sous différents aspects. Ils peuvent avoir été *dessinés à main levée* ou, au contraire, à l'aide d'instruments graphiques (règles, normographes, ...) et peuvent avoir ou non subi des dégradations (photocopies successives, ...). *La qualité* d'une image se définit alors sur des critères d'orientation, de bruits (taches noires parasites ou blancs induisant des discontinuités), et de distorsion des formes. Face à ces défauts, dus aux techniques d'acquisition, mais aussi à la mauvaise qualité du document original lui-même, les méthodes choisies devront être plus ou moins robustes et il conviendra donc de bien définir le type et la nature des documents traités par chaque système.

1.2. Evolution des techniques d'analyse

1.2.1. Analyse du document sur la base de sa structure physique

Il peut sembler hasardeux de tenter de **définir une structure commune** à tous ces documents. Une étude des dessins décrits précédemment permet cependant de mettre en évidence un certain nombre de points communs.

Il est possible de mettre en évidence une organisation en couches pour chacun de ces documents [Tombre92]. Il s'agit de la **structure physique** des documents, qui dans tous les cas se compose :

- *d'une couche graphique* qui contient la majeure partie de l'information. On peut parfois la diviser en plusieurs autres sous couches selon des critères purement structurels ou géométriques (épaisseur du trait, style du trait, hachures, ...). Ces sous couches dépendent fortement du domaine d'application auquel on s'intéresse, c'est-à-dire du contenu logique des documents.

- *d'une couche texte* qui se présente généralement sous la forme de groupes de composantes connexes de petites tailles constituant soit un mot, soit un nombre, soit une annotation. La position du texte par rapport au graphique indique à quelle partie du dessin se rapporte la cote ou la légende; il est donc important de conserver les positions relatives de ces composantes. Le vocabulaire utilisé est souvent, dans chaque dessin, relativement réduit et les caractères composant le texte sont plus ou moins aisément séparables par des critères topologiques selon la nature du dessin.

- *d'autres couches éventuelles* composées, par exemple, des petits symboles (têtes de flèches, disques, ...)

L'autre point commun entre ces différents dessins concerne les formes des objets à extraire des images. Tout système d'analyse de documents techniques manipule **des entités de type structurel** formées à partir de segments de droite et d'arcs tels que lignes, polygones, cercles, courbes. D'ailleurs, dans la plupart des systèmes [Joseph92,Cugini84,Vaxivière94], une étape de vectorisation permet d'obtenir une description de l'image sous forme de suites de vecteurs plutôt que sous forme de pixels. Le premier pas vers une description, puis vers une interprétation de type structurel, est ainsi réalisé. Il peut être suivi d'une phase d'identification d'entités géométriques simples, obtenues par combinaison des vecteurs de base, telles que les polygones [Kasturi90].

Le problème de l'interprétation automatique de document semble donc alors résider dans la phase de séparation puis de reconnaissance de chacune de ces couches, et dans le choix de l'ordre d'extraction le plus approprié. Cela ne va pas sans difficulté puisque, comme nous venons de le voir, toutes les couches sont constituées d'objets géométriques eux-mêmes formés à partir de primitives d'un même type (arcs et segments).

1.2.2. Les acquisitions : méthodes d'analyse

La plupart des systèmes d'analyse de documents techniques réalisent, en premier lieu, une identification des formes (composantes connexes, formes fines, formes pleines, ...) présentes dans l'image [Kasturi92]. Les objets constituant les dessins sont ainsi regroupés suivant les méthodes de traitement que l'on pourra leur appliquer par la suite. Une fois cette séparation réalisée, chaque couche continue à être traitée indépendamment, par l'intermédiaire de méthodes variées (vectorisation, détection de contours, maillage, ...), chacune étant choisie en fonction du type de forme à extraire (figure 1.6).

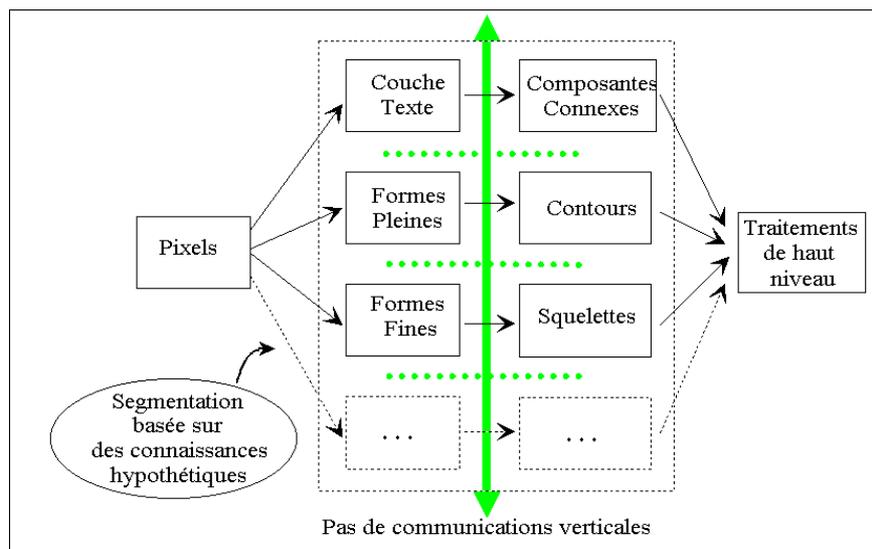


Figure 1.6 : Méthode classique d'analyse de bas niveau

Cette façon d'aborder le problème peut être suffisante pour certaines applications, notamment pour les organigrammes ou pour tout autre document comportant simplement une couche Texte et une couche Formes Fines (ou un nombre limité de couches). La reconstruction des entités (géométriques) est alors réalisée assez simplement si chacune de ces couches peut être analysée indépendamment et si les formes à regrouper ne sont pas dispersées dans différentes couches.

Au plus bas niveau se trouvent les algorithmes opérant directement sur les pixels de l'image pour en extraire des informations plus structurées, telles que suivi de contours, vectorisation, filtrage, ... Des outils spécifiques, dédiés à cette *phase d'analyse lexicale*, ont été développés de façon à traiter chaque forme de la manière la plus adéquate possible.

Analyse morphologique

Certains proposent de privilégier l'extraction des zones d'intérêt au moyen de traitements de type morphologique (ouverture, érosion/dilatation) pour éliminer les formes fines d'un dessin. Kasturi utilise ce principe pour localiser et extraire les régions pleines d'un document [Kasturi90]. Une succession d'érosions permet d'éliminer les segments, arcs et symboles de l'image. Reste, ensuite, à redonner aux objets encore présents dans l'image leur forme initiale par dilatations successives. La dilatation n'étant pas exactement l'opération inverse de l'érosion, des traitements supplémentaires sont encore nécessaires pour obtenir une restitution exacte des symboles pleins originels.

En règle générale le recours au codage du contenu de l'image est indispensable afin de structurer, pour ensuite traiter et décrire l'image. Cette transformation doit posséder les caractéristiques suivantes :

- conservation de l'information intéressante contenue dans l'image,
- réduction de la place nécessaire au stockage,
- simplification et adaptabilité du nouvel espace de représentation au traitement à réaliser.

Il est difficile de trouver une représentation donnant pleine satisfaction à ces trois contraintes. De manière pratique on choisit un compromis en privilégiant les aspects importants du problème traité. Comme nous l'avons vu précédemment, l'une des caractéristiques principales des documents techniques est qu'ils sont constitués essentiellement à partir de deux primitives élémentaires : le *segment de droite* et l'*arc*.

Nous allons donc voir les différentes méthodes permettant d'exploiter cette structure spécifique pour passer de la structure matricielle de l'image à une représentation plus adaptée aux dessins étudiés.

La représentation de l'image sous forme de vecteurs est la plus fréquemment utilisée, même si l'on a parfois fait appel à d'autres techniques. Ce passage de la forme matricielle de l'image à une description sous forme de vecteurs se nomme *vectorisation*.

Le problème de la vectorisation

De nombreux systèmes effectuent d'abord une squelettisation [Antoine91, Luo95] pour obtenir des segments, des arcs et des courbes (primitives élémentaires) ayant une épaisseur de 1 pixel. Les traitements suivants se trouvent ainsi simplifiés. Cependant, la squelettisation pose certains problèmes comme l'apparition de barbules, la perte d'information (épaisseur du trait), ou l'inadaptabilité au traitement des formes pleines.

Aussi, plutôt que d'exécuter des analyses complémentaires, d'autres techniques basées sur *le suivi de trait* tentent de pallier ses inconvénients.

Pour détecter les traits présents dans un document, certains travaux proposent de se déplacer de point en point en restant soit à l'extérieur de l'objet (suivi de contours), soit à l'intérieur de l'objet (suivi de traits); l'information extraite est généralement codée à l'aide des directions de Freeman.

Abe [Abe86] propose une méthode de suivi de contours adaptée à l'analyse des organigrammes. Grâce à cette technique, les boucles fermées, utiles par la suite, sont extraites durant la vectorisation. Cugini [Cugini84] emploie une technique d'appariement de contours pour obtenir une représentation, sous forme de vecteurs, des projections orthogonales. D'autres systèmes couplent l'information contours avec l'information squelette pour améliorer la qualité des résultats [Leplumey95, Hori93, Tanigawa94]. Il ressort de ces études que la mise en correspondance des contours est moins sensible au bruit que la squelettisation mais que sa mise en place est beaucoup plus complexe.

Suivi de traits

Paquet par contre propose une méthode de suivi de trait. Dans une image binaire, une ligne peut être vue comme un ensemble de pixels noirs entre deux régions blanches. Partant d'un point initial, il s'agit alors de progresser à l'intérieur de la ligne jusqu'à l'extrémité opposée.

Pour progresser, Paquet [Paquet90, Paquet91] définit comme direction locale de la ligne (figure 1.7), la direction de Freeman permettant le plus grand déplacement à l'intérieur de la ligne. La distance entre la frontière (transition Noir/Blanc) de droite et celle de gauche détermine l'épaisseur locale du trait dans le cas des lignes simples (pas de croisement).

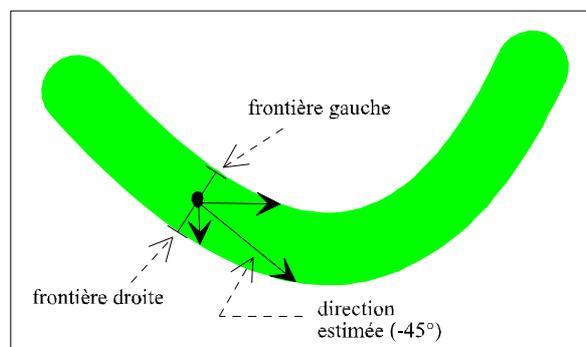


Figure 1.7 : Suivi de ligne

L'algorithme de suivi estime, à chaque étape, les directions successives de la ligne et fournit la liste des points particuliers : jonctions (épaisseur anormale) ou extrémités. Paquet propose d'utiliser un pas de progression proportionnel à l'épaisseur moyenne de la ligne suivie. Cette méthode est utilisée par Ogier [Ogier94a] pour l'analyse de plans cadastraux. L'équipe de Dori utilise une méthode similaire (Orthogonal Zig-Zag) permettant, en plus, la détection d'arc de cercle [Dori95].

Certains ont préconisé l'utilisation d'images en niveaux de gris

Les algorithmes présentés jusqu'à présent travaillent sur des images binaires. Le système de lecture de dessins mécaniques ANON, développé par Joseph et son équipe [Joseph89], travaille directement sur les images en niveaux de gris. L'image originale est divisée en régions, l'étude de l'histogramme des niveaux de gris de chaque région fournit le seuil de séparation fond / objets.

A partir d'un premier pixel ligne détecté, l'algorithme de suivi effectue une exploration circulaire du voisinage (centrée sur ce point). Si ce cercle intersecte une ou deux suites de points noirs, l'algorithme teste la présence d'un lien à l'intérieur du cercle par suivi de l'axe médian du trait (figure 1.8), s'il rencontre plus de deux groupes de pixels noirs, l'algorithme prend la décision de suivre celui ayant le niveau de gris le plus élevé. Le processus est répété à partir de l'intersection détectée tant qu'un critère de noirceur est vérifié (seuil déduit de celui séparant les points objets du fond). L'épaisseur du trait est évaluée après chaque étape par sondage suivant la direction perpendiculaire à celle de la ligne. Le diamètre du cercle d'exploration est déduit de cette épaisseur estimée. Un deuxième seuil sur les niveaux de gris permet une tolérance aux légères coupures du trait.

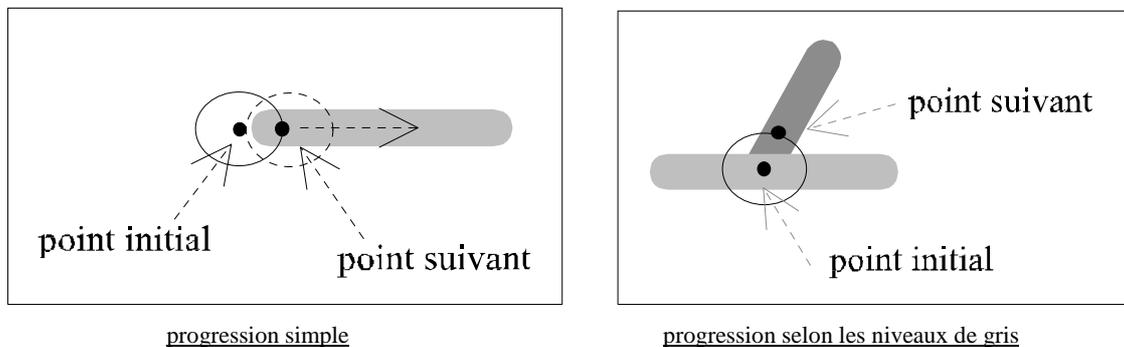


Figure 1.8 : Suivi du trait selon les niveaux de gris

Utilisation des graphes dans le codage des images

Dans le but d'obtenir un codage de l'image sous forme de vecteurs, les méthodes par «codage des segments» (Run Length Coding) balayent l'image ligne par ligne. Les sections horizontales (suites de pixels homogènes) noires et blanches de chaque ligne sont extraites puis codées par des paramètres relatifs à leur position et à leur longueur. Plusieurs algorithmes de codage fondés sur ce principe sont décrits dans la littérature [Ramachandra80, Pavlidis86].

L'algorithme de Pavlidis permet d'obtenir un codage des objets de l'image sous forme de graphes (figure 1.9 : Line Adjancy Graph). Les noeuds du graphe correspondent aux milieux des segments (ou sections) noirs. A chaque noeud (segment) est associé un couple (a,b) de valeurs appelé degré, où a désigne le nombre de segments en contact avec celui-ci sur la ligne de dessus et b le nombre de segments chevauchants situés sur la ligne de dessous. Les noeuds sont reliés par une arête s'ils correspondent à deux segments en contact situés sur deux lignes voisines. Pour un noeud extrémité, $a = 0$ ou $b = 0$, tandis que pour un noeud jonction $a > 1$ ou $b > 1$.

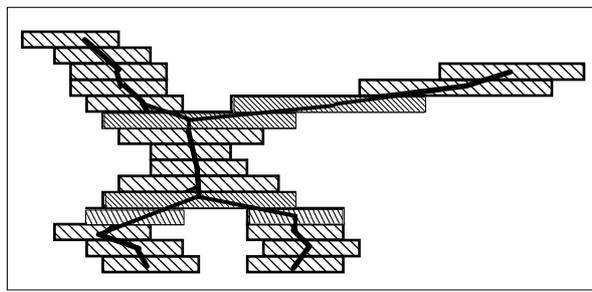


Figure 1.9 : Graphe d'adjacence des lignes

Ramachandran utilise un principe similaire pour effectuer un codage de l'image sous forme de vecteurs. Il effectue un parcours du graphe et, par regroupement, selon des critères de pentes et de longueurs sur les sections, obtient un codage sous forme de vecteurs (figure 1.10).

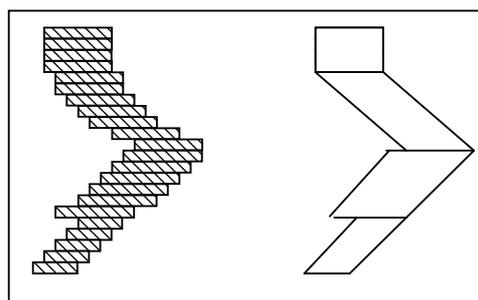


Figure 1.10 : «Vector Coding»

Ce type de technique a souvent été utilisé pour vectoriser des images de type plan cadastral [Madej91, Antoine91], car il s'adapte bien à la détection de zones hachurées. Bley [Bley84] utilise une technique voisine pour vectoriser des images de circuits électriques.

Pour accélérer la phase de vectorisation, il est possible de sous-échantillonner l'image initiale à l'aide d'un quadrillage de dimensions adaptées à l'épaisseur des traits [Lin85]. Cette idée a été reprise et améliorée par Vaxivière [Vaxivière94a] [Vaxivière94b]. L'ajout d'un mécanisme de fusion et de division des mailles facilite ensuite le suivi des traits. Avec cette méthode, il est nécessaire de caractériser précisément les différents types d'intersections entre les traits. Arias propose une description hiérarchique des différents types d'intersections pouvant apparaître dans des plans téléphoniques [Arias93] [Arias94]

Dans tous les cas, il se pose le problème du choix de l'approximation : c'est à dire la méthode de sélection des points critiques. Une approximation se réalise à partir d'une entité de référence qui peut être le segment de droite, l'arc de cercle, ou une courbe d'ordre supérieur. L'approximation polygonale concerne l'approximation à partir de segments de droite. Le calcul d'approximation est fondé soit sur le calcul d'une erreur minimale pour un nombre de segments donnés, soit sur la recherche d'un nombre de segments minimal pour garantir la précision demandée. On peut distinguer 3 grandes classes de méthodes : les méthodes fondées sur la notion de courbure, les méthodes de découpages récursifs et les méthodes d'approximations itératives.

Après l'étape d'approximation polygonale, l'image est représentée par une suite de vecteurs dont on connaît le point initial, le point final et parfois l'épaisseur moyenne du trait. Roosli a proposé récemment d'intégrer 2 primitives (l'arc et le segment) dans un même temps pour approximer les différentes parties des dessins [Roosli95]. Il intègre de plus des contraintes de parallélisme, de perpendicularité, et de concentricité pour obtenir des résultats plus lissés.

Avec la Transformation de Hough, l'objectif envisagé est de définir un changement de représentation de l'image pour que l'ensemble des points appartenant à une forme géométrique spécifique soit identifié par un seul point dans l'espace transformé. Pour cela on exploite la représentation paramétrique des formes géométriques à détecter. Pao [Pao91] propose d'utiliser ce principe pour détecter les droites et ellipses entrant dans la composition de dessins mécaniques.

Analyse par les composantes connexes

De nombreux systèmes d'analyse de documents techniques effectuent un étiquetage des **composantes connexes**. L'analyse des composantes détectées procure d'importantes informations pour la suite du traitement, plus particulièrement pour l'extraction des composantes texte ou l'extraction de certains symboles électriques. La Transformée de Hough permet de regrouper en mots ou phrases les composantes connexes alignées correspondant (parfois!) aux caractères [Fletscher88] [Gloger92]. Les caractères imprimés correspondent, lorsqu'ils ne touchent pas une autre partie du dessin, aux petites composantes connexes. Des traitements complexes et variés peuvent alors être réalisés pour, d'une part supprimer les composantes ne correspondant pas à des caractères (pointillés, ...), et d'autre part localiser les caractères manquants. Desseilligny [Desseilligny94] [Desseilligny95] étudie très précisément tous les voisinages des composantes connexes de petites tailles pour les regrouper en mots. Pour cela il utilise des critères de taille, de voisinage, d'orientation et effectue même une reconnaissance de la police.

Des techniques ont été développées pour extraire les caractères attachés aux graphiques en minimisant à la fois l'altération du caractère et celle du graphique [Joseph91][Boatto92].

Analyse syntaxique

Intervient ensuite, durant une *phase d'analyse syntaxique*, les algorithmes qui effectuent une étude des données fournies par le niveau inférieur : segmentation (regroupement/séparation), calcul d'attributs, ... en vue de la reconnaissance. Pour cette phase, les travaux réalisés sur les schémas électriques sont à rapprocher de ceux effectués sur les plans cadastraux. La localisation des entités à reconnaître se base sur l'étude des vecteurs (issus de la vectorisation) : recherche des boucles fermées (polygones) et sur l'étude de la position des zones de texte. Pour traduire les relations entre les primitives de bas niveau, Antoine [Antoine92] préconise l'utilisation de modèles hiérarchiques d'objets. Après la localisation, de nombreuses méthodes à base de graphes structurels permettent de mettre en correspondance les modèles d'une base de données avec les formes détectées [Kuner86, Lee92]. Il se pose malheureusement rapidement des problèmes d'explosion combinatoire. L'utilisation d'heuristiques permet, dans certains cas, de résoudre ces problèmes par génération d'hypothèses (boucles fermées = symboles dans les schémas électriques) [Habacha93b] [Okazaki88].

Lai [Lai91] propose d'utiliser le squelette à partir de ses points particuliers (jonctions, liens, extrémités) pour localiser et caractériser les lignes pointillées. Les résultats obtenus sont intéressants et permettent notamment la localisation des courbes aussi bien que des droites.

Cette grande variété des méthodes ne fait que refléter les difficultés du problème.

1.2.3. Les problèmes qui subsistent

Même si ces travaux ont permis, comme on vient de le voir, de nombreuses avancées en ce qui concerne les traitements de bas niveau et l'extraction de primitives, de nombreuses difficultés subsistent encore :

- Les formes pleines ainsi que certaines jonctions ou extrémités sont altérées durant la squelettisation (préalable à la vectorisation).

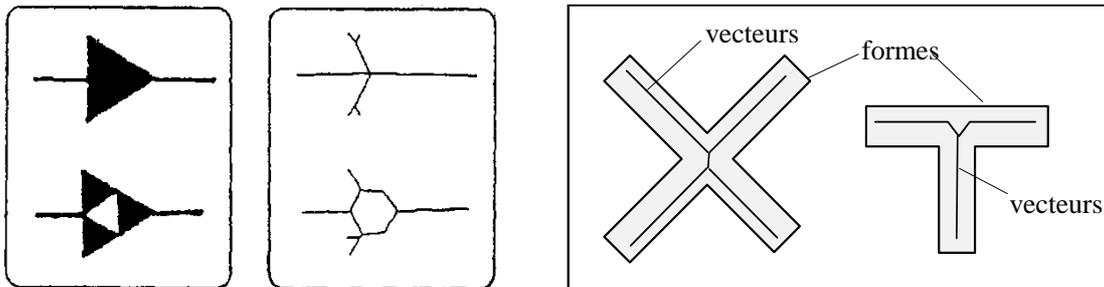


Figure 1.11 : Problèmes lors de la squelettisation

- Quelle que soit la technique de vectorisation, le résultat de l'approximation polygonale n'est pas toujours celui escompté. Les points critiques (points de contrôle) fournis ne sont pas toujours représentatifs de la forme initiale comme on peut le voir sur la figure suivante.

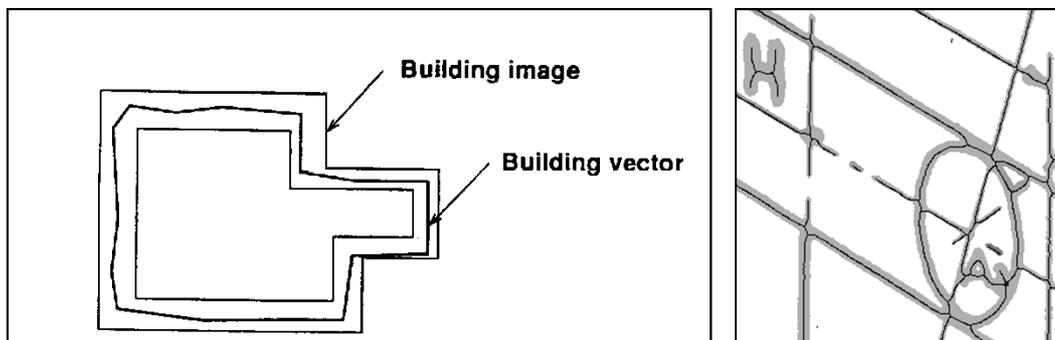


Figure 1.12 : Problèmes lors de l'approximation

- La séparation des caractères du graphique pose des difficultés lorsque la taille des composantes connexes n'est plus significative : c'est à dire quand un caractère est connecté à une autre partie du dessin ou quand il s'agit de manuscrit.

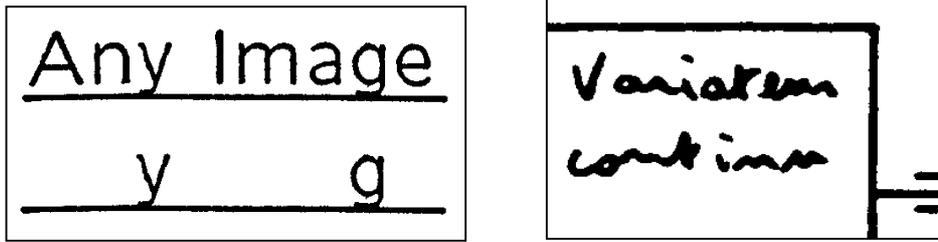


Figure 1.13 : Problèmes lors de la localisation

- Les caractères isolés sont également très difficilement localisables et sont souvent confondus avec d'autres petits symboles. Les pointillés ne doivent pas être confondus avec les caractères.

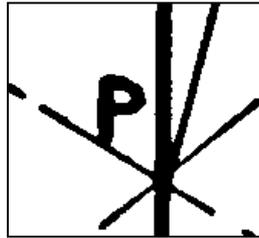


Figure 1.14 : Problème lors de la localisation

- Peu d'articles s'étendent sur l'extraction et le codage des courbes, cercles et autres formes non rectilignes si nombreuses dans le manuscrit ainsi que sur leur mise en relation avec les autres parties du document (liaison entre une droite et une courbe).

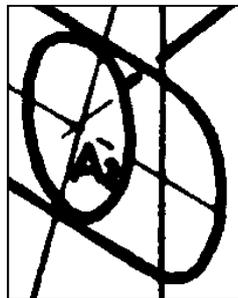


Figure 1.15 : Problème de représentation

De même, une fois les primitives de base extraites, les systèmes d'interprétation de documents complexes se heurtent à d'autres difficultés.

- Après la séparation des différentes couches que comporte un schéma, il devient très difficile d'effectuer la reconstruction et la reconnaissance d'entités de haut niveau puisque leurs composants peuvent être dispersés dans plusieurs des couches. Les représentations engendrées sont très différentes les unes des autres, la comparaison et la mise en correspondance des résultats obtenus pour chaque couche deviennent quasiment impossibles. La plupart des réalisations ne fournissent pas des solutions satisfaisantes; les recherches s'interrompent donc à ce niveau en raison d'une interprétation isolée de chacune des couches.

Pourtant, beaucoup de travaux [Antoine92,Bélaïd92] indiquent que cette segmentation en couches doit obligatoirement être réalisée avant toute autre analyse puisque les traitements d'extraction sont spécifiques à chaque type de forme qu'il est possible de rencontrer dans de tels documents; ils estiment qu'aucun traitement universel n'est applicable.

- Il me semble que les retours en arrière sont pratiquement irréalisables sans remise en cause de la totalité de l'analyse. Les travaux de Ogier [Ogier94b] sur l'analyse de la cohérence vont aussi dans ce sens; les vérifications sont réalisées après chaque construction d'objet élément du cadastre pour éviter la propagation des erreurs.

1.3. Vers de nouvelles méthodes

1.3.1. Interprétation se basant sur la structure logique

Toutes ces difficultés montrent la nécessité d'aborder le problème sous un nouvel angle. Les techniques utilisées dans le domaine très voisin de la lecture automatique de partitions musicales semblent instructives [Ramel93, Ramel94]. Là aussi un problème de segmentation complique l'interprétation puisque les symboles musicaux doivent être séparés des lignes de portées (figure 1.16).

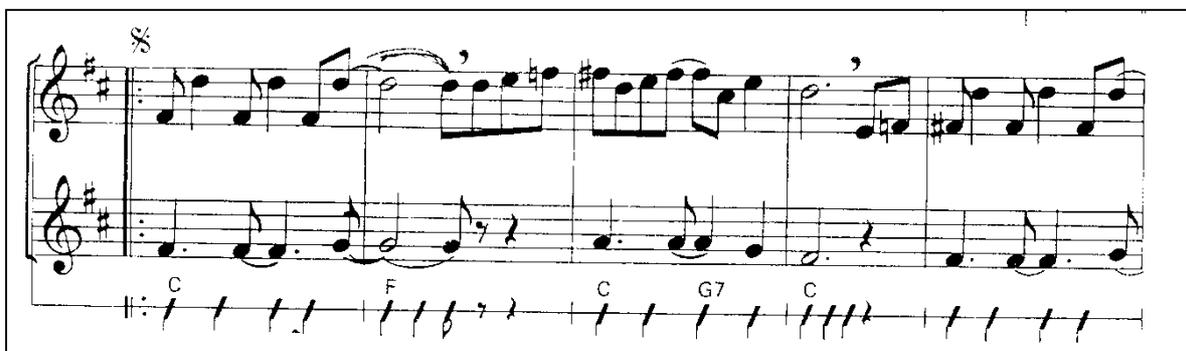


Figure 1.16 : Extrait de partition musicale

Même si, contrairement aux documents auxquels on s'intéresse, la structure logique d'une partition musicale est connue a priori, un parallèle peut être mené. Poulain [Poulain94] se base sur des critères géométriques pour considérer une partition comme la superposition de 3 couches : une pour les lignes horizontales, une pour les lignes verticales; les autres formes constituant la troisième couche. Fahmy [Fahmy91] distingue, lui, 2 couches en rapport avec la sémantique musicale : la couche lignes de portées et la couche symboles musicaux. Ce dernier préconise, de plus, de justifier chaque étape de l'analyse par une ou plusieurs règles tirées de la syntaxe musicale. Depuis, de nombreuses équipes se sont penchées sur le problème de traduction de la syntaxe musicale sous forme de grammaires [Baumann95].

De même, les techniques d'interprétation de documents techniques ont évolué et la décomposition en couches n'a plus été réalisée en fonction de critères géométriques (couches des traits forts, couches des traits fins, ...), c'est à dire des traitements à mettre en place, mais plutôt sur des *critères sémantiques*, c'est à dire en fonction du sens des informations fournies par les regroupements judicieux des formes de l'image, en utilisant sa structure logique.

Il est à noter qu'avec ce principe, seules les méthodes d'extraction des primitives de base et les modèles de description des objets constituant les dessins peuvent être communs à tous les systèmes d'interprétation. L'ordre d'extraction des entités dépend lui du domaine d'application. Ainsi, il est possible de décrire la structure logique des différents types de documents :

- Pour les documents de type projections orthogonales, on distingue généralement une couche Cotation (Texte + lignes de référence + flèches), une couche Contours visibles, une couche Contours cachés, une couche Axes de symétrie, ...

- Pour les schémas de principe (et les schémas cinématiques) on distingue le plus souvent une couche Légendes/cotations, une couche Symboles, et une couche Liens de connexion.

- Pour les documents de type plans, les couches à extraire sont, par exemple, la couche Réseaux routiers, la couche Immeubles, la couche Canalisations, ... pour obtenir des cartes dites "géocodées".

Ces exemples illustrent la difficulté de définir une *structure logique* pouvant être commune aux différents types de documents; celle-ci varie énormément suivant le type d'application et même le résultat escompté. La stratégie d'interprétation doit cependant se baser sur ces connaissances et, de plus, les exploiter dès le bas niveau, aussi bien que dans tous les niveaux de l'analyse, afin d'employer des outils spécifiques et adaptés durant tout le traitement.

C'est en tenant compte de ce type d'information qu'il faut définir, par exemple, l'ordre d'extraction des différentes entités constituant un dessin, ou essayer de lever certaines ambiguïtés concernant l'appartenance d'une primitive à une entité plutôt qu'à une autre. On limite ainsi les incohérences et leur propagation dans les niveaux supérieurs.

1.3.2. Méthodes syntaxiques et méthodes sémantiques

Les différents documents obéissent tous à des règles strictes, définies par différentes normes, qui précisent quelles entités et quelles relations entre entités sont susceptibles d'apparaître. On ne parle plus alors de structure mais plutôt de syntaxe propre au schéma. Ces règles varient selon le type de documents et d'applications auxquels on s'intéresse : les bases de données d'un système de CAO mécanique sont totalement différentes de celles d'un système de simulations de circuits électriques ou même de celles d'un système de CAO d'un autre type. La connaissance de ces différentes règles permet de vérifier la cohérence de la représentation obtenue.

De nombreuses méthodes de reconnaissance (structurelles, syntaxiques, statistiques, ...) peuvent être appliquées. Cependant, il est préférable de leur adjoindre des programmes d'Intelligence Artificielle, fondés sur des bases de connaissances spécifiques au domaine considéré. *L'analyse sémantique* peut en effet ainsi venir en aide à l'analyse syntaxique.

Le système ANON, développé au sein de l'équipe de Joseph [Joseph92] utilise une grammaire de type LR1 pour gérer les connaissances. Les règles de cette dernière ont pour but de contrôler l'activation des mécanismes de détection des primitives de bas niveau. Dans le même esprit, Habacha [Habacha93a] définit des liens contextuels entre le texte, les symboles, et les liaisons constituant les schémas électriques. La détection d'une entité peut alors déclencher des recherches d'entités d'un autre type dans son voisinage.

De la même façon, Cugini s'intéresse aux aspects mécaniques des dessins et essaie d'extraire des «form-features» [Brun94] par reconstruction à partir de primitives de base [Cugini88, Cugini90, Bottoni92].

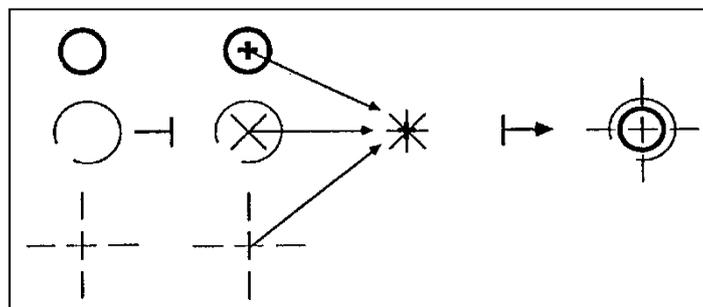


Figure 1.17 : Construction d'une forme caractéristique

Pour cela, son équipe a défini un langage de description des formes caractéristiques et l'a traduit sous forme de grammaire (figure 1.17).

En France, l'équipe de Tombre travaille depuis de nombreuses années sur l'élaboration d'un prototype de système d'interprétation automatique de dessins techniques. Une des premières versions, nommée Cellestin III, avait pour but de localiser et reconstruire les blocs fonctionnels (vis, engrenages, ...) constituant le mécanisme (figure 1.18) [Tombre91] [Vaxivière92].

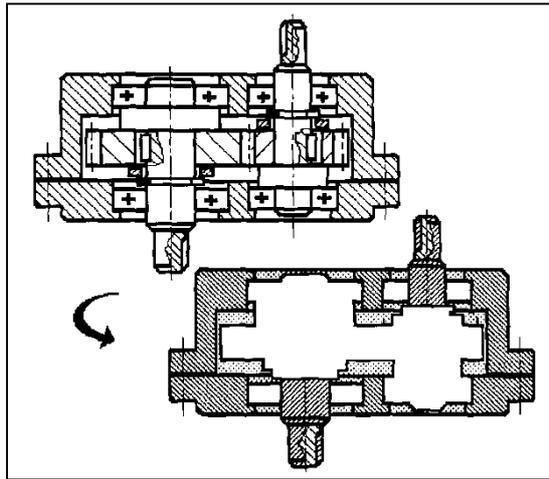


Figure 1.18 : Blocs fonctionnels

Le système a depuis évolué pour donner naissance à Cellestin IV. Les connaissances utilisées concernent alors non seulement les entités (cohérence interne de l'objet) et les liens entre entités (cohérence externe ou globale), mais aussi la sémantique du domaine puisque les dessins représentent des objets du monde réel, régis par des lois fondamentales venant de la mécanique, de la physique, ...

Ces connaissances sont stockées à l'aide d'un blackboard (cf paragraphe 1.4) qui est utilisé et maintenu à jour par différents spécialistes. La coordination des spécialistes se base sur des règles d'activation. Leur rôle est d'analyser les symétries (recherche des blocs fonctionnels), de vérifier la possibilité d'assemblage/désassemblage du mécanisme et d'analyser la cohérence cinématique du mécanisme [Vaxivière94b, Vaxivière94c, Collin93, Dori95].

Dans tous ces systèmes, la gestion des connaissances est très complexe et de ce fait, les dessins étudiés restent assez simples. Certains chercheurs ont alors choisi d'étudier plus précisément les cotations et légendes se trouvant sur les projections orthogonales [Lai93]. Ces données constituent les éléments d'un véritable langage [Dori93], il peut alors être opportun de vérifier que leur regroupement forme un tout cohérent, concorde avec les autres informations extraites, et possède un sens non seulement formel mais aussi physique.

Dans le cadre du développement du système MDUS, l'équipe de Dori s'est longuement étendue sur l'analyse syntaxique des chaînes de cotations dans les projections orthogonales. Ce qui a abouti à la réalisation d'une méthode d'extraction, d'analyse et de vérification de cohérence des cotations (norme ANSI) modélisées à l'aide de graphes [Dori91, Dori92, Wang92].

Comme la plupart des systèmes, celui-ci commence par détecter les têtes de flèches pour localiser les cotations. Une fois la zone d'étude définie, la cotation entière est reconstruite par extractions successives de ses constituants (texte, lignes de référence, ...). Les différentes dispositions possibles des constituants sont généralement décrites à l'aide de grammaires attribuées. Les travaux de Collin prouvent que l'utilisation d'une plex-grammaire (figure 1.19) est bien adaptée à ce problème [Collin91, Collin92].

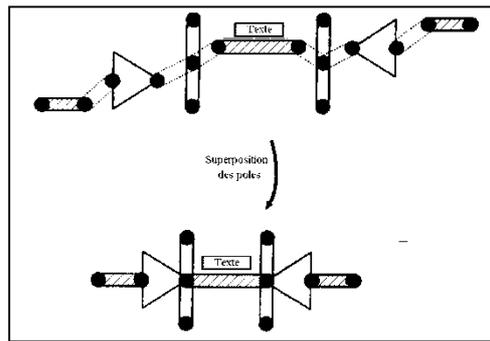


Figure 1.19 : Grammaire des cotations (de Collin)

Parallèlement à tous ces travaux, des études sur la reconstruction 3D à partir des vues 2D ont aussi été menées afin d'analyser la cohérence des données extraites, mais aussi afin d'atteindre le but initial de la lecture de projections orthogonales !

Il ne s'agit pour l'instant que de prémisses car la tâche est difficile. Les travaux se basent sur des données reflétant la compréhension totale et sans erreur des vues 2D; de plus; les dessins étudiés sont très rudimentaires.

A partir de points particuliers, supposés représenter les sommets du volume, les arêtes sont reconstruites puis combinées pour former des faces qui à leur tour seront regroupées pour constituer le volume complet [Lysac91, Wang93, AhSoon95]. Il est également possible de chercher à localiser des configurations particulières (composantes 3D) sur les différentes vues 2D qu'il faut ensuite combiner par géométrie constructive [Marston95, Marti91].

Weiss (de l'équipe de Dori) tente d'utiliser les chaînes de cotations pour effectuer la reconstruction 3D par fusion des chaînes de cotes des différentes vues 2D [Weiss95].

1.3.3. Les problèmes qui subsistent

Tous ces exemples sont intéressants, mais il faudrait, aujourd'hui, se donner les moyens d'étendre ce type de techniques à l'ensemble des éléments qui constituent le dessin. Ces travaux apportent des réponses diverses aux questions posées et suscitent les remarques suivantes :

- La «volonté» d'effectuer une analyse ascendante des documents techniques (extraction de primitives puis reconstruction) augmente les difficultés auxquelles sont confrontés les systèmes d'interprétation.
- La détection ou la reconstruction d'entités possédant une sémantique est une tâche difficile, elle nécessite des connaissances précises sur le document. La difficulté se situe principalement au niveau de l'amorçage du processus lorsque le contexte sur lequel on se base est faible (recherche des boucles fermées, des têtes de flèches, des hachures, ...).
- La soustraction des entités reconnues (hachures, textes, ...) de l'image initiale ou de son modèle de représentation pose encore beaucoup de problèmes. Ces entités rendent pourtant plus difficile la détection des autres données.
- La prise en compte des données extraites est ardue : la gestion du contexte doit permettre de générer des hypothèses c'est-à-dire de faciliter l'utilisation des connaissances obtenues et d'augmenter les possibilités de coopération entre les processus d'analyse. Les techniques de reconnaissance pure deviennent alors insuffisantes et doivent être complétées. L'adjonction d'artifices doit permettre la prise en compte du *contexte* de l'image et du *but* poursuivi.

1.4. Intelligence artificielle distribuée

Les résultats obtenus en analyse d'images de documents techniques soulèvent de nombreuses interrogations. Pour essayer d'y répondre Nagy [Nagy95] propose de mettre de côté les recherches concernant l'extraction de primitives (traits, hachures, caractères, ...) pour lesquelles on obtient maintenant des résultats satisfaisants, et de s'intéresser plus spécialement aux aspects coopération entre processus d'analyse des formes qui posent encore de nombreux problèmes. Nous allons rappeler ici sommairement l'état de l'art concernant les méthodes de gestion des connaissances et celles permettant la coopération entre processus utilisables en reconnaissance de formes ou en vision par ordinateur.

Le but de l'intelligence artificielle est l'acquisition par l'ordinateur de certaines compétences d'un être humain, et de leur utilisation dans l'accomplissement d'une tâche nécessitant à la fois des connaissances, de l'expérience, et une certaine capacité de raisonnement. Dans la réalité, les individus travaillent généralement en groupe, mettant en commun des savoirs et des savoir-faire afin de collaborer à la réalisation d'un même but. Il est donc naturel qu'une approche IA reproduise le fonctionnement de cette multiplicité d'experts co-oeuvrant : on parle alors de l'Intelligence Artificielle Distribuée. Ces méthodes de résolution distribuée des problèmes conduisent à la définition de nouveaux domaines d'investigations.

Un système de **résolution distribuée** est conçu comme une communauté de systèmes travaillant de manière coordonnée ce qui pose alors de nouveaux problèmes théoriques : modélisation de la communication entre experts, nature de la structure de contrôle, type d'organisation des données et des traitements (hiérarchiques ou non), ...

Dans les **systèmes multi-agents**, chaque expert (agent) dispose d'une connaissance plus spécialisée mais parcellaire et utilise un mécanisme de raisonnement relativement simplifié.

Cette simplification vient cependant compenser la complexification des interactions entre agents pendant la phase de résolution. Chaque expert manipule généralement des données incomplètes du fait de sa vision partielle de l'environnement. Des données incertaines, fondées sur des informations variables propres à chaque expert, doivent être manipulées à un niveau local aussi bien qu'à un niveau global. Des données incorrectes peuvent être fournies par un agent, aussi est-il nécessaire que ces données soient reconnues comme fausses à un niveau global.

1.4.1. Organisation des données et des traitements

Une première méthode, dite **hiérarchique**, représente la tâche à accomplir par une décomposition hiérarchique de la tâche globale en une succession d'étapes élémentaires. La représentation d'un niveau supérieur est une simplification, une abstraction, du niveau à partir duquel sont extraites les informations.

Cette représentation n'est pas satisfaisante dès lors que l'on souhaite modéliser les processus de résolution humains. En effet le cerveau adopte une méthode plus **opportuniste** de résolution des problèmes; celle-ci suppose une représentation non hiérarchique des données et des opérations de traitement.

Pour se rapprocher de ce fonctionnement, des systèmes ont été développés permettant la co-exploitation d'une ou plusieurs ressources. L'interaction exploite alors la possibilité de mettre en jeu plusieurs points de vue sur une donnée ou sur la façon de la traiter : on parle de fusion dans le premier cas et de compétition dans le second. Ces deux procédés sont naturellement perçus comme complémentaires ; il s'agit dans les deux cas de multiplier les points de vue, qu'ils soient descriptifs ou opératoires, pour enrichir l'analyse et en exploiter les conflits éventuels. L'intérêt de ces approches est de gérer explicitement et de manière rationnelle les conflits de décision et la multiplicité des représentations.

La **compétition** de compétences consiste à produire des solutions (descriptions ou décisions) qui seront analysées par des techniques de **fusion**. Le principe est d'appliquer différents opérateurs sur un ensemble de données, puis de conserver le meilleur résultat ou un compromis des résultats obtenus.

L'accès concurrent à une ressource peut être simulé par des processus interruptibles gérés par un séquenceur, selon cette approche l'exécution des tâches est réellement entrelacée et il devient possible de bénéficier d'une réelle interaction entre les traitements (compétition d'accès).

1.4.2. Modèles de communication entre agents

Les agents ont besoin d'échanger des informations. On distingue 2 modèles de communication correspondant à 2 types d'architectures :

- communication par messages
- communication par partage d'information

Les modèles multi-agents fondés sur la distribution totale aux différents agents, des connaissances, des méthodes utilisées, et des résultats, nécessitent un modèle de communication par message. Un agent ne peut alors manipuler que sa base de connaissance locale et envoyer des messages aux agents dont il connaît l'existence. Une hiérarchie dans les agents doit donc définir les modalités de communication ; ceci suppose que l'on puisse isoler des îlots de sous-problèmes.

Les systèmes de résolution opportunistes supposent que les agents travaillent sur un même espace contenant les données contextuelles ainsi que les états partiels du problème en cours de résolution. La communication est réalisée par le partage de l'information, rassemblée dans un espace commun.

1.4.3. Structure de contrôle

Tout système multi-agents impose l'existence d'une forme de contrôle. Deux modes de contrôle peuvent être adoptés pour régir la coopération entre tâches (qui peuvent être implantées comme des processus, des procédures, ou des agents) :

- Le mode centralisé, selon lequel un superviseur décide de la tâche à entreprendre au vu de l'état courant de la résolution (souvent représenté dans un Blackboard). L'organisation est alors hiérarchique, l'agent superviseur peut orienter les recherches pour aboutir à la solution par un mécanisme de focalisation de l'attention.
- Le mode décentralisé dans lequel le contrôle est distribué ; chaque agent possède des informations sur les compétences des autres agents ainsi qu'un plan de résolution personnelle permettant de guider la résolution.

Dans les deux cas, l'accent est mis sur le caractère incrémental et constructif de la stratégie de résolution, chaque spécialiste apportant un complément d'information susceptible d'être exploité par un autre.

Le principal problème réside alors dans l'adaptation, c'est à dire dans la capacité de sélectionner et d'ajuster les opérateurs ou traitements de manière optimale étant donné un contexte de résolution : couples situation - but ou système - environnement.

1.4.4. Le modèle Blackboard

Description

Un système **Blackboard** peut être considéré comme une méthode de résolution de problème par construction **incrémentale** de la solution (évolution pas à pas vers la solution par résolution de sous-problèmes).

Il s'agit de décrire l'état général d'un **processus coopératif** à un ensemble de **spécialistes** en train de résoudre un problème spécifique. Chaque spécialiste regarde le **tableau noir** et est attentif à l'information qui pourrait l'intéresser (c'est-à-dire relative à son domaine). Aussitôt qu'une telle information apparaît, il exécute le travail prévu et transmet les résultats aux autres spécialistes. Une fois le travail terminé, chaque spécialiste écrit les résultats sur le tableau noir pour que les autres prennent connaissance des résultats.

Ainsi, aucun spécialiste n'a besoin de connaître l'existence des autres spécialistes. Ceci permet une intégration facile, à la fois des méthodes utilisant une approche d'analyse « bottom-up » avec celles utilisant l'approche « top-bottom ». De même, l'analyse peut aussi bien être dirigée par les données que par les buts à atteindre.

Architecture

On peut généralement décomposer une application Blackboard en 3 modules :

- **L'espace solution** : Il s'agit du tableau noir lui-même. Il décrit l'état actuel du problème en cours de résolution. Ce dernier est souvent constitué d'une organisation hiérarchique d'objets. Il peut être divisé en différents ensembles de données.

- **Les sources de connaissances** : Il s'agit des spécialistes eux-mêmes, ainsi que de l'ensemble des connaissances utilisées par les experts. Elles sont souvent codées sous forme de **règles** : SI condition ALORS action

Pour exécuter le travail chaque spécialiste utilise des procédures de traitement. Chaque action est soumise à une condition.

- **Le contrôle (ou le directeur)** : Son but est d'optimiser la résolution du problème. Il supervise l'action des spécialistes en assurant une bonne coopération. Il dirige le focus d'attention des spécialistes et gère l'enchaînement des travaux : c'est lui qui choisit la règle à exécuter parmi toutes les règles activables à un instant donné.

Problèmes

Les principaux problèmes résident dans l'enchaînement des différents travaux (activation des spécialistes) :

- Il se produit parfois des divergences ou **contradictions** entre résultats fournis par 2 spécialistes différents.
- Il peut aussi se produire des **silences** pendant lesquels plus aucune règle n'est applicable.

Le Blackboard est une entité purement **conceptuelle** car rien n'est défini sur la façon de le mettre en place informatiquement.

1.5. Conclusion

Avant de conclure, je pense qu'il est nécessaire de donner un aperçu des quelques systèmes commerciaux qui ont été développés et mis sur le marché. La plupart de ces systèmes permettent de numériser le document, de pré-traiter le document par un ensemble de filtres pour éliminer le bruit, et d'effectuer une partie des opérations suivantes :

- séparation texte/graphiques
- description vectorielle de la partie graphique du dessin (vecteurs, arcs, cercles)
- reconnaissance de primitives graphiques plus évoluées : lignes tiretées, hachures, flèches, ...
- reconnaissance des symboles et notamment du texte, ...

Malgré quelques succès relatifs, ces systèmes sont loin de répondre aux besoins exprimés. Les principales causes de ce rejet ont été mises en évidence dans ce chapitre. Parmi toutes les remarques faites au cours de cet état de l'art, il me semble ressortir principalement :

- que les problèmes d'interprétation proviennent des nombreux niveaux de représentation nécessaires pour passer de l'image à une interprétation finale : phase lexicale, puis syntaxique, puis sémantique. Il me paraît donc indispensable d'effectuer des recherches susceptibles de contribuer à la découverte d'une représentation simple et efficace des documents techniques et surtout qui soit capable de supporter toutes les phases de l'analyse (chapitre 2). A mon avis, les points cruciaux concernent :

- le type d'information à extraire de l'image,
- la manière d'extraire les informations ainsi que leur gestion,
- l'intégration de connaissances sur le domaine durant les analyses,
- la coopération entre les différentes analyses (lexicale, syntaxique et sémantique).

- que c'est en utilisant au maximum les informations obtenues au cours des divers traitements, c'est-à-dire en tenant compte d'un contexte le plus général possible, que l'on parviendra à une interprétation globale des documents.

Nous allons proposer une technique permettant la description de l'image dans sa globalité et fournissant une représentation sur laquelle puissent se baser différents processus capables de réaliser à la fois l'extraction, l'analyse et l'interprétation (chapitre 3).

- qu'il n'est pas possible de limiter à deux phases successives le processus d'interprétation de tels documents puisque même l'homme n'est pas capable d'un tel exploit ! Il a besoin de voir un document à plusieurs niveaux pour l'appréhender totalement. Nous essaierons de reproduire ce phénomène (qui traduit d'ailleurs la nécessité de la présence d'un contexte pour comprendre) en privilégiant l'extraction des évidences tout en enrichissant le contexte et en limitant les erreurs quitte à augmenter le nombre de cycles nécessaires à l'interprétation.

Chapitre 2

Perception globale d'un document

2.1. Bases conceptuelles

2.1.1. Perception humaine

Le but des recherches en Reconnaissance de Formes est de réaliser des machines capables de rivaliser avec le système visuel humain. Il semble donc intéressant d'établir des correspondances entre le système artificiel que l'on crée et le système humain.

L'homme peut, grâce à un système visuel très perfectionné, reconnaître toutes sortes d'objets et comprendre sans difficulté son environnement. Les performances du système visuel humain peuvent être expliquées, d'une part, par sa faculté d'utiliser le contexte environnant et les connaissances préalablement acquises, et d'autre part par son mode de fonctionnement [Gordon89]. Selon B. Julesz [Lecas92], la vision humaine opère selon 2 modes distincts :

- ♦ la vision pré-attentive (pré-consciente) permet l'acquisition instantanée de l'image, sans scrutation, et en couvrant un large champ visuel.
- ♦ la vision attentive, limitée à de petits domaines, correspond à une scrutation pas à pas avec une attention bien marquée.

L'activité perceptive peut être représentée par une période d'oscillation entre 2 phases : la « perception » : acquisition de l'image, puis la confrontation avec l'appris (au niveau de la mémoire associative : analyse, reconnaissance et interprétation). Ce cycle perdure jusqu'à ce que la compréhension totale de la scène soit obtenue. L'homme n'utilise pas une seule « image » lorsqu'il interprète un phénomène, mais il effectue des acquisitions successives, chacune de ces focalisations ayant un objectif précis et contribuant à faire évoluer, aussi bien la représentation mentale de la scène en cours que ce qui avait été préalablement appris, jusqu'à les faire converger vers un tout cohérent [Fortin92].

Durant l'approche globale (vision pré-attentive), notre système visuel ne tient pas compte des détails. La vue élaborée durant cette phase sert de contexte par la suite. Marr parle d'élaboration d'une esquisse primaire [Marr82], par extraction d'indices visuels, permettant d'obtenir une vue d'ensemble.

Ensuite, par une approche plus locale de la scène (déplacements oculaires), en utilisant des connaissances pour cibler les recherches, sont trouvées les informations précises et pertinentes qui conforteront le raisonnement et contribueront à l'interprétation.

Durant cette phase d'exploration visuelle, le système humain focalise son attention sur des points particuliers de la scène, il est à la recherche d'informations précises, dont la présence est supposée à des endroits donnés et sur lesquelles il possède aussi des connaissances.

Il est à noter que chez l'Homme, toute l'information semble se trouver centralisée au niveau du cerveau.

2.1.2. Les documents techniques

Même si la lecture intelligente de documents techniques est une tâche difficile, qui nécessite un entraînement et se trouve par là même réservée aux spécialistes ayant les connaissances requises, tout le monde est capable d'extraire certaines informations ou caractéristiques (textures, densité fond/forme, ...) d'un dessin. Par la suite, nous appellerons ces dernières : indices. Pour obtenir ces données, aucune connaissance relative au domaine n'est nécessaire. C'est à ce niveau (entrée en jeu des connaissances relatives au domaine) que se situe la séparation entre la vision pré-attentive (perception globale du document) et la vision attentive (lecture et compréhension du document). Une étude des différentes formes contenues dans le document, de leur type, de leur fréquence d'apparition et de leur disposition, suffit généralement pour cette première caractérisation.

Ce chapitre tente donc de définir un certain nombre d'indices qui seront pris en compte lors de cette première étape. Ensuite, il sera possible de choisir en terme de quelles primitives ou paramètres « informatiques » pourront être traduits les indices, pour être facilement utilisés durant toute la phase d'interprétation.

Du point de vue de l'analyse de la forme, on remarque que les formes à décrire sont constituées par des traits et des formes pleines. Les traits ont globalement une forme allongée à côtés parallèles alors que les formes pleines sont d'aspect plus varié. Par ailleurs, il ne faut pas oublier les petits symboles et les zones de texte qui constituent les annotations, éléments qui se reconnaissent souvent à leur texture particulière.

Les dessins techniques, et plus particulièrement les schémas cinématiques, possèdent, pour les experts mécaniciens, deux caractéristiques intéressantes que nous serons amené à extraire, en priorité. Ce sont, d'une part, les directions privilégiées dans l'image et d'autre part, les épaisseurs des traits.

2.1.3 Modélisation

Muni de ces informations, il est nécessaire de choisir un modèle de représentation et la méthode de mise en place qui lui correspond. Ce modèle devra reprendre les points forts du système humain mentionnés au début de ce chapitre. Les primitives de description choisies devront être à même de décrire le document durant toute la suite de l'analyse.

L'utilisation de toutes les primitives doit permettre, dans tous les cas, la description et la restitution de toutes les informations contenues dans l'image initiale avec le moins possible de pertes. D'ailleurs tout au long du traitement qui fait évoluer l'image initiale, nous conservons intact un exemplaire de l'image initiale, ce qui permet à tout moment de procéder à des vérifications avant de prendre une décision. Evidemment les primitives utilisées doivent être plus évoluées que le *Pixel*. Le modèle de représentation que nous avons choisi est aussi plus approprié aux images spécifiques que nous avons à examiner. Le choix des primitives de description, outre son adéquation avec les formes à décrire, se fonde sur leurs aptitudes à emmagasiner les connaissances recherchées et celles que l'on a pu acquérir à chaque niveau d'abstraction.

Une forme est décrite de manière équivalente par son contour et par les pixels de la région noire qui la constituent. Les directions peuvent être obtenues par une étude locale des contours ou par une étude des voisinages variables de chaque point de la forme. Cette dernière méthode est plus lourde; le contour convient mieux à l'étude des directions. Quant à l'épaisseur, elle est surtout intéressante pour étudier les traits; elle correspond à la distance entre deux frontières du trait. Ces motifs nous ont donc conduit à choisir la description par les *contours* au moyen de la primitive *Vecteur*.

Les connaissances livrées par les primitives de description *Vecteurs* sont intéressantes mais insuffisantes car l'ensemble des vecteurs ne permet pas d'étudier l'épaisseur des traits. Pour obtenir et gérer aisément des données supplémentaires, il a fallu choisir un outil de description encore plus évolué. Motivé, par la nature des formes à extraire, nous avons choisi le *Quadrilatère*; décision renforcée par l'étude des travaux de Rosenfeld [Rosenfeld86] et de Ade [YläJääski92] concernant les représentations axiales des formes. A partir d'un ensemble de segments, ce dernier construit tous les couples (Ribbons) pouvant être générés et leur associe des attributs. L'analyse des valeurs prises par chaque attribut lui permet ensuite, d'effectuer des classements de formes.

Notre idée étant de ne mettre en correspondance que les vecteurs décrivant les frontières d'une même forme (plus particulièrement des traits), nous avons choisi de réaliser cette étape d'appariement de manière à ne construire les Quadrilatères que dans certaines conditions, celles qui correspondent à la définition d'un trait.

Pour mettre en place un modèle appliquant ces principes, nous utilisons une organisation hiérarchique des primitives de description (figure 2.1). A chaque primitive de représentation correspond une classe d'objets. La gestion des connaissances, la vérification de la cohérence des résultats et la coexistence des formes se voient ainsi facilitées tout en permettant, à un niveau de la hiérarchie, d'utiliser toutes les informations exprimées dans les niveaux inférieurs.

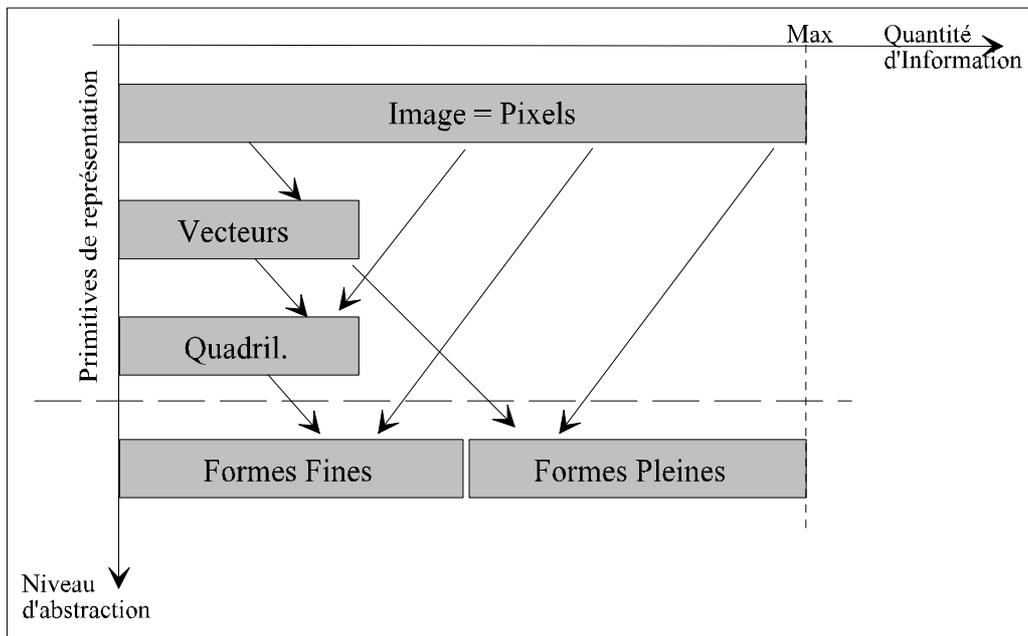


Figure 2.1 : Hiérarchie des classes d'objets

Pour gérer les connaissances qui lui sont associées, chaque classe d'objets possède des attributs et une méthode de construction. Au bas de la hiérarchie se trouvent les objets *Pixels* qui constituent l'information de base. La progression dans la hiérarchie permet, à chaque niveau, d'acquérir des connaissances plus précises sur le contenu de l'image (figure 2.1).

L'ensemble des composantes connexes est aussi une source importante d'information dans le cas des documents techniques. Une technique d'extraction des composantes fait donc partie du modèle que nous avons mis en place, elle concerne la globalité du document et n'apparaît pas dans la hiérarchie de la figure 2.1.

La période durant laquelle sont construits les vecteurs et les quadrilatères, éléments de base de notre modèle de représentation et qui permettent d'obtenir une première ossature du document, constitue cette phase d'approche globale. A l'aide de cette représentation de base, on ne distingue que les formes pleines des formes fines, les directions et épaisseurs des formes fines, et leur disposition dans l'image. L'ensemble de ces primitives élémentaires constitue cependant une source très importante d'information et sera utilisé comme base du contexte tout au long de l'analyse.

2.2. Le modèle de description

2.2.1. Construction des Vecteurs

Notre modèle de représentation se base, au départ, sur le contour des formes contenues dans l'image initiale. Cette information nous a semblé la plus robuste et la plus adaptée à la description des documents techniques codés sous forme d'images binaires. De cette manière, les formes fines aussi bien que les formes pleines ne subissent aucune dégradation importante.

Extraction du contour

Dans notre système, les contours sont obtenus par suivi des frontières des formes contenues dans l'image initiale. On utilise, pour cela, un algorithme proche de celui proposé par Pavlidis dans [Pavlidis82].

Cet algorithme considère toutes les composantes connexes (amas E de pixels noirs) qui constituent les formes de l'image. On désigne par P_a le pixel de départ élément du contour de E (1^{er} pixel noir rencontré lors du parcours des lignes de l'image), par P_c le pixel courant dont le voisinage 8-connexe va être examiné et par D l'orientation de la recherche conformément aux numéros de la table de voisinage donnée figure 2.2.

3	2	1
4	P_c	0
5	6	7

Figure 2.2 : Table de voisinage

Cette table définit l'ordre de recherche et de sélection du pixel à suivre pour progresser. Cette progression permet de suivre la frontière de l'amas de pixels noirs considéré (figure 2.2). Le suivi est terminé lorsque le contour est fermé, c'est à dire quand le pixel courant correspond au pixel de départ Pa. L'algorithme recherche alors un nouveau point de départ.

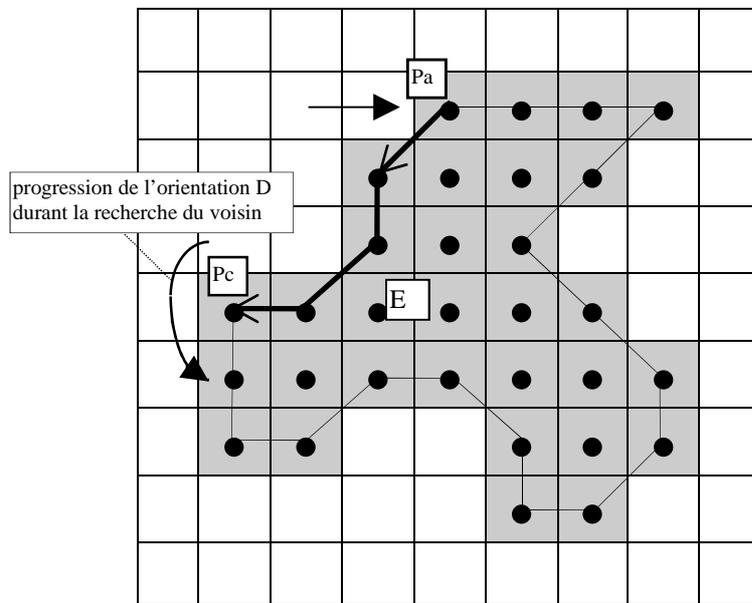


Figure 2.3 : Suivi de contour

Approximation polygonale du contour

La plupart des formes présentes dans les images possèdent des frontières qui peuvent, en presque tous les points, être approximées localement par une droite. Une approximation polygonale des contours nous permet donc de réduire la quantité de données à traiter sans trop de perte d'information.

L'algorithme utilisé est incrémental : les données sont introduites séquentiellement et traitées au fur et à mesure. Dans notre cas, il s'agit des points du contour ordonnés par le suivi. Ces points sont successivement ajoutés et regroupés jusqu'à ce qu'un certain critère, précisé dans la figure 2.4, et remis à jour à chaque ajout d'un point, ne soit plus vérifié. Les extrémités du segment sont alors le point initial et le dernier point à avoir satisfait le critère.

Nous avons choisi de reprendre la méthode d'approximation polygonale proposé par Wall [Wall84]. Cette méthode s'attache à rester dans l'espace discret de l'image et se base sur un calcul d'aire où le seuil concerne l'erreur cumulée autorisée par unité de surface du segment approximant. On détermine ainsi les sommets de la ligne polygonale : les points de contrôle.

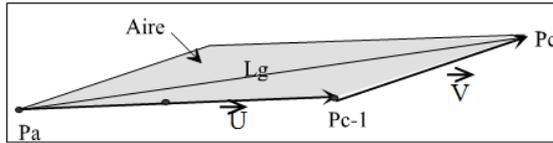


Figure 2.4 : Critère de l'approximation

$$Erreur = Erreur + \|\vec{U} \wedge \vec{V}\|$$

$$Erreur = |Erreur + |(x_{c-1} - x_a)(y_c - y_{c-1}) - (y_{c-1} - y_a)(x_c - x_{c-1})|$$

$$\text{Condition d'arrêt : } Erreur > Lg \times Epsilon$$

$$Erreur > \sqrt{(x_c - x_a)^2 + (y_c - y_a)^2} \times Epsilon$$

Les segments obtenus sont stockés dans une liste chaînée en respectant l'ordre fourni par le suivi des contours.

Ensuite, pour améliorer les performances de sélection des points de contrôle et diminuer la variabilité suivant le seuil choisi, l'étude des frontières est exécutée à différents niveaux de détails. Pour cela, la phase d'approximation polygonale est opérée de manière itérative (en utilisant toujours le même algorithme d'approximation) sur les points nouvellement obtenus. Le nombre de points de contrôle est ainsi réduit par fusion de certains des segments obtenus lors de l'étape précédente. Ce procédé est répété jusqu'à stabilisation : c'est à dire jusqu'à ce que plus aucune fusion ne soit possible. Le nombre d'itérations avant stabilisation varie généralement entre 2 et 5 suivant les caractéristiques du contour (courbure, manuscrit ou imprimé, ...).

Définition du Vecteur

Le sens du vecteur est choisi de manière à ce que la forme soit toujours placée à droite du vecteur (figure 2.5). Chaque représentant de la classe Vecteurs possède les attributs suivants :

- Pixel initial (x_i, y_i) : P1
- Pixel final (x_f, y_f) : P2
- Longueur (exprimée en pixels)
- Angle (par rapport à l'horizontale)
- Forme (référence à l'entité de niveau supérieur dont il fait partie)

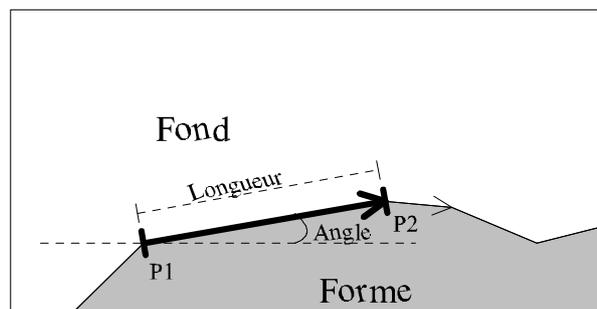


Figure 2.5 : Objet Vecteur

Résultats

Cette phase du traitement fournit, en sortie, une suite SV ordonnée de Vecteurs, définis chacun par 2 points de contrôle et représentant les frontières des formes de l'image. La figure 2.6 montre les résultats obtenus sur 2 images différentes; un grossissement de chacune des images des vecteurs est aussi fourni pour prouver la précision de cette représentation.

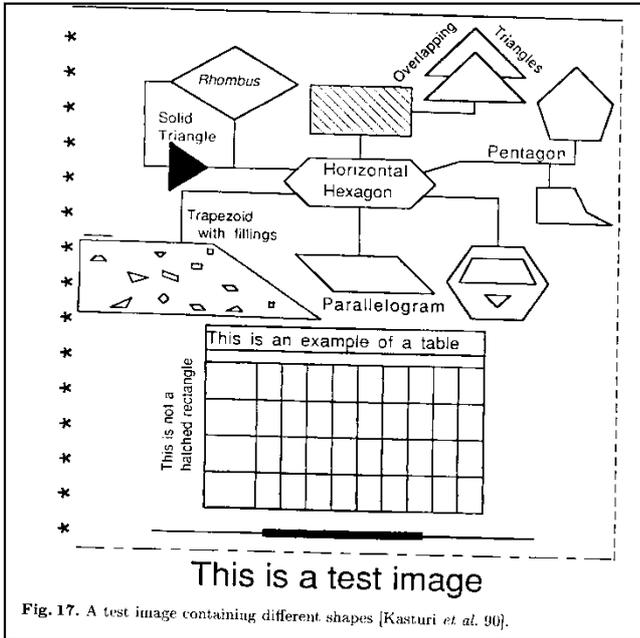


Fig. 17. A test image containing different shapes [Kasturi et al. 90].

Image initiale 1

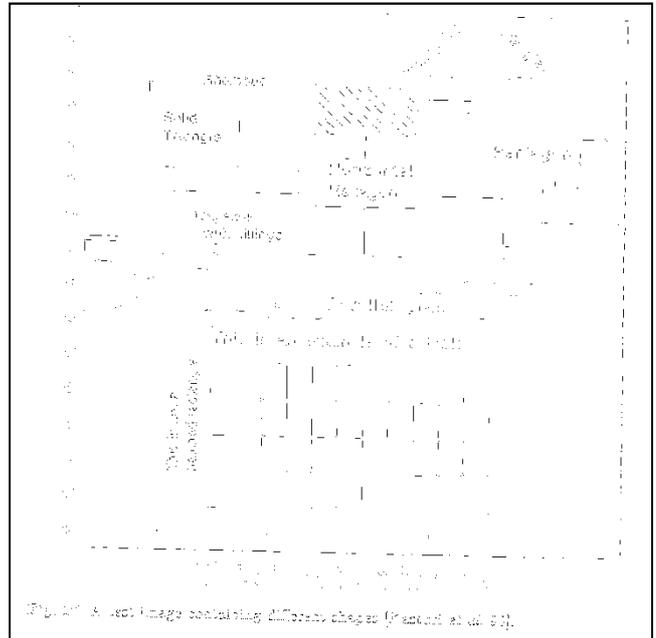


Fig. 18. A test image containing different shapes [Kasturi et al. 90].

Image des vecteurs 1

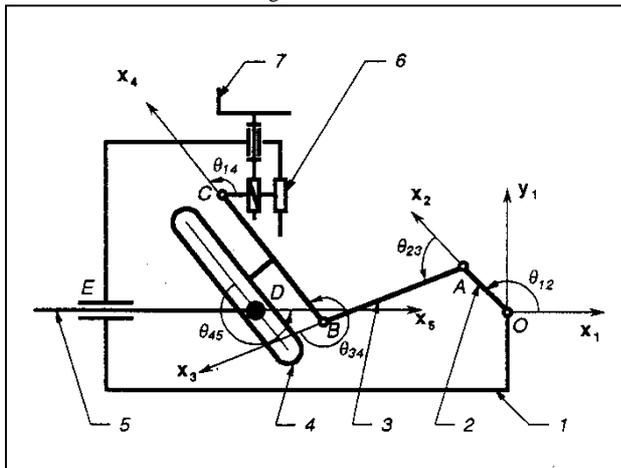


Image initiale 2

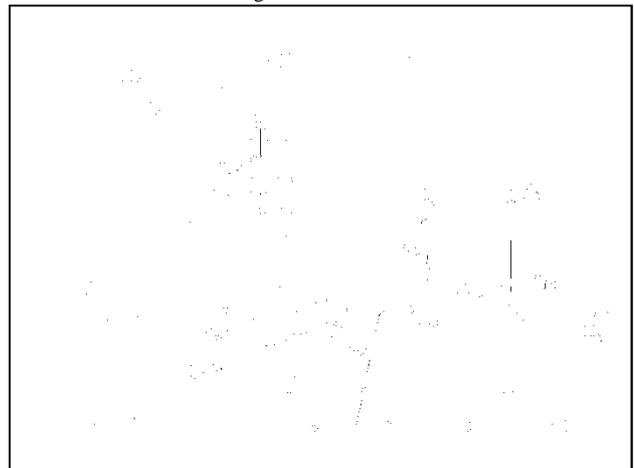
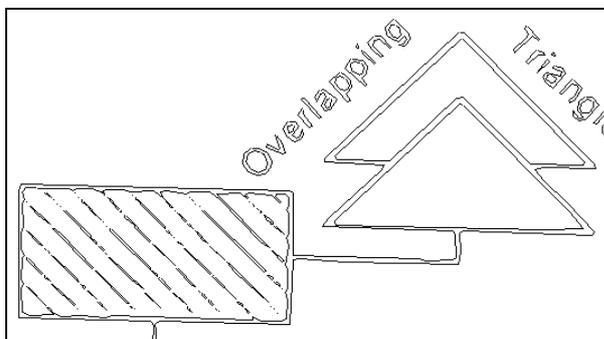
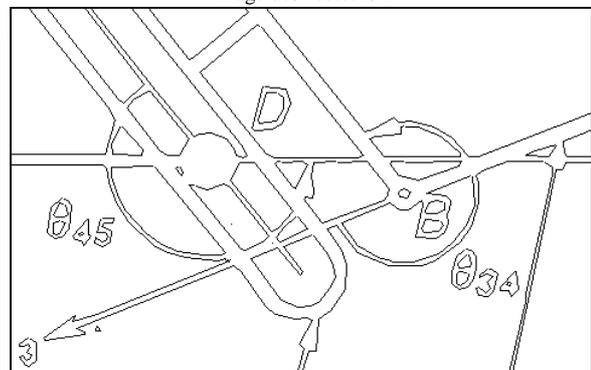


Image des vecteurs 2



Grossissement d'une partie de l'image des vecteurs 1



Grossissement d'une partie de l'image des vecteurs 2

Figure 2.6 : Images des vecteurs

La valeur du paramètre *Epsilon* servant de seuil à l'approximation a été choisie de manière empirique de façon à respecter au maximum l'information initialement fournie par l'image (*Epsilon=0,7*) notamment les courbes et arcs de cercle.

Il est à noter que la chaîne de vecteurs obtenue pour chaque contour dépend du point de départ rencontré lors du parcours des lignes de l'image. Compte-tenu des méthodes utilisées dans la suite, ce problème n'est cependant pas gênant.

2.2.2. Composantes connexes

Nous avons vu dans le chapitre précédent que les composantes connexes constituaient une source importante d'information très souvent utilisée durant le processus d'interprétation des documents techniques (notamment pour la localisation des caractères). Notre modèle de description exploite aussi cette primitive élémentaire.

La plupart des algorithmes de détection des composantes connexes effectuent un étiquetage des pixels noirs par étude des adjacences locales entre points d'un même objet. Dans le cas qui nous concerne, les paramètres intéressants sont la position et la taille des composantes connexes. Il s'agit alors plus d'obtenir les rectangles englobants correspondant à chaque composante que d'effectuer un véritable étiquetage. De plus, il est possible d'extraire ces rectangles durant la phase de suivi des contours.

Pour cela, il est nécessaire de définir la notion de contours internes et externes d'une forme. Nous appelons contours internes, les contours correspondant aux trous que contiennent certaines formes. Ainsi, alors que toute forme possède un contour externe, seulement certaines possèdent un ou plusieurs contours internes (figure 2.7).

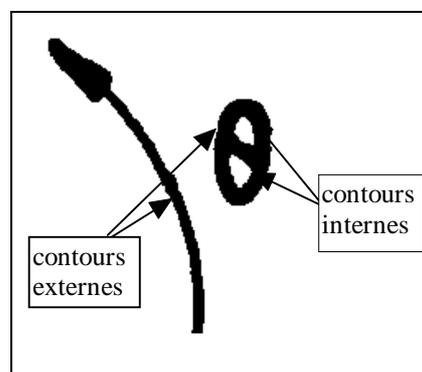


Figure 2.7 : Contours internes et externes

Détection des contours externes et internes

L'algorithme de suivi que nous avons élaboré effectue aussi bien le suivi des contours internes que des contours externes. Cependant, pour obtenir le rectangle englobant d'une composante connexe, seul l'étude du contour externe est utile. Il nous a donc fallu trouver un critère de discrimination de ces 2 types de contours. Le critère que nous avons choisi se base sur l'étude des valeurs successives correspondant aux différentes orientations (directions de Freeman) utilisées durant le suivi du contour [Rosenfeld70, Chassery91].

La somme des différences des directions successives prises lors du suivi (jusqu'au retour au point de départ) permet de discriminer les 2 types de contours (figure 2.8). On montre que cette somme S_d est égale à :

- $S_d = -8$ pour les contours internes
- $S_d = +8$ pour les contours externes

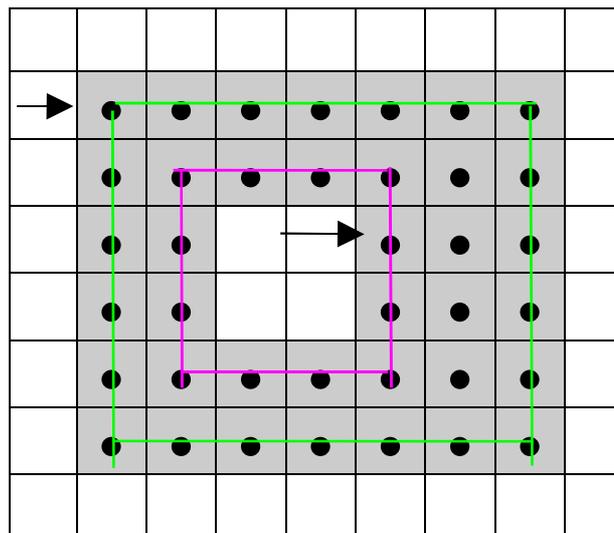
Si D_1 est l'orientation de recherche permettant d'arriver sur le futur pixel courant et D_2 l'orientation la précédant (conformément aux numéros de la table de voisinage donnée figure 2.2), alors la différence entre 2 orientations successives D_d peut être calculée (dès lors que l'on possède 2 directions successives, la première orientation du contour est soustraite à la dernière lors du bouclage) :

$$D_d = D_1 - D_2$$

$$\text{Si } (D_1 - D_2) > 4 \text{ alors } D_d = D_d - 8$$

$$\text{Si } (D_1 - D_2) < -3 \text{ alors } D_d = D_d + 8$$

Le retour-arrière ne se produisant jamais, $(D_1 - D_2) = -4$ ne se réalise jamais.



Contour interne : 664442220006 => 0-200-200-200-20 => $S_d = -8$

Contour externe : 6666600000022222444444 => 000020000200002000002 => $S_d = 8$

Figure 2.8 : Caractérisation des contours internes et externes.

Seuls les contours pour lesquels la somme S_d obtenue est positive ($=+8$) correspondent à des contours externes de composantes connexes noires. Les informations relatives à ces dernières sont stockées sous forme d'un ensemble d'objets.

Exemple de résultats

Si l'on reprend l'exemple de la figure 2.3 pour illustrer ce "théorème" alors les différentes directions utilisées durant le suivi sont : 565466010760123311444, la séquence des différences entre directions est alors : 1-1-12021-1-1-121110-20300. La somme de ces différences donne bien $S_d = 8$ ce qui correspond bien à un contour externe.

Les figures 2.9 et 2.10 présentent des exemples de résultats concernant l'extraction des composantes connexes à l'aide de la méthode décrite ci-dessus.

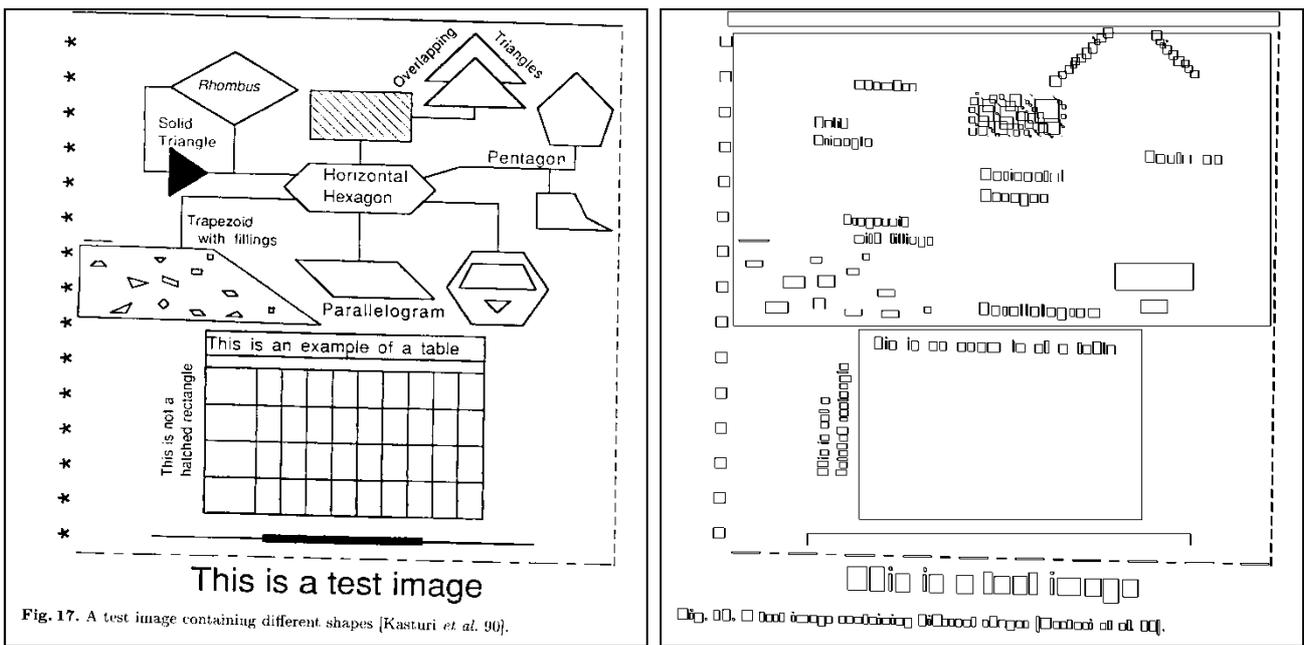


Image initiale 1 Image des composantes connexes 1

Figure 2.9 : Composantes connexes extraites (rectangles englobants)

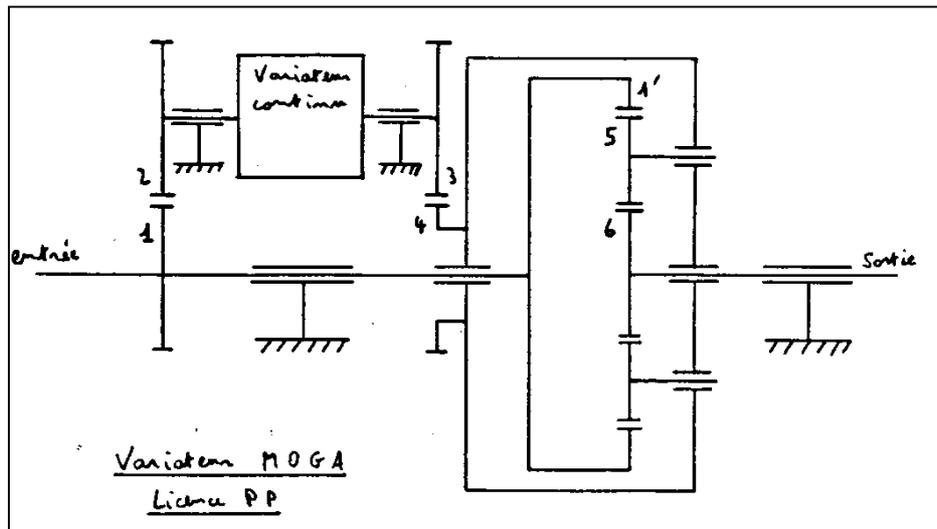


Image initiale 2

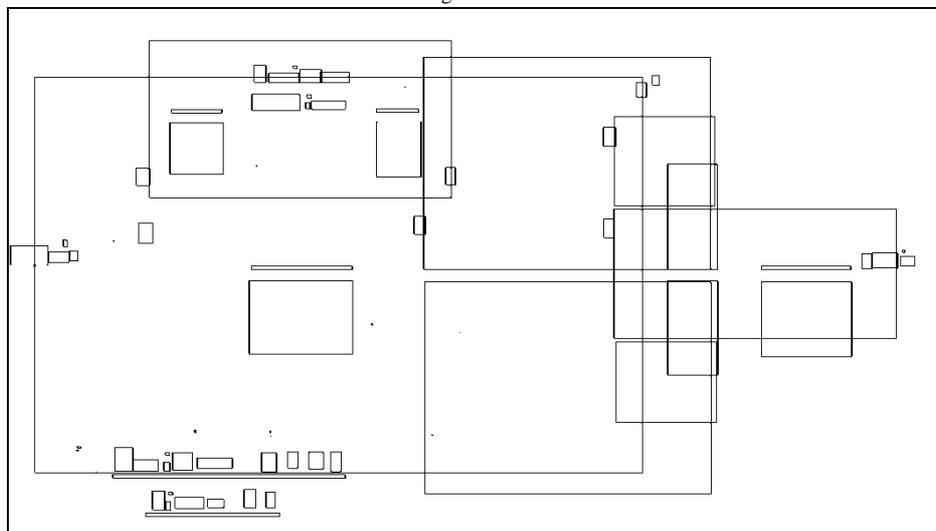


Image des composantes connexes 2

Figure 2.10 : Composantes connexes extraites (rectangles englobants)

Il est important de noter que cette méthode permet aussi d'avoir une information sur le nombre de trous contenus dans une composante connexe ou de s'intéresser aux « composantes blanches ou intérieures » contenues dans le document. Ce résultat est intéressant pour la recherche de boucles fermées comme dans l'exemple de la figure 2.11.

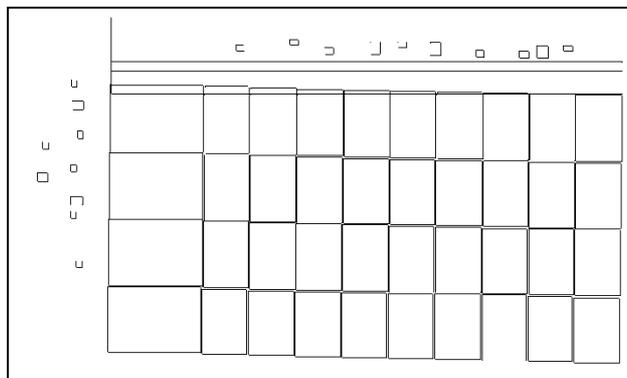


Figure 2.11 : Composantes connexes internes (rectangles englobants)
correspondant au tableau contenu dans l'organigramme (figure 2.9)

2.2.3. Construction des Quadrilatères

La représentation de l'image sous forme de chaînes de Vecteurs correspondant aux contours des formes permet de ne pas perdre d'information et donne une approximation très fidèle du dessin, mais elle n'est pas facilement exploitable et ne fournit pas suffisamment d'information sur la structure du document. La structuration supplémentaire décrite ci-dessous permet de faire évoluer ce modèle de manière hiérarchique et d'acquérir des données intéressantes.

Une étape d'appariement des *Vecteurs* de la suite SV permet la mise en place des primitives *Quadrilatères*. Chaque *Quadrilatère* est défini par une paire de *Vecteurs* (chacun appartenant à l'une des deux frontières opposées d'une Forme Fine). Cette description peut être comparée à celle utilisée par Kasturi basée sur des trapèzes [Kasturi92]. Dans notre cas, la construction des Quadrilatères s'effectue en plusieurs étapes :

- Appariement des vecteurs (éléments de SV) pour construire les Quadrilatères
- Tri des Quadrilatères suivant leur proximité de manière à reconstruire une « chronologie » du tracé
- Fusion de certains Quadrilatères voisins

Appariement

L'algorithme d'appariement commence par sélectionner le vecteur V1 de longueur maximum dans la liste des vecteurs encore non appariés, puis recherche le vecteur le plus proche (suivant la distance euclidienne entre points) de chacune des extrémités de V1 et vérifiant certains critères. Pour que la mise en correspondance de 2 vecteurs ait lieu, des conditions (critères d'appariement), correspondant chacune à une propriété physique des *Traits*, doivent être remplies. Ces critères ont été choisis, pour l'instant, "empiriquement". Une illustration est présentée figure 2.12.

Propriété du trait	Critère d'appariement
Un trait est une forme fine, ses deux bords sont proches l'un de l'autre.	Minimum de distance entre extrémités de 2 vecteurs (un vecteur ne peut être apparié qu'avec le vecteur le plus proche de lui).
Les bords d'un trait sont parallèles ou de pente très voisine.	La différence entre les directions des 2 vecteurs doit être inférieure à 30°.
Un trait comprend la zone délimitée par les 2 bords.	Présence de pixels Objets entre les extrémités (P1Q2 et P2'Q1). Les vecteurs doivent être de sens opposé.
Un trait est une forme allongée.	La longueur du 1er vecteur doit être supérieure aux épaisseurs calculées à partir des extrémités.

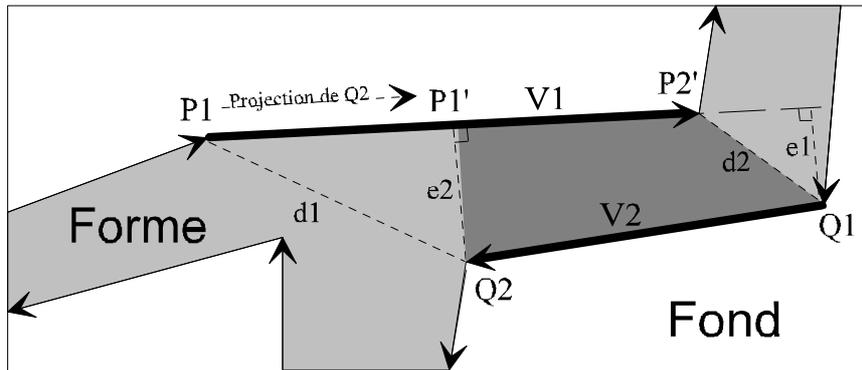


Figure 2.12 : Appariement des Vecteurs

Pour améliorer l'appariement, certains vecteurs peuvent être décomposés en 2 vecteurs fils, le point de coupure est obtenu par projection d'une extrémité d'un vecteur sur le support de l'autre (figure 2.12). Les longueurs des deux vecteurs constituant le *Quadrilatère* sont, de cette manière, plus semblables. Le second vecteur fils reste dans la liste des vecteurs à étudier. Le processus d'appariement s'arrête lorsque plus aucune mise en correspondance n'est réalisable (selon les critères d'appariement imposés).

Chacun des *Quadrilatères* de l'ensemble obtenu est défini par les 2 vecteurs qui le constituent (avec leurs attributs respectifs) et possède des attributs supplémentaires précisés sur la figure 2.13 :

- Attributs hérités du Vecteur 1
- Attributs hérités du Vecteur 2
- Epaisseurs $e1$ et $e2$
- Forme (référence à l'entité de niveau supérieur dont il fait partie)

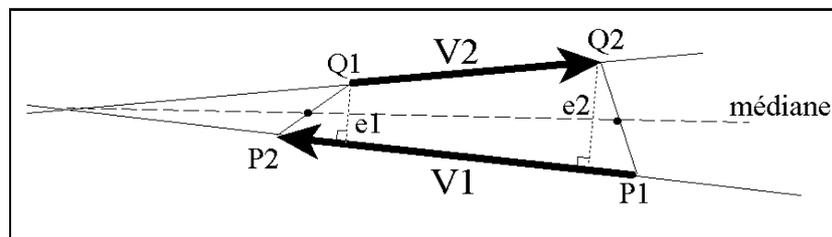


Figure 2.13 : Objet Quadrilatère

Fusion

Du fait de la localisation des points de contrôle (provenant de l'approximation polygonale) et de la méthode d'appariement, une phase de fusion des *Quadrilatères* est nécessaire pour

obtenir des objets se rapprochant au mieux des traits effectifs sur l'image (Quadrilatères de dimensions optimales et en nombre réduit). L'amélioration apportée est visible sur la figure 2.14. Cette fusion sera effectuée après une étape de tri des quadrilatères selon leur proximité.

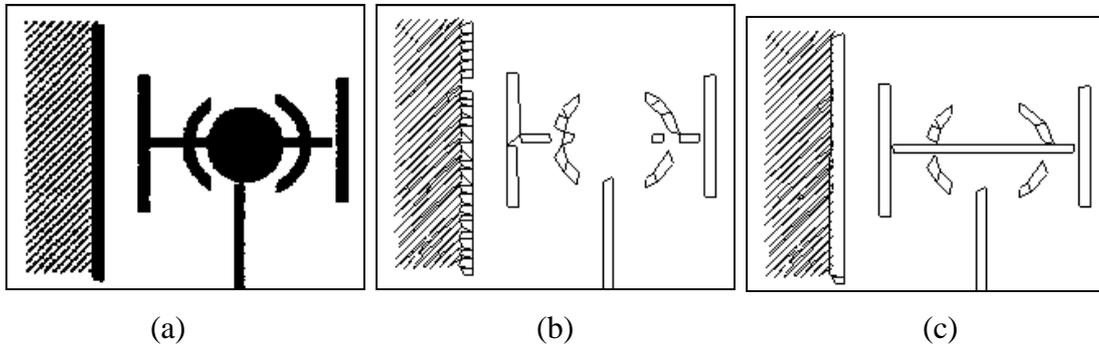


Figure 2.14 : Utilité de la phase de fusion des Quadrilatères
 (a) Image initiale, (b) avant fusion (c) après fusion

Pour cela, on définit une distance δ ou plutôt un écart entre 2 Quadrilatères (figure 2.15) :

Si $|\text{angle}Q - \text{angle}Q'| < 30^\circ$ **ET** $\text{Noir_entre}(Q, Q')$ **alors**
 $\delta(Q, Q') = \text{Min}(P1P2' + Q2Q1', P1Q2' + Q2P1')$

Sinon
 $\delta(Q, Q') = + \infty$

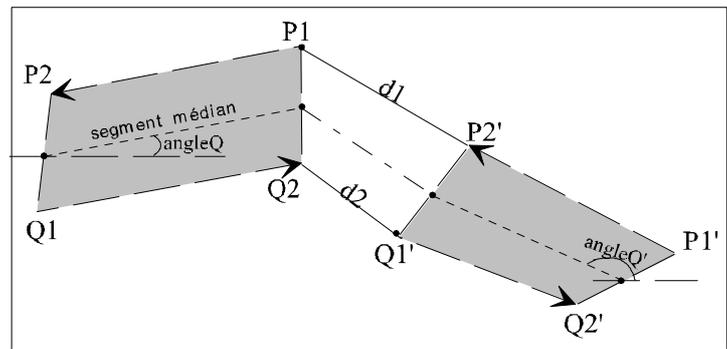


Figure 2.15 : Distance entre Quadrilatères

Une fois la distance entre Quadrilatères définie, il reste à choisir la manière de sélectionner les Quadrilatères servant de point de départ à la reconstitution des différents tracés constituant le dessin. Une fois ce choix réalisé, une recherche de son plus proche voisin permet de définir de proche en proche le nouveau Quadrilatère courant et ceci jusqu'à la rencontre d'un Quadrilatère sans voisin (suivant) marquant la fin d'un tracé.

Le choix du Quadrilatère à l'origine d'un tracé, nécessite aussi l'utilisation d'une distance entre Quadrilatères (distance euclidienne). Cette fois-ci, nous avons choisi de rechercher l'élément ayant le voisin le plus éloigné de lui, la probabilité que ce Quadrilatère corresponde à l'extrémité d'un tracé est ainsi accrue (figure 2.16).

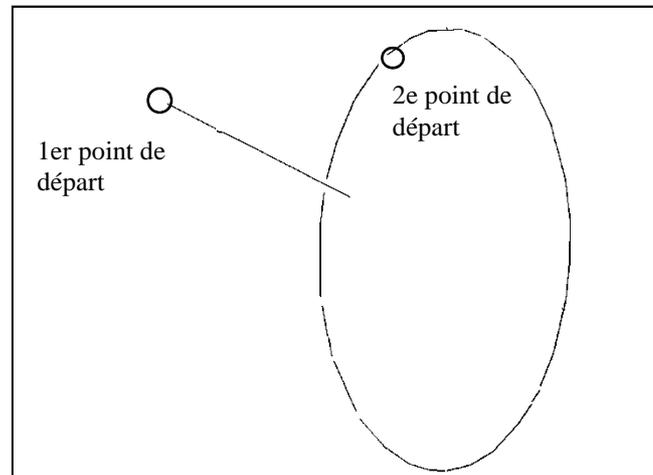


Figure 2.16 : Choix d'un point de départ aisé pour une droite, indifférent pour un cercle

Ce tri permet de revenir à la manière dont a été construite l'image de départ. Une chronologie plausible du tracé du dessin est reconstituée et stockée dans la suite SQ (stockage des *Quadrilatères* construits).

En se basant toujours sur le principe d'analyse de l'image à différents niveaux de détails, la phase de fusion des Quadrilatères est réalisée suivant le même algorithme d'approximation polygonale que celui utilisé lors de la construction des vecteurs, cette fois-ci, appliqué aux extrémités des segments médians des *Quadrilatères* (figure 2.17). L'algorithme tient cependant compte des attributs de chaque *Quadrilatère* (notamment de l'épaisseur) pour autoriser ou refuser la fusion de certains Quadrilatères. De plus, il consulte l'image initiale pour vérifier la présence de pixels noirs entre les 2 quadrilatères avant d'autoriser leur fusion.

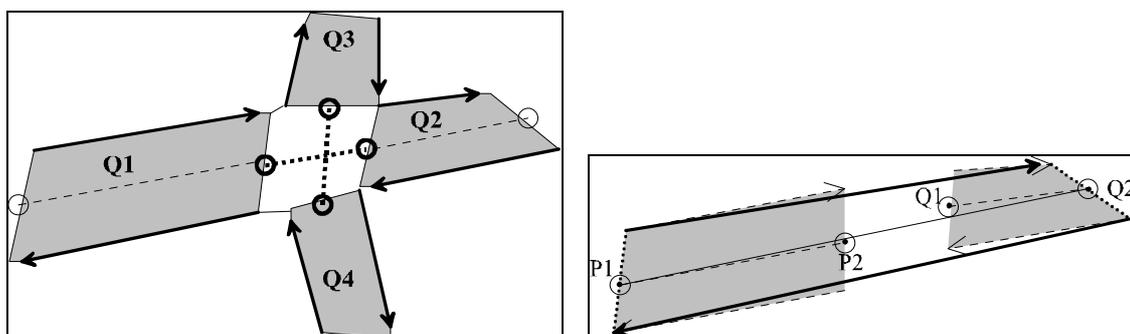


Figure 2.17 : Fusion des Quadrilatères

De la même manière que lors de la fusion des Vecteurs, le processus s'arrête lorsque plus aucune fusion n'est possible. Le nombre de cycles Tri - Fusion varie la plupart du temps entre 3 et 5. Des exemples de résultats sont fournis figures 2.18 et 2.19.

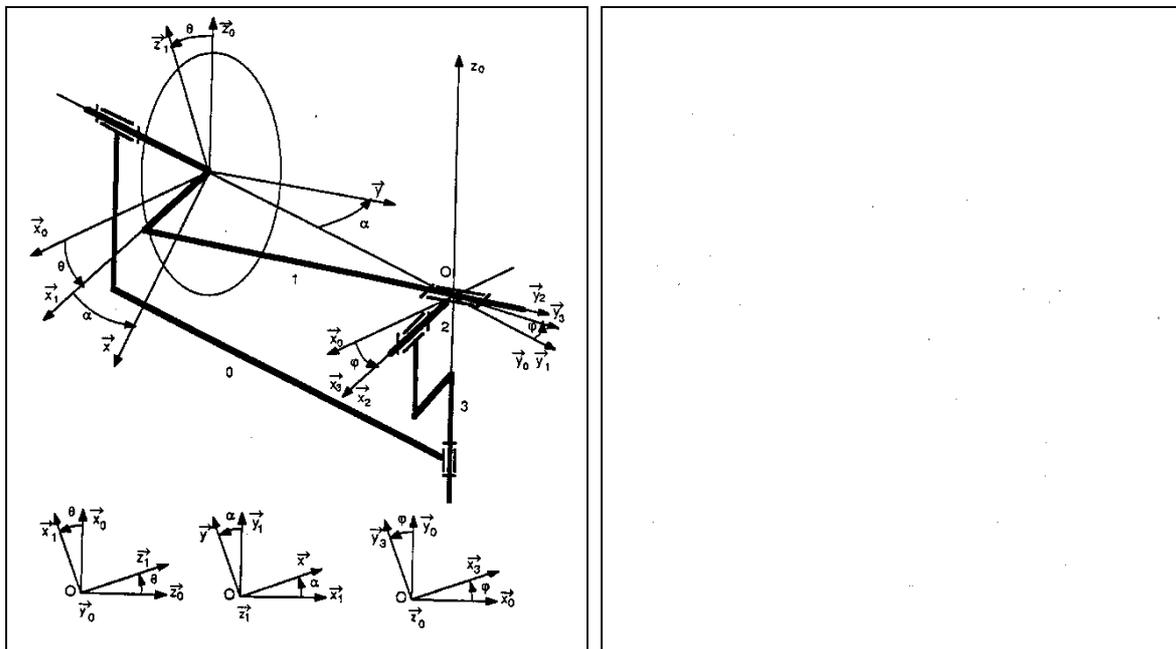


Image initiale

Axes médians des quadrilatères

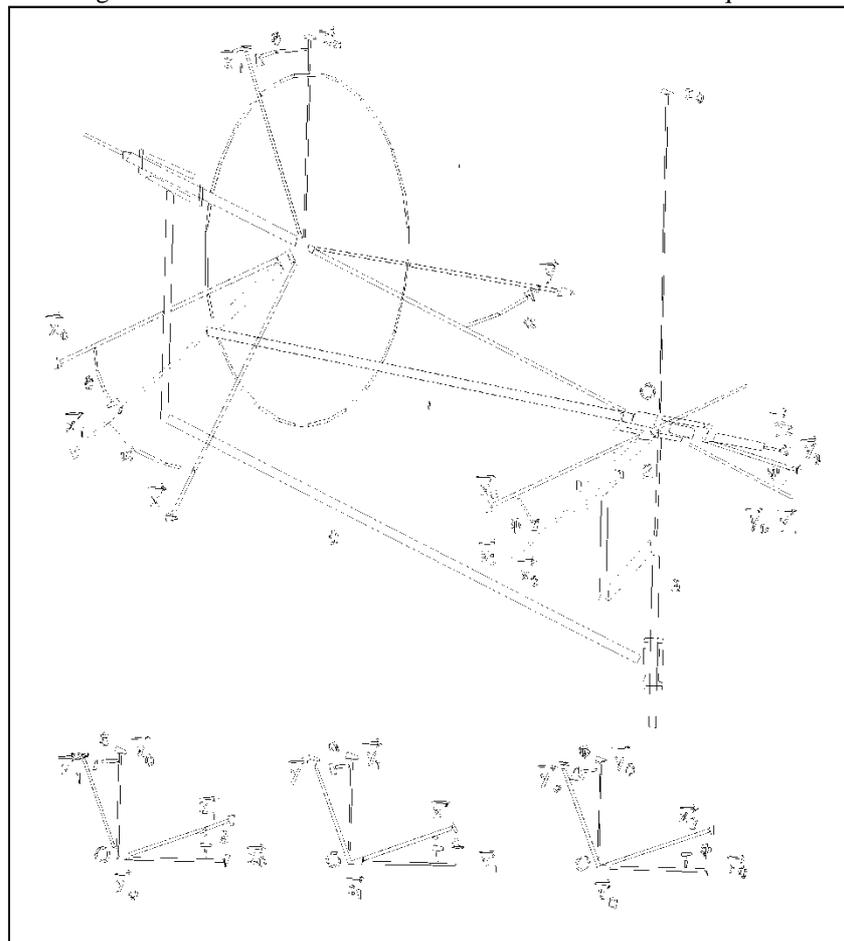


Image des quadrilatères

Figure 2.18 : Images des Quadrilatères et représentations axiales correspondantes

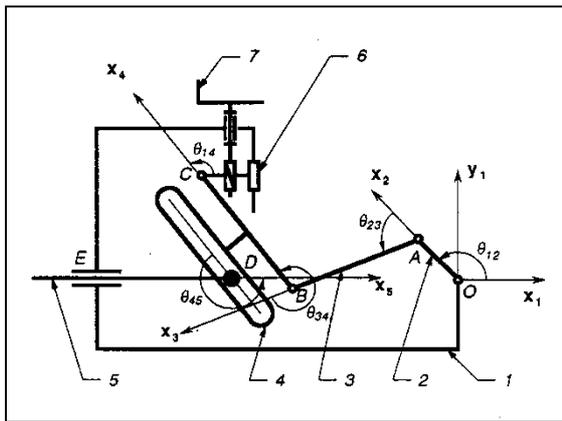
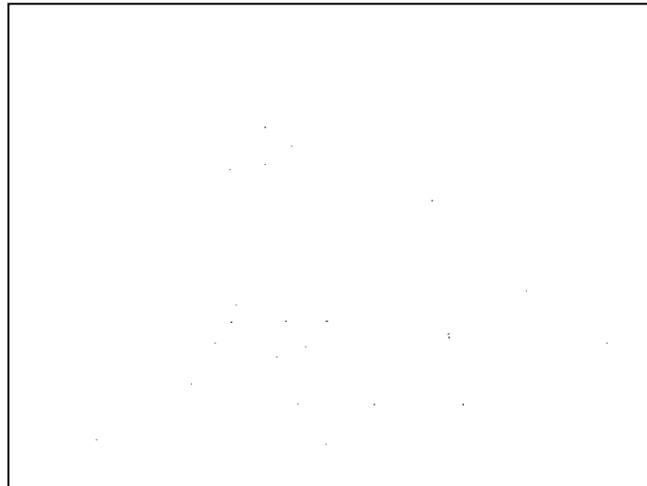


Image initiale



Axes médians des quadrilatères

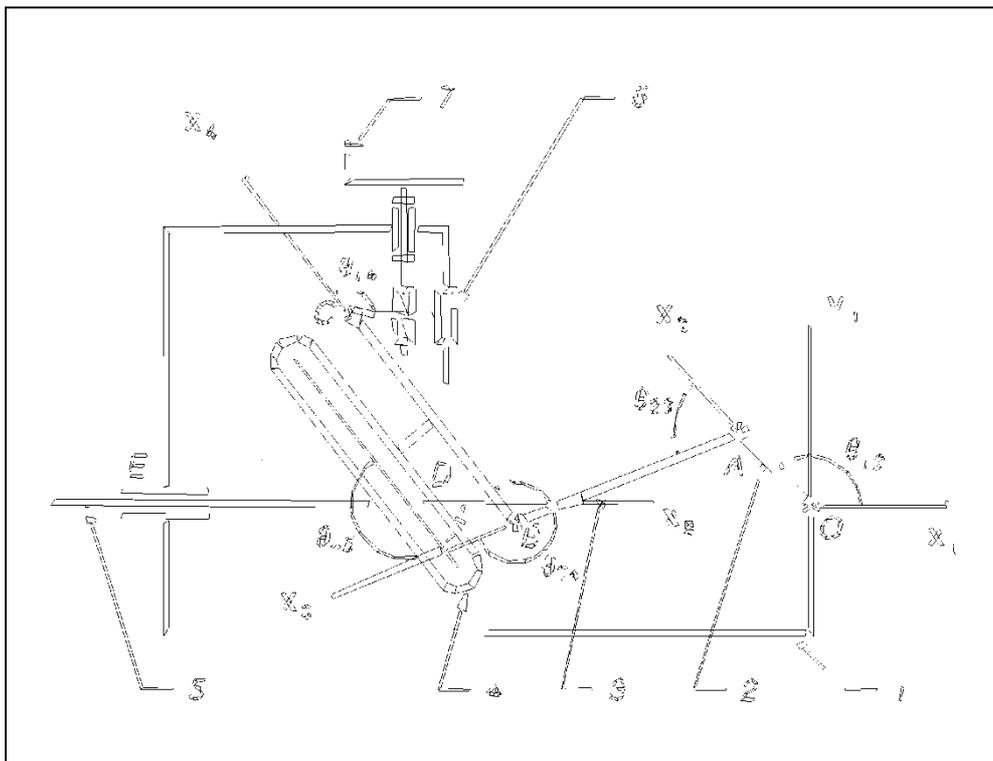


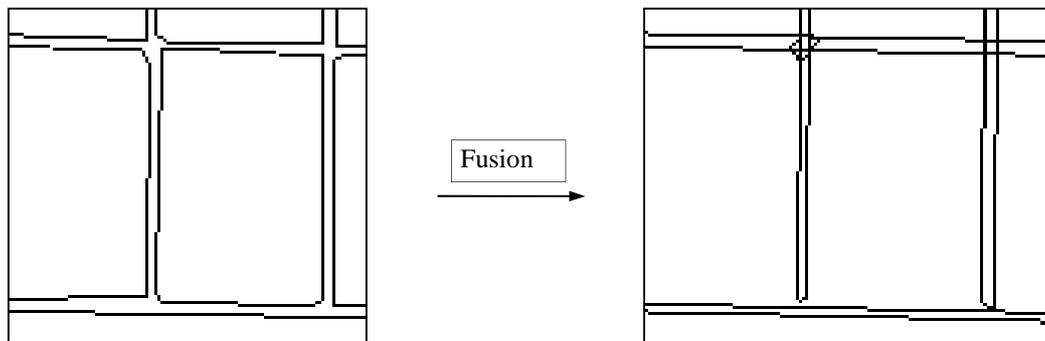
Image des quadrilatères

Figure 2.19 : Images des Quadrilatères et représentations axiales correspondantes

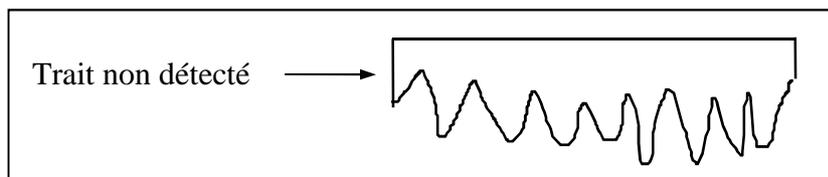
L'ensemble des objets Quadrilatères finalement obtenu est représentatif des formes fines présentes dans le document initial. Il ne faut cependant pas voir cette étape comme une véritable étape de vectorisation du document mais plutôt comme la méthode que nous avons choisie pour obtenir une description primaire des formes fines d'un document. Partant de ces données, des étapes supplémentaires permettent d'obtenir une vectorisation de haute qualité.

En effet, cette représentation ne traduit pas directement les intersections entre traits du document original. La phase de fusion des Quadrilatères (sensible aux changements de direction) engendre parfois l'apparition de bruits au niveau de certaines intersections (figure

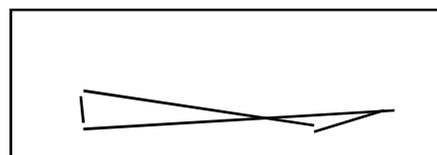
2.20a). Cette représentation ne localise pas certains types de traits (appariement impossible - figure 2.20b) relevant d'ailleurs plus de la détection de contours subjectifs. Enfin, la discrétisation provoque la dégénérescence de certains Quadrilatères lorsque l'épaisseur des traits est insuffisante (figure 2.20c).



(a) Quadrilatères parasites



(b) Contours subjectifs



(c) Quadrilatères dégénérés

Figure 2.20 : Quelques difficultés

2.3. Apport du modèle de description

Les différentes étapes effectuées se marquent bien dans les figures 2.27 et 2.28 où, à partir de l'image initiale (en haut à gauche) on a extrait d'abord des vecteurs (en haut à droite), les quadrilatères (en bas à gauche), tandis que ne subsistent que les vecteurs non appariés (en bas à droite) dont l'importance sera mise en évidence plus loin. On sera frappé par la faible perte d'information entraînée par les opérations précédentes.

La méthode de description choisie présente plusieurs avantages. En premier lieu, elle apporte une vision globale du document, se montre capable d'évoluer, et de s'adapter aux diverses situations, aussi bien que de contenir une grande quantité d'information dont elle peut assurer le partage.

2.3.1. Une Vision Globale

Ce modèle de représentation évolutif ne doit pas être confondu avec les modèles de représentation décrivant la structure d'un type précis de documents, utilisé par de nombreux systèmes de lecture automatique. Ces derniers sont généralement statiques et traduisent à l'aide de règles une structure et des connaissances figées auxquelles doivent ensuite se conformer les documents à interpréter. Dans ce dernier cas, la notion d'évolution des interprétations jusqu'à l'émergence d'une solution pertinente n'est d'ailleurs pas du tout reproduite alors que nous verrons, dans la suite, comment il est possible de faire évoluer notre représentation initiale du document jusqu'à ce qu'elle traduise la compréhension du contenu du document.

Contrairement à ce type de modèle de description, celui construit par notre système peut être comparé à la représentation mentale d'une scène utilisée par le système visuel humain. Sa construction se base sur des extractions d'informations dans l'image pour les traduire en terme d'objets. Son mode de construction se fonde sur une analyse préalable de la structure des documents à interpréter (connaissances a priori). Ce modèle fournit des informations sur l'allure générale du document, vision globale qui donne accès aux principales caractéristiques du document (directions, épaisseur, types de formes). Cette partie a pour but de présenter plus précisément les différentes caractéristiques fournies.

2.3.1.1. Directions privilégiées

Avant l'étape d'appariement, une connaissance concernant les directions privilégiées de l'image est déjà contenue dans l'ensemble des objets *Vecteurs* (attribut Angle) et peut être traduite visuellement par le calcul d'une rose des directions (figure 2.21).

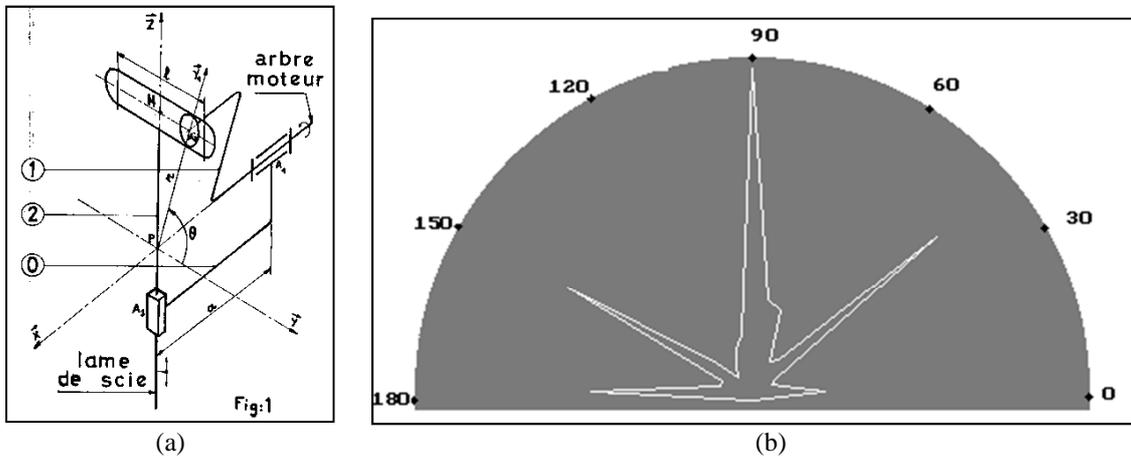


Figure 2.21 : (a) Image initiale et (b) Rose des directions

Cette information est primordiale pour les experts mécaniciens qui, lorsqu'ils lisent un dessin mécanique, que ce soit une projection orthogonale ou un schéma cinématique, s'intéressent particulièrement aux parallélismes, orthogonalités, et axes de symétrie mettant en rapport certains traits du document (figure 2.22). Ces éléments sémantiques fondamentaux leur permettent d'examiner la cohérence du dessin. En outre, dans notre cas, cette information est très utile lors de la localisation et de la reconnaissance des symboles normalisés ainsi que lors de la phase de reconstruction 3D du mécanisme.

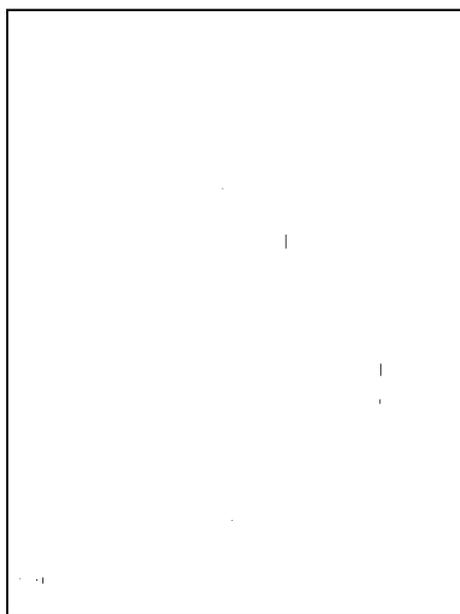


Figure 2.22 : Sélection des traits selon leur direction

2.3.1.2. Épaisseur des traits

Les objets de la classe Quadrilatère sont quant à eux représentatifs des formes fines rectilignes présentes dans le document initial. La suite SQ permet de générer un « squelette » ainsi que l'histogramme des épaisseurs des formes fines présentes dans le document (figure 2.23). Il suffit pour cela d'utiliser les attributs épaisseur et longueur de chaque quadrilatère. Les accumulateurs associés aux différentes épaisseurs sont incrémentés de la longueur de chaque quadrilatère ayant l'épaisseur correspondante.

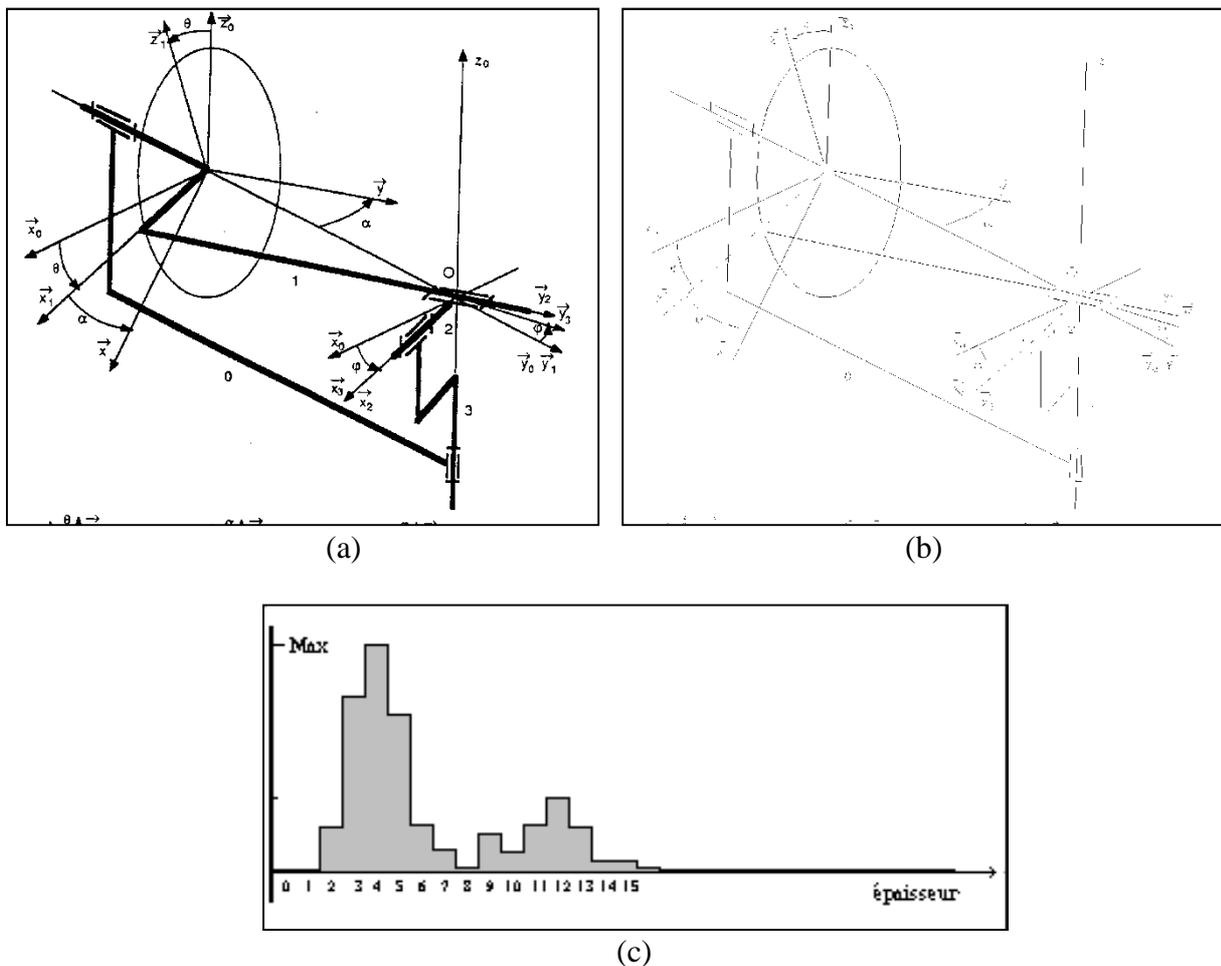


Figure 2.23 : (a) Image initiale , (b) Représentation axiale, (c) Histogramme des épaisseurs

L'information épaisseur des lignes est aussi importante que l'information direction puisqu'une partie des traits (traits fins) correspond aux annotations, légendes ou axes, tandis que la seconde partie (traits épais) représente le mécanisme mécanique lui-même. Il est donc important que le modèle de description choisi pour représenter des documents techniques permette de sélectionner les traits suivant un critère d'épaisseur (figure 2.24).

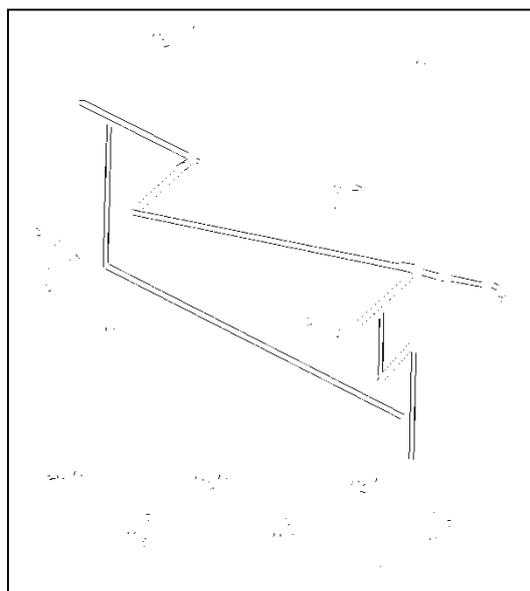


Figure 2.24 : Sélection des traits suivant leur épaisseur

Notre système utilise l'information épaisseur dès les traitements de bas niveau (mise en place du modèle de représentation - phase de fusion des quadrilatères). De cette manière, la distinction entre les différents types de traits présents dans le document est réalisée dès le départ, ce qui n'est pas possible lorsque l'on utilise un mécanisme de vectorisation basé sur la squelettisation. En effet, avec notre méthode, 2 traits successifs d'épaisseurs différentes se traduisent par la présence de plusieurs quadrilatères (correspondant aux différentes épaisseurs) dans la représentation et non par un seul vecteur (figure 2.25).

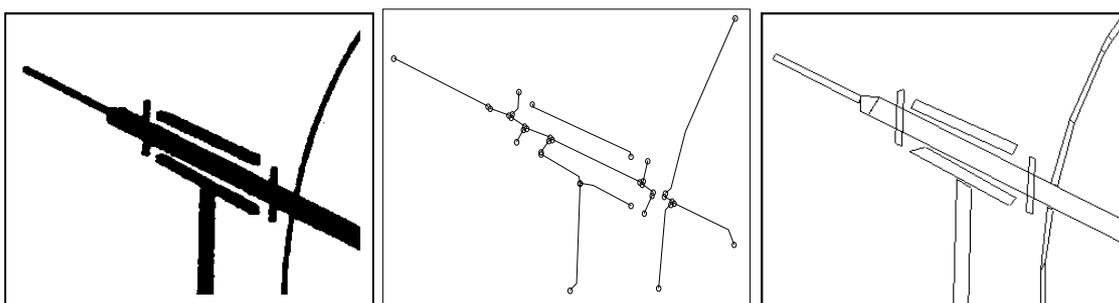


Figure 2.25 : Différence entre lignes brisées et Quadrilatères

L'utilisation de l'information épaisseur n'est pas toujours aisée. Comme on peut le voir sur les histogrammes des épaisseurs de traits, il n'y a pas seulement 2 valeurs distinctes d'épaisseurs de traits dans une image. Certains quadrilatères (ceux qui correspondent aux caractères, ...) ont une épaisseur intermédiaire et rendent difficile le classement en 2

catégories disjointes (traits épais/trais fins) des différents traits dès ce niveau de l'analyse. Il est préférable d'accumuler d'autres connaissances et d'utiliser leur fusion pour ne pas réaliser, dès ce niveau, des erreurs d'interprétation. Aussi, ce classement des traits n'est-il pas effectué immédiatement.

Le calcul des histogrammes des épaisseurs de traits à plusieurs instants du traitement montre (figure 2.26) que les 2 pics (correspondant aux 2 principaux types de traits) sont de plus en plus marqués au fur et à mesure que l'interprétation avance et prouve donc qu'il est préférable d'attendre.

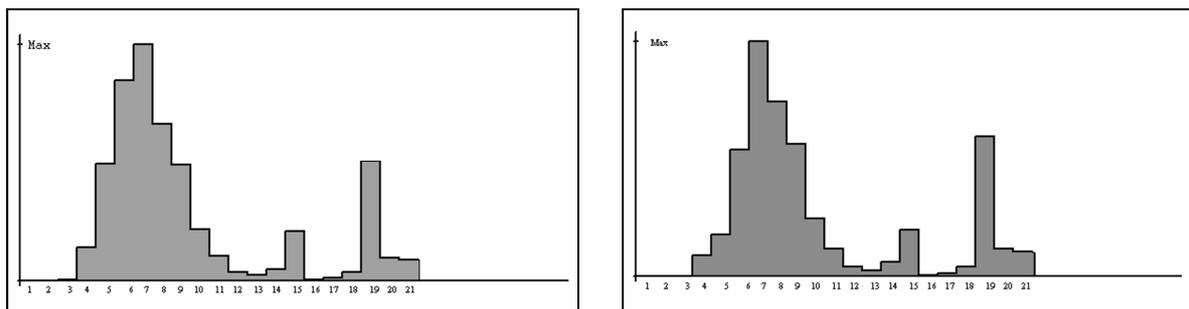


Figure 2.26 : Histogrammes avant et après extraction du texte

2.3.1.3. Type de forme

Les résultats obtenus par le changement de représentation effectué permettent de constater que les formes fines sont décrites par l'ensemble des quadrilatères. Par contre, les vecteurs pour lesquels l'appariement n'a pas été réalisé correspondent aux formes pleines ou à des parties de caractères non rectilignes (figures 2.27 et 2.28).

Pour faire évoluer cette représentation, il est nécessaire d'étudier plus précisément chacun de ces ensembles. C'est ce qui est fait dans la suite de la chaîne d'interprétation.

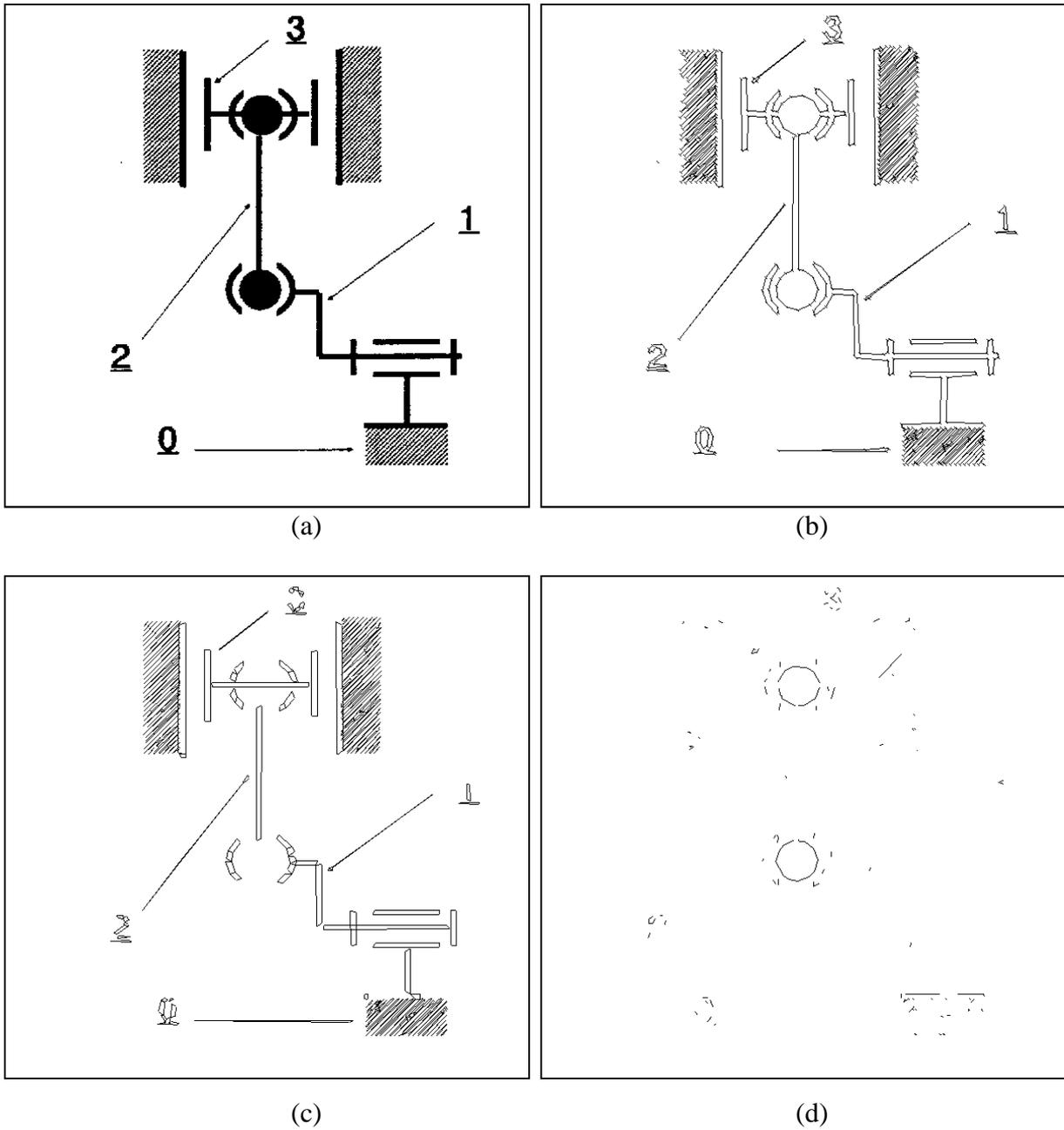
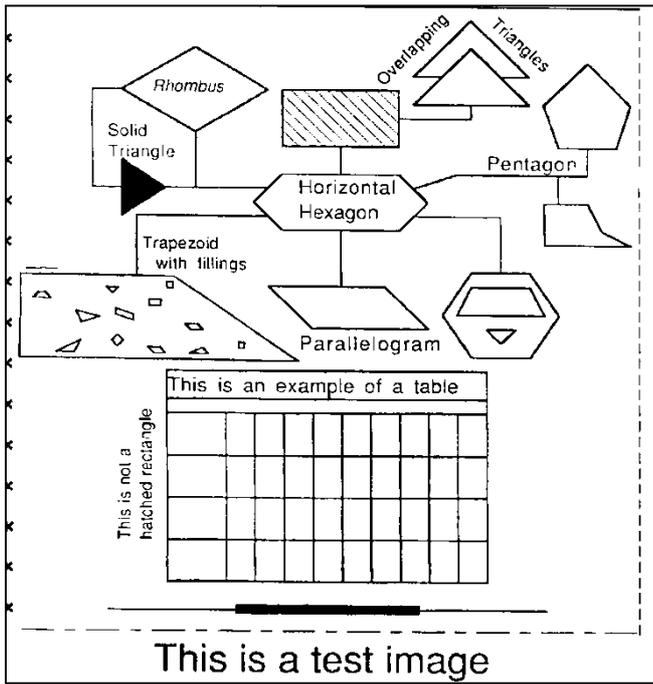
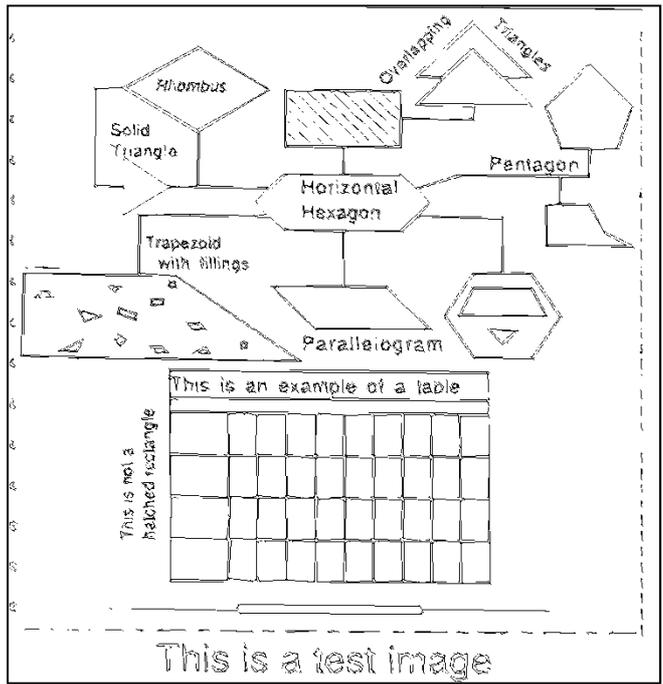


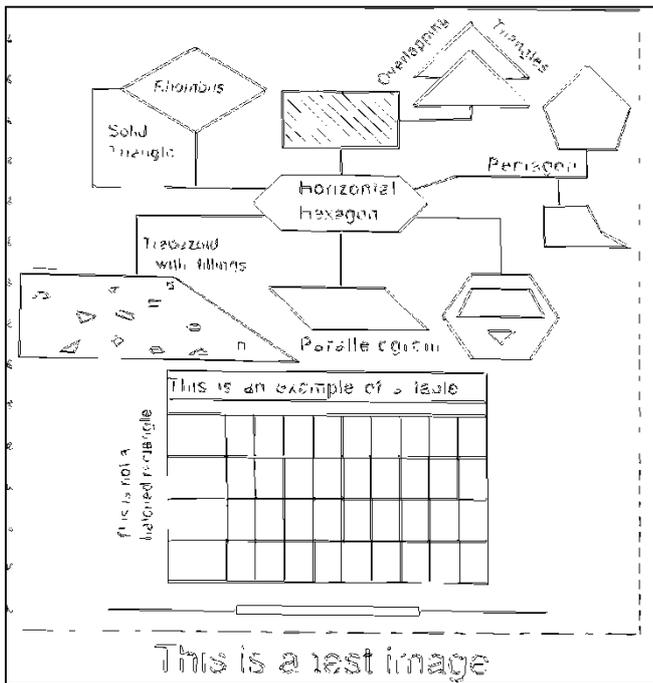
Figure 2.27 : (a) Image initiale, (b) Vecteurs, (c) Quadrilatères, (d) Vecteurs non appariés



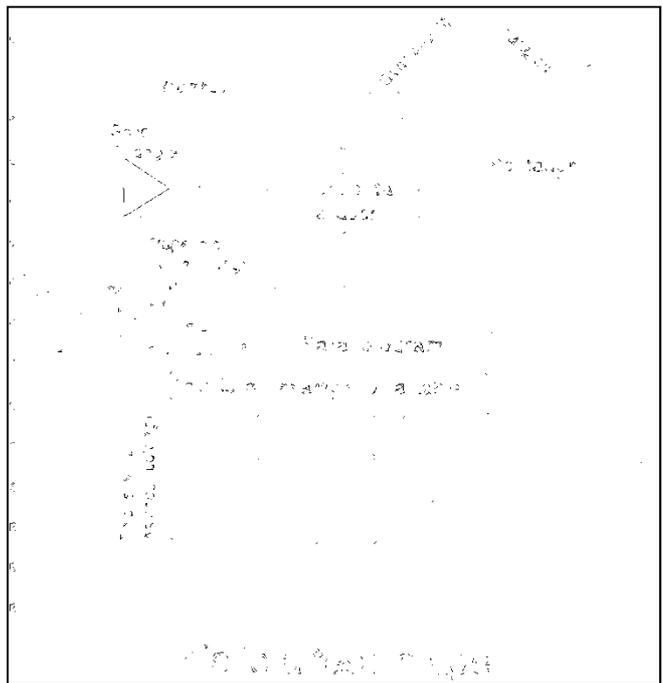
(a)



(b)



(c)



(d)

Figure 2.28 : (a) Image initiale, (b) Vecteurs, (c) Quadrilatères, (d) Vecteurs non appariés

☆ Etude des Formes Pleines

Après l'étape d'appariement, la suite SV de Vecteurs ne contient plus que les objets n'ayant pu être mis en correspondance avec aucun autre. Les vecteurs qui ont pu être appariés ont été enlevés de SV et ont donné lieu à un quadrilatère de la suite SQ. La plupart des vecteurs restant alors dans SV constituent un élément de la frontière d'une forme pleine (figures 2.27d et 2.28d). Les ensembles de vecteurs (de SV) correspondant aux frontières d'une forme pleine peuvent chacun être caractérisés par :

- la présence de pixels Objets entre les extrémités des vecteurs successifs constituant la frontière (une forme pleine a un contour fermé)
- une somme des longueurs des vecteurs successifs importante (le périmètre d'une forme pleine est une valeur élevée)

On peut alors extraire les formes pleines du dessin en procédant, dans SV, à la recherche des éléments de contours quasi fermés. On se réfère aussi à l'image pixelaire.

Dans ce but, les vecteurs, éléments de SV, sont triés selon la distance qui sépare leurs extrémités (figure 2.29) :

$$\text{Successeur}(V_i) = V_s \text{ avec } V_s \text{ vérifiant } \text{dist}(V_{iF}, V_{sI}) = \text{Min}\{\text{dist}(V_{iF}, V_{jI})\},$$

pour j variant de 1 à $\text{card}(SV)$ et $i \neq j$

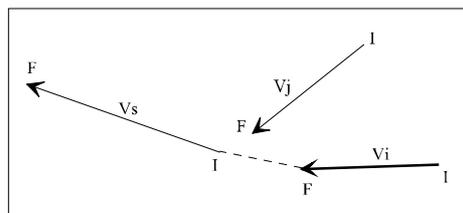


Figure 2.29 : Choix du successeur d'un vecteur

De la même manière que pour la fusion des quadrilatères, ce tri nécessite le choix d'un point de départ. Celui-ci n'influence cependant pas le résultat final. La méthode utilisée fait appel au même procédé que celui décrit et mis en place pour le tri des Quadrilatères.

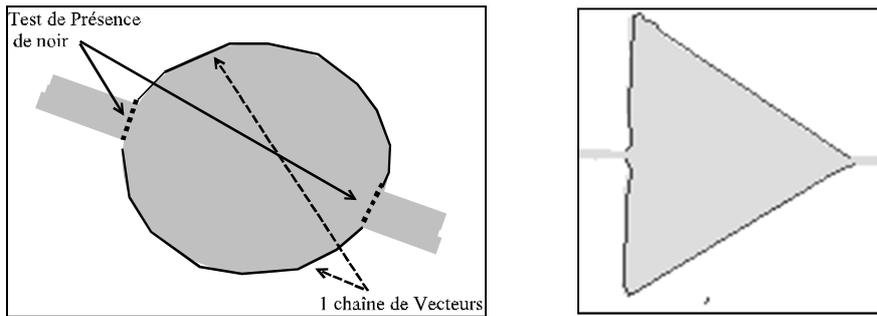


Figure 2.30 : Recherche des chaînes de contours des formes pleines

Une fois les vecteurs classés, on recherche (dans SV) les lignes polygonales de longueurs significatives. L'algorithme nécessite l'examen de l'image initiale pour tester la présence de noir entre 2 vecteurs consécutifs de SV triée (figure 2.30). S'il n'y a pas suffisamment de noir, il s'agit d'une extrémité de la ligne polygonale ne correspondant donc pas à une forme pleine (comme cela est le cas dans la zone hachurée en haut à droite de l'image 2.27). Les lignes fermées délimitent une zone noire dans l'image initiale (objet toujours à gauche des vecteurs) et traduisent la présence des *Formes Pleines*. Ces dernières sont extraites et stockées (figure 2.31) sous forme d'objets (avec attributs).

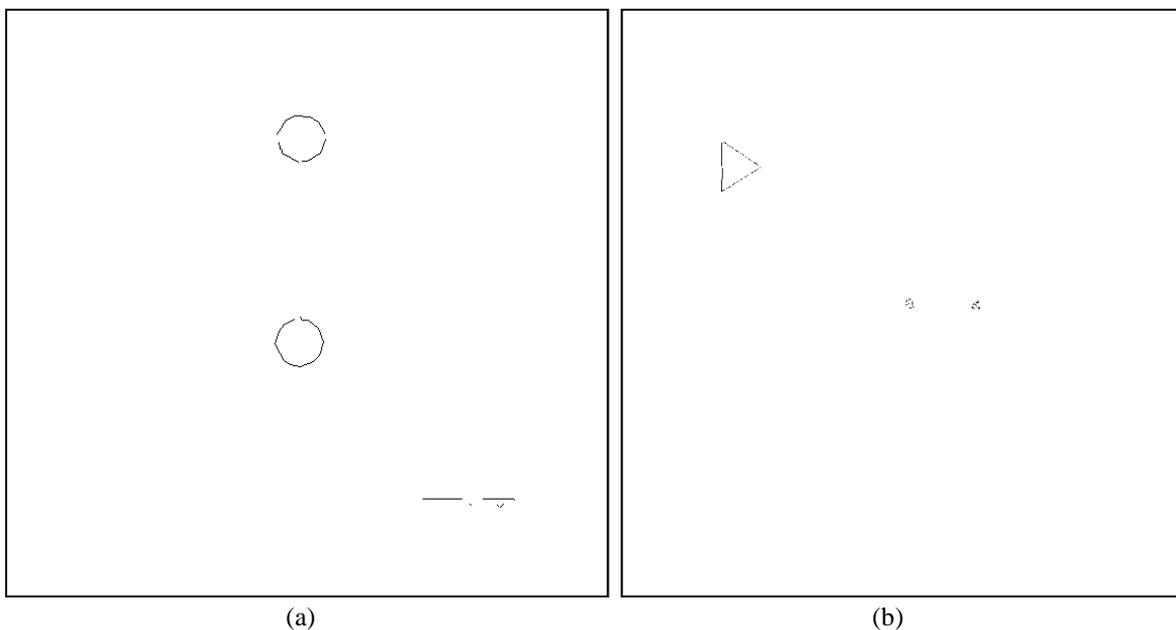


Figure 2.31 : Images des Formes Pleines extraites

(a) de l'image 2.27 et (b) de l'image 2.28

Le choix du périmètre minimal des formes pleines (longueur du contour polygonal inclus dans SV) ainsi que la distance maximale autorisée entre les extrémités de 2 vecteurs

successifs, constituant une ligne polygonale, n'est pas trivial. Ces seuils ont été choisis de manière à limiter le nombre de rejets abusifs. Ainsi, les lignes ayant une longueur même assez faible sont acceptées en tant que formes pleines. On voit un exemple de ce cas dans le résultat (figure 2.31b) où le triangle plein a été détecté mais aussi 2 formes de faible taille qui correspondent à 2 minuscules du mot parallélogramme. Dans la même optique, les regroupements de vecteurs ayant des extrémités éloignées est possible à condition que suffisamment de pixels noirs soient présents entre leurs extrémités. Figure 2.31a, on voit que la ligne noire, en bas à droite, limitant les hachures n'a pas conduit à la formation d'un quadrilatère car un de ses bords n'est pas assez marqué. Les vecteurs correspondant à cette zone sont alors pris comme frontière d'une forme pleine puisque les critères précédemment définis ne sont pas incompatibles avec ce type de forme.

☆ Etude des Formes Fines

La liste SQ est, quant à elle, représentative des formes fines présentes dans l'image initiale puisque les critères d'appariement ont été choisis dans ce but.

Comme nous l'avons déjà précisé, cette représentation ne peut pas raisonnablement être considérée comme une vectorisation de haut niveau pour 2 raisons principales :

- d'une part elle ne traduit pas toujours les connexités entre les différents traits présents dans le dessin (figure 2.32).

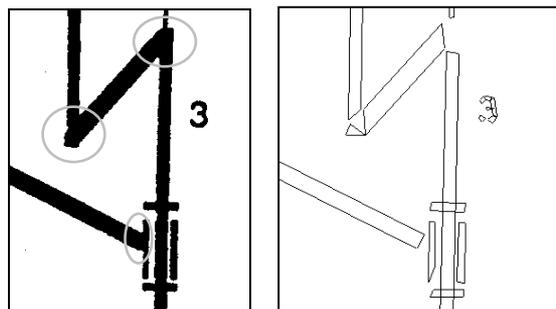


Figure 2.32 : Connexités non traduites

- d'autre part, un trait dans l'image initiale engendre parfois plusieurs quadrilatères dans le modèle de représentation (figure 2.33).

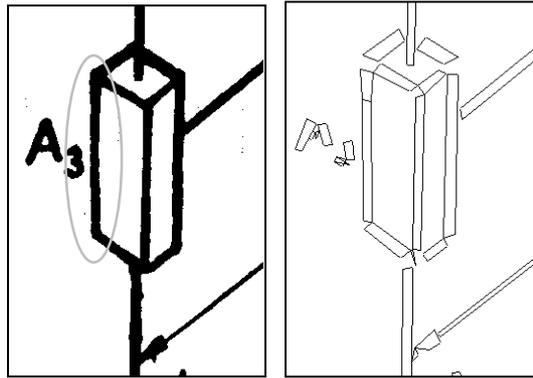


Figure 2.33 : A un trait correspondant deux quadrilatères

Il est bon de rappeler que les quadrilatères sont représentatifs de toutes les formes fines : c'est à dire, non seulement des traits, mais aussi des courbes et donc de certaines parties de caractères. Ces phénomènes sont évidemment normaux, explicables par la technique choisie et ne doivent donc pas poser de problème. Ce modèle de représentation obtenu a pour but de faciliter la gestion des connaissances extraites grâce aux attributs associés aux différents objets. Il s'agit de primitives de description de bas niveau permettant d'étudier la disposition relative des formes dans le document. En ce sens, le modèle assure donc complètement le rôle qui lui avait été assigné.

Pour mettre en évidence l'étendue de notre modèle et son intérêt, nous l'avons expérimenté sur des documents divers et nous présentons ici les résultats obtenus à partir d'images de partitions musicales. Les problèmes sont différents des nôtres, principalement l'extraction des notes sur les lignes de portées présente des difficultés particulières [Leplumey93]. Sur les figures 2.34 à 2.38 qui suivent, les différentes étapes du traitement sont résumées. Ce mode de représentation permet l'extraction des lignes (figure 2.35), la localisation des formes pleines (têtes de notes noires - figure 2.36) et la sélection d'un type de lignes, par exemple les lignes horizontales (correspondant normalement aux portées de la partition - figure 2.37). Comme on le verra dans le chapitre suivant, il est également possible de localiser les zones de texte grâce à ce modèle de description (figure 2.38).

THE ENTERTAINER RAG
 by Scott Joplin - The Raggy Doctor

Michael's Ragtime Echo

22

Figure 2.34 : Image initiale

Figure 2.35 : Image des quadrilatères

Figure 2.36 : Formes pleines

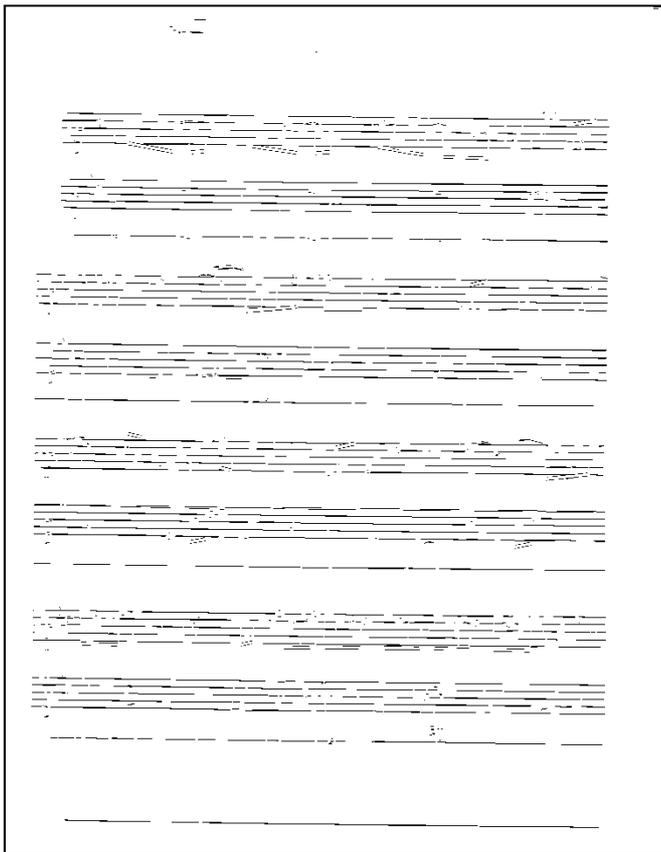


Figure 2.37 : Lignes horizontales

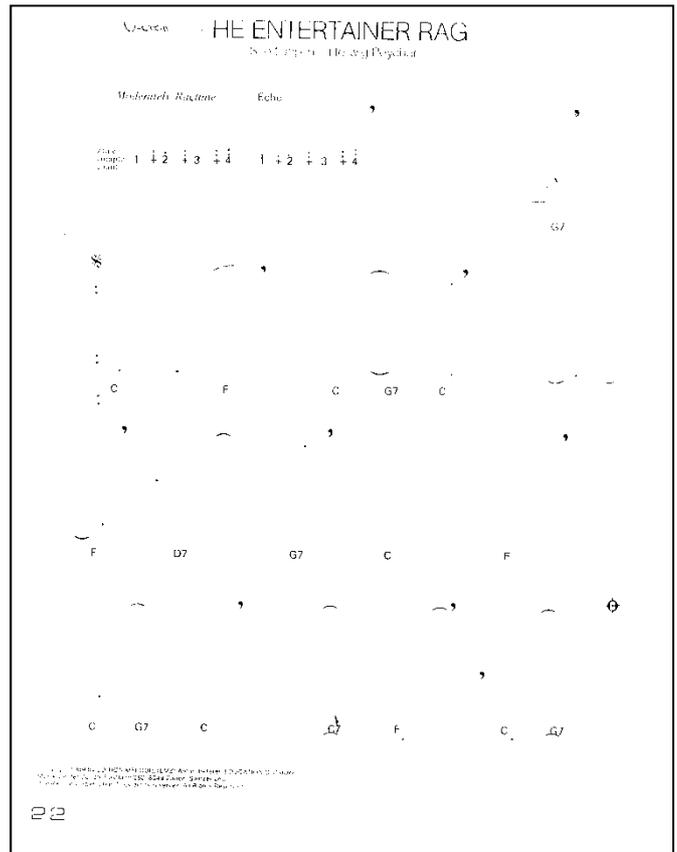


Figure 2.38 : Zones de texte

Néanmoins, cet ensemble de primitives de description décrit une partie encore trop locale de l'image et n'apporte pas suffisamment de renseignements sur l'aspect global du document. Les travaux déjà réalisés montrent clairement qu'il est nécessaire de tenir compte du contexte le plus général possible lors des différents traitements à mettre en oeuvre. Ainsi, à ce stade du traitement, il nous apparaît nécessaire d'organiser les informations extraites de manière à mieux mettre en évidence les rapports qui existent entre elles. Une représentation sous la forme d'un graphe nous semble la plus adaptée pour relier les différentes formes qui constituent le dessin en fonction de la description locale, et la plus apte à traduire les liens topologiques existant entre les primitives (pour passer à un niveau plus global de description). L'ensemble des primitives de représentation (Vecteurs, Quadrilatères, Symboles Pleins) constitue les noeuds du graphe. Les arcs reliant ces noeuds traduisent les relations topologiques qui existent entre les primitives.

A ce niveau de l'analyse, la quantité de données à gérer est bien réduite mais reste encore importante. Une image du type de celles présentées jusqu'à présent qui comportent environ 3000x1000 pixels a été ramenée à environ 1000 quadrilatères auxquels peuvent s'ajouter 1000

vecteurs non appariés. Il ne paraît pas envisageable d'étudier une à une ces primitives pour construire un graphe décrivant les relations qui existent entre ces objets. De plus, il n'est pas toujours utile de mettre en relation entre eux, les quadrilatères non représentatifs du mécanisme (parties de caractères, ou d'annotations, ...). En utilisant toujours une démarche d'extraction progressive (hiérarchique), et à l'aide de l'outil de description mis en place, des entités de niveau plus élevé peuvent être extraites de manière réfléchie et robuste. En effet, il est possible de reconstituer d'abord certaines entités : courbes, annotations complètes, et autres formes en tirant parti de leurs particularités et sans utiliser de structuration supplémentaire. Ces localisations ou reconstructions d'objets permettent de réduire leur quantité, d'en faciliter la gestion et d'obtenir un graphe qui sera plus représentatif de la sémantique du mécanisme mécanique.

La construction du graphe structurel sera donc effectuée plus tard dans la chaîne des traitements lorsque l'interprétation rendra indispensable son utilisation. Ce report se justifie seulement par la volonté d'optimiser le temps de traitement et d'adapter les différentes étapes à la forme même des objets à détecter. Sans doute aurait-il été plus cohérent de construire la représentation structurelle durant la mise en place du modèle de description. Néanmoins, ce délai ne change rien à la philosophie employée ni aux résultats obtenus. Nous avons choisi une présentation de la méthode employée proche de la pratique plutôt que l'élaboration trop théorique d'une méthode que nous aurions dû trahir lors de la mise en oeuvre.

2.3.2. Partage de l'information

Chez l'Homme l'information est, semble-t-il, centralisée, au niveau du cerveau et donc aucune difficulté de coopération ne semble survenir. Tout est différent dans le cas de systèmes informatiques où se posent les problèmes de transfert des connaissances et de communication des résultats intermédiaires entre les différents processus d'analyse. Le modèle de représentation que nous venons de décrire a été choisi pour contenir toutes les données utiles et pour permettre leur évolution durant l'ensemble de l'analyse.

Nous allons maintenant mettre l'accent sur les principes d'interaction qu'il est possible de mettre en oeuvre grâce à ce modèle de description. Ces principes découlent tous de la manière de partager, de traiter, et de faire évoluer l'information.

Ce modèle se base sur l'image initiale de la scène observée, et évolue jusqu'à former un tout cohérent et interprété. Les connaissances a priori sur la structuration des documents à traiter se trouvent autant dans la manière dont on choisit de faire évoluer le modèle de représentation, que dans la structure du modèle de représentation lui-même. En modifiant le plan régissant l'enchaînement des processus faisant évoluer le modèle du document, le système peut s'adapter et analyser des documents de structure variée (partitions musicales).

Notre approche est fondée sur la coopération entre niveau global et niveau local. La méthode d'analyse comporte deux phases qui coopèrent, l'une utilisant les informations fournies par l'autre (figure 2.39) [Ramel96a] :

- vision globale durant laquelle est constituée une représentation initiale du document;
- vision locale durant laquelle des processus d'analyse locale travaillent à partir de ce contexte et le font évoluer jusqu'à la compréhension totale du document.

Au niveau local, nous définirons des spécialistes qui devront coopérer, grâce au modèle de représentation, pour extraire les Formes Pleines, le Texte, les Courbes, les Eléments Mécaniques, ... Les spécialistes travaillent en mode prédiction-vérification (génération d'hypothèses à l'aide du modèle, puis vérification dans l'image initiale). Ils contribuent à l'évolution de la représentation initiale, et sont activés selon un ordre pré-défini.

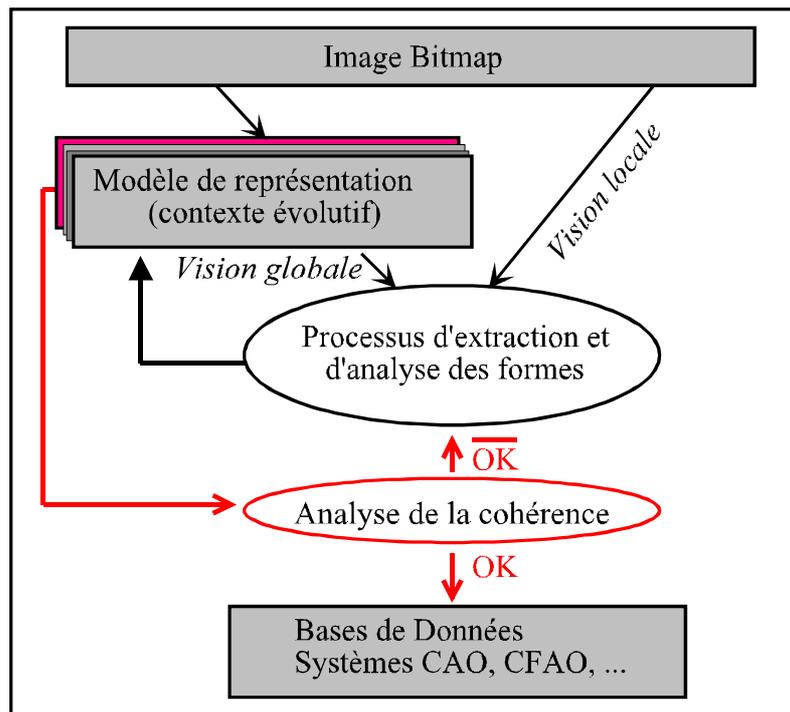


Figure 2.39 : Architecture du système

D'autre part, il est nécessaire de maintenir à jour ce qui correspond, dans le système, au médium de communication : c'est à dire de modifier et faire évoluer le modèle de représentation du document vers la solution finale. Par ce point, l'architecture de notre système est proche de celles des **systèmes distribués** travaillant sur un **espace commun (blackboard)**.

Coopération

Notre système d'interprétation fonctionne selon un mode contrôle centralisé, qui permet une réelle souplesse dans le déroulement d'un plan d'action, grâce au partage d'informations réalisé par le biais du modèle de représentation (qui peut être vu comme un blackboard). A défaut d'un parallélisme réel, l'accès à une ressource (le modèle de description) permet de gérer la coopération entre les différents processus d'analyse. Les processus travaillent alors de manière incrémentale par appropriation ou modification de certaines primitives du modèle. Chaque processus réalise sa tâche, s'interrompt, et à la fin de chaque cycle d'analyse une demande de lancement de processus complémentaires est faite si nécessaire. Selon cette approche, l'exécution des tâches est réellement entrelacée, et il devient possible de bénéficier d'une réelle interaction entre les traitements.

Pour qu'un contrôle incrémental soit efficace, il est nécessaire que chaque processus puisse s'interrompre de lui-même pour demander une coopération avec un autre processus (facilité de changement de contexte, liberté de l'information), de telle sorte que le lieu et la manière ne soient pas imposés et que des informations complémentaires soient obtenues.

Dans le cas d'un contrôle décentralisé, tout agent doit pouvoir décider de l'opportunité de déclencher un autre agent. Dans ce but, il doit pouvoir s'interrompre dès que :

- son comportement n'est plus adapté
- certains événements sont rencontrés

et mettre alors en oeuvre le processus dont la compétence est adaptée à la nouvelle situation.

Afin de reprendre les avantages de tels systèmes, nous avons choisi d'utiliser un mécanisme de cycles perceptifs successifs permettant à différents spécialistes d'extraire d'abord les **évidences** et d'être appelés plusieurs fois si nécessaire durant l'interprétation. De cette manière, l'accent est mis sur le caractère incrémental et constructif de la stratégie de résolution, chaque spécialiste apporte un complément d'information susceptible d'être exploité par un autre. Dans tous les cas, l'effort réside dans la mise en commun des informations extraites ou dans l'exploitation, de manière simultanée, de plusieurs points de vue sur une donnée.

Compétitions

Il s'agit de coordonner l'application d'actions ou la façon de les traiter (compétition) pour enrichir l'analyse et exploiter les conflits éventuels en tenant compte à la fois de la complexité des tâches à mettre en oeuvre (problème de la planification) et de la nécessité de préserver la réactivité du système aux situations imprévues (problème de l'opportunisme). La compétition de compétences consiste à produire des solutions (descriptions ou décisions) conflictuelles : celles-ci seront ensuite analysées par des techniques de fusion. Le principe est d'appliquer différents opérateurs sur un même ensemble de données, puis de conserver le meilleur résultat ou un compromis entre les résultats obtenus. Ce mécanisme n'est pas implémenté dans notre système.

La compétition de compétences peut aussi être considérée comme concurrence dans l'accès à une ressource que l'on cherche à s'approprier. C'est la méthode qui se rapproche le plus de celle employée par notre système dans son état actuel. Tous les processus d'analyse (**spécialistes**) communiquent à travers ces primitives de base, se les disputent, les associent ou les font évoluer pour constituer des objets plus complexes. Chaque spécialiste analyse les

données auxquelles il a accès selon son point de vue de manière à infirmer, confirmer ou enrichir la description du document. Pour cela, des zones de focalisation sont obtenues par consultation du modèle de description : recherche d'un élément ou d'une association de plusieurs éléments dans une configuration particulière.

Pour réaliser leur tâche, les différents spécialistes peuvent à tout moment accéder à l'image initiale pour vérifier leurs hypothèses mais les résultats fournis ne sont jamais remis en cause. Il s'agit, ici, d'une première ébauche de solution coopérative qui demande à être complétée; nous y reviendrons dans le chapitre suivant (paragraphe 3.3).

Il nous reste alors à définir les différents spécialistes dont nous venons de parler. Ce sera l'objet du chapitre suivant.

Chapitre 3

Lire et comprendre un document

La méthodologie utilisée dans notre système d'interprétation, rappelons-le, exploite deux phases complémentaires:

- la première consiste en la mise en place d'une vision globale du document par construction du modèle de description (dans son état initial). Il s'agit de la phase de perception globale du document décrite dans le chapitre précédent.

- la seconde, la phase de lecture et d'interprétation du document, est fondée sur un principe de prédiction - vérification d'hypothèses. Les prédictions sont faites en utilisant la vision globale tandis que les vérifications sont faites par des analyses locales. Cette étape, à la fois de lecture et de compréhension, sera décrite dans ce chapitre.

Il conviendra d'abord de préciser la méthodologie proposée à partir des résultats fournis par les études menées sur les mécanismes de lecture et d'interprétation de l'Homme [Bruce93], de décrire ensuite les différents processus d'analyse locale qui doivent être mis en oeuvre dans ce mécanisme, ce qui nous permettra de comprendre comment s'organise leur travail respectif. Nous serons alors en mesure, dans une troisième partie, de décrire le fonctionnement du processus d'analyse contextuelle des schémas (gestion des incohérences - reconstruction 3D) et de présenter les résultats expérimentaux obtenus.

3.1. La méthodologie

Les traitements présentés jusqu'alors étaient adaptés à la plupart des différents types de documents techniques (schémas électriques, cadastres, organigrammes, dessins mécaniques, ...) puisque le modèle de description utilisé doit convenir à la description de n'importe quelle image binaire constituée majoritairement de traits accompagnés de formes pleines et de textes.

Si l'on se réfère de nouveau à l'Homme, il est (à ce niveau) encore possible d'exploiter certaines observations sur le fonctionnement de son système visuel. En effet, une fois l'esquisse primaire obtenue, pour comprendre ce qui l'entoure, le système visuel humain focalise son attention sur des points particuliers de la scène, guidé par la vision globale et par les connaissances accumulées auparavant : il s'agit d'une phase d'exploration visuelle durant

laquelle il est à la recherche d'informations précises, dont la présence est supposée exister à des endroits donnés, et sur lesquelles il possède aussi des connaissances [Denhière92].

L'idée sous-jacente que l'on a décidé d'adapter dans notre système, est de tirer parti des informations apportées par le modèle de description, pour extraire les entités contenues dans le document, (de quelque niveau d'abstraction qu'elles soient) de la manière la plus robuste possible, rejetant les éléments dont la certitude est insuffisante [Ramel96].

On remarquera en outre, en analysant des exemples, que l'architecture choisie permet une adaptation facile du système à différents types de documents (bien que cette étude concerne plus particulièrement les schémas cinématiques mécaniques). Pour que cela soit possible, il a été nécessaire de mettre en place une répartition des connaissances intelligentes pouvant être aisément complétées ou modifiées.

Comme nous l'avons déjà laissé entendre, cette distribution des connaissances dans le système se base sur l'utilisation d'un ensemble de processus indépendants les uns des autres, experts chacun dans un domaine particulier, et réalisant une tâche précise en fonction des connaissances qu'il possède. Ces processus contiennent les connaissances sur le domaine (c'est à dire particulières à un type de document) et une partie de leur travail consiste à les coupler avec les informations globales fournies par le modèle de description. A ces différents processus nous donnerons, tout au long de ce mémoire, le nom de **spécialistes**.

Le modèle de description constitue la source d'information commune, partagée par les spécialistes : c'est le vecteur de communication évoluant grâce aux travaux des spécialistes jusqu'à, finalement, traduire la compréhension totale du document.

Grâce à cette architecture, il suffit d'ajouter ou d'enlever des spécialistes pour apporter de nouvelles connaissances au système et donc l'adapter à l'interprétation d'un autre type de document ou améliorer ses performances. Par exemple, dans le cas de l'analyse de partitions musicales, un spécialiste aura en charge l'extraction des têtes de notes.

Un spécialiste ne communique jamais directement avec les autres, il utilise obligatoirement le modèle de description dont il peut consulter l'état à un instant donné, et qu'il peut analyser de son point de vue pour le faire évoluer si cela est possible. Cette évolution est hiérarchique : le travail de certains spécialistes ne peut commencer qu'une fois que d'autres ont entamé le leur. L'ordre de localisation et de reconnaissance des entités découle de la facilité et de la sûreté avec lesquelles les travaux d'analyse locale de l'image peuvent être exécutés (d'ailleurs souvent en rapport avec le nombre d'occurrences d'un objet dans le document).

La puissance du système découle aussi du fait que tous les spécialistes n'extraient que les évidences pour faire se développer **progressivement** le contexte de manière à créer de nouvelles évidences pour d'autres spécialistes, et finalement aboutir à l'interprétation du document avec un minimum d'**erreurs de lecture**.

Grâce à cette technique, les entités faciles à détecter sont localisées et reconnues les premières, lorsque le contexte apporte encore peu d'information et que cette information est de bas niveau sémantique donc mal structurée et trop abondante. Tandis que, les entités difficiles à déchiffrer sont détectées les dernières, lorsque le modèle de description a atteint un niveau sémantique suffisant, c'est à dire quand les données sont davantage structurées et réduites en quantité mais de meilleure qualité.

3.2. Les Processus d'Analyse Locale

3.2.1. L'extraction du texte

La plupart des systèmes de lecture automatique de documents techniques utilisent uniquement la taille et la position des composantes connexes pour extraire les zones de texte d'un document [Lai93]. Des méthodes lourdes ont été implémentées pour tenter de détecter les caractères cachés (caractères en contact avec d'autres parties du dessin), la plupart sont basées sur des études de l'alignement et de l'espacement entre composantes [Desseilligny94]. Mais, comme nous l'avons vu dans le chapitre 1, elles ne sont pas encore suffisamment efficaces et ne résolvent pas les problèmes posés par le manuscrit (caractères liés les uns aux autres).

De récents travaux en OCR [Elliman90, Lorette92] ont prouvé qu'il était possible de distinguer les zones textuelles d'un document par l'analyse de sa texture. On a pu en effet remarquer, sur la figure 2.26c, que les caractères se présentent sous la forme de nombreux quadrilatères de faibles longueurs et d'orientations variées. Ainsi, les attributs définis dans notre modèle, la taille, la forme et l'agencement des quadrilatères constituant la suite SQ sont bien représentatifs de la texture de l'image initiale.

Les zones de texte se caractérisent par des regroupements de quadrilatères de petite taille. On le remarque sur les figures 3.1a₂ et 3.1b₂ où n'ont été représentés que les quadrilatères ayant une longueur inférieure à 150 pixels. Le spécialiste chargé d'extraire le texte commence donc

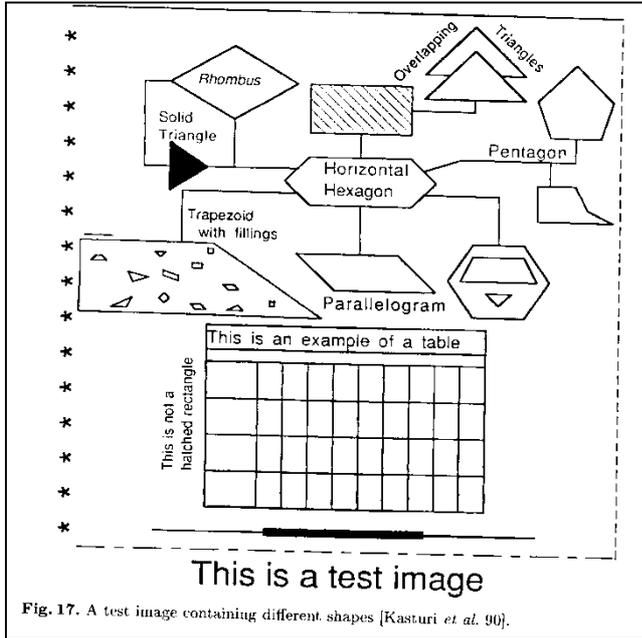
par étudier l'ensemble des quadrilatères en vue d'extraire de SQ les «amas» de quadrilatères de petite taille susceptibles de correspondre aux zones de texte. Un amas est un ensemble comprenant au moins deux petits quadrilatères distants de moins de 50 pixels. Cette focalisation par génération d'hypothèses n'est pas suffisante pour conclure car d'autres formes : arcs de faible rayon de courbure, pointillés, ... génèrent aussi des quadrilatères répondant à ces mêmes critères (faibles longueurs).

Dans un premier temps, une simple analyse locale (étude des valeurs prises par les différents attributs des quadrilatères constituant les amas construits) permet d'éliminer facilement certaines des zones (hachures, pointillés où les quadrilatères sont de même direction, bruits où les quadrilatères sont très isolés); les amas correspondants sont alors rejetés.

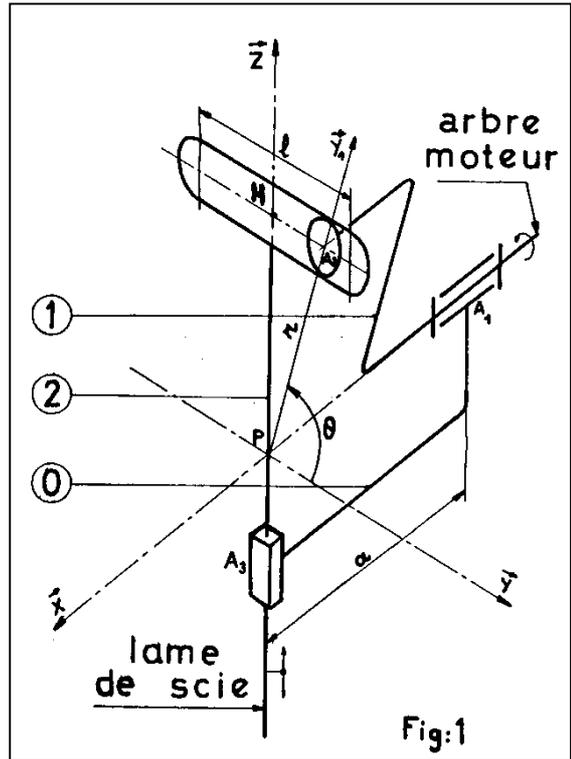
Les regroupements qui subsistent après ce premier élagage constituent les zones susceptibles de contenir du texte (aussi bien imprimé que manuscrit).

Lors d'une seconde étape, les informations fournies par le modèle de représentation (zones de focalisation) avec hypothèse de présence de texte sont fusionnées avec celles fournies par l'image initiale dont on a extrait les composantes connexes. Dans de nombreux cas, notamment pour les caractères imprimés de taille moyenne, cette fusion d'informations permet de prendre une décision finale. Si les regroupements de quadrilatères correspondent à des composantes connexes (amas et composantes de tailles comparables) on peut considérer comme zone de texte, la composante. Ainsi, on voit sur la figure 3.2 les résultats de cette fusion correspondant respectivement aux images a1 et b1 de la figure 3.1. On peut noter en particulier que la figure a3 ne contient plus, ni les zones de hachures du haut, ni les pointillés de droite. Par contre les étoiles à gauche peuvent être des caractères.

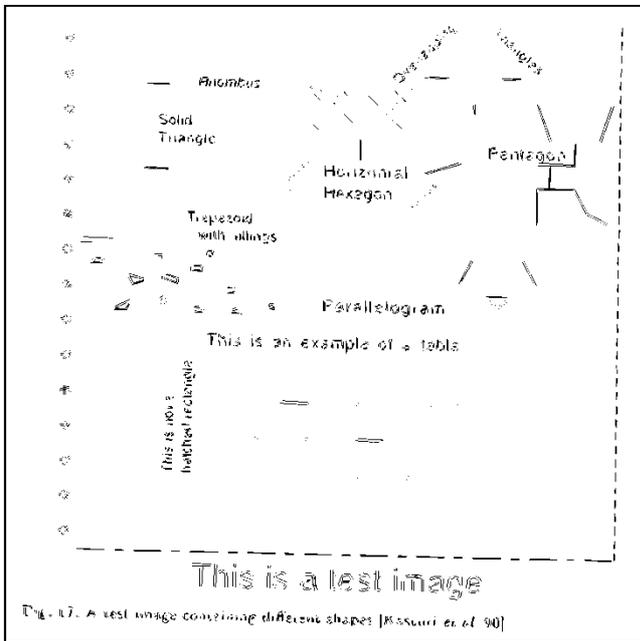
Les primitives de description incluses dans ces zones sont alors remplacées par des entités de niveau supérieur qui indiqueront la présence du texte. Ainsi, progressivement, le modèle de description évolue et se simplifie.



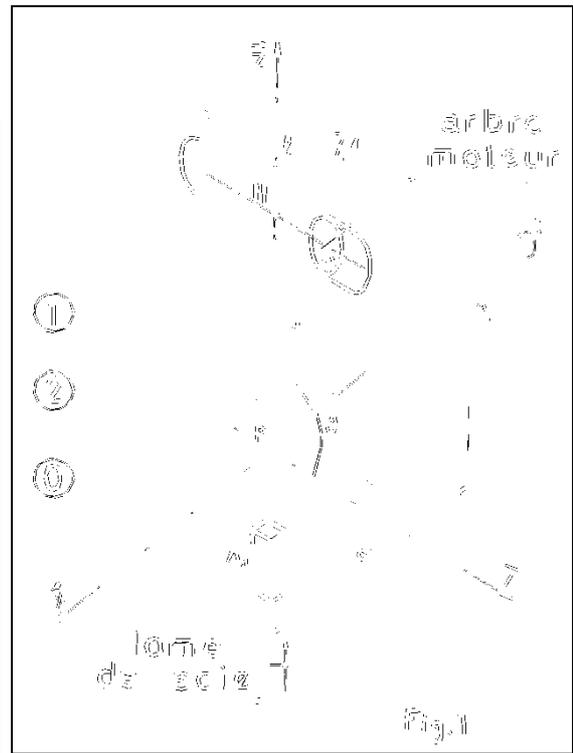
a₁



b₁



a₂



b₂

Figure 3.1 : Quadrilatères susceptibles de composer un texte

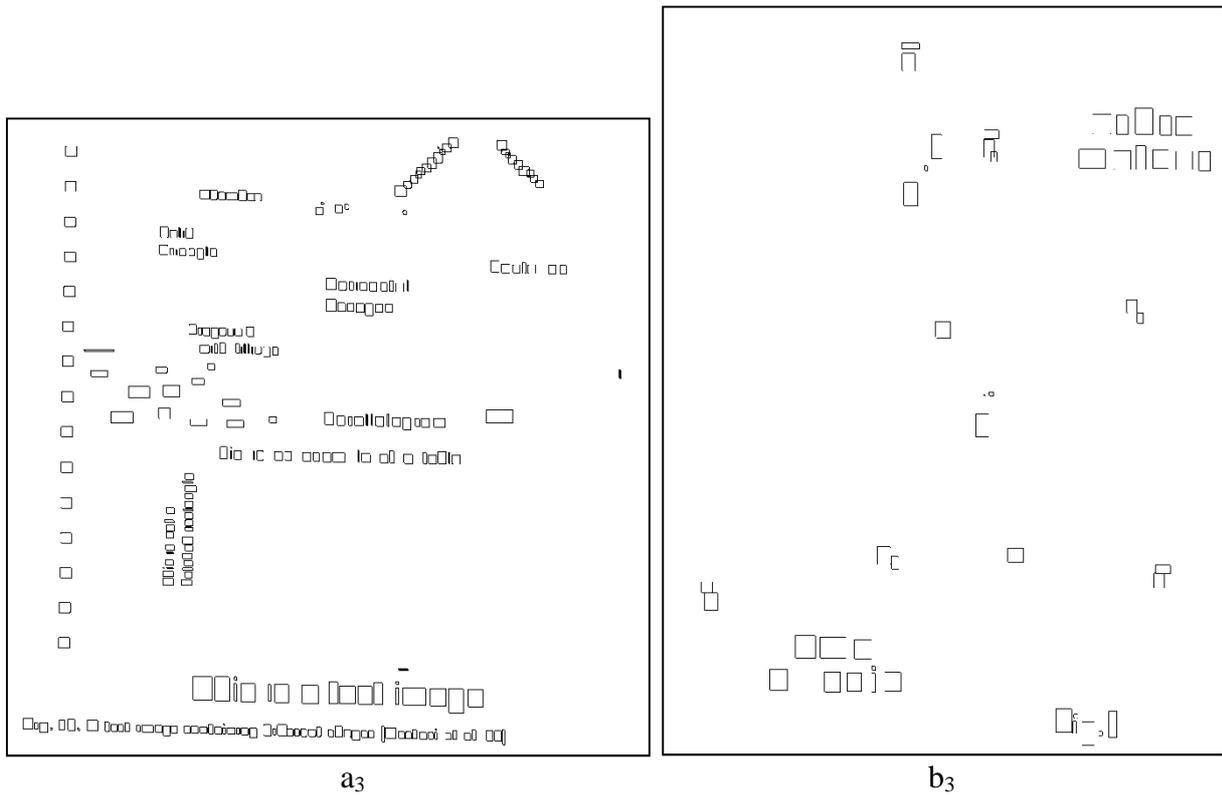


Figure 3.2 : Caractères facilement localisés

Les amas qui subsistent, trop différents des composantes connexes, correspondent à des zones ambiguës et doivent subir une étude plus poussée (analyse locale).

A partir de chaque amas de quadrilatères, il est possible de générer une **composante de texte virtuelle** (zone de focalisation) constituée par le rectangle englobant l'ensemble des quadrilatères regroupés pour former l'amas; la figure 3.3 montre quelques exemples de zones où la décision n'a pas pu être prise



Figure 3.3 : Composantes de texte virtuelles

En étudiant la taille et la position par rapport aux zones de texte déjà localisées, on procède à une confrontation avec le contexte, et certaines ambiguïtés peuvent alors être levées. Enfin dans les cas où le doute subsiste, on peut faire appel à un module d'OCR ou interroger l'utilisateur pour prendre d'autres décisions.

Plusieurs activations successives de ce spécialiste ne donnent pas des résultats identiques. Si le modèle de description a évolué entre les activations, les composantes de texte virtuelles peuvent être différentes. Nous avons choisi, dans un premier temps, de consulter l'utilisateur pour toutes les zones ambiguës ; les composantes de texte virtuelles construites sont présentées à l'expert qui confirme la présence de texte ou infirme l'hypothèse. La figure 3.4 présente les zones textuelles finalement extraites des images initiales présentées, respectivement figure 3.4a et figure 3.1b₁. Le texte en grisé dénote les zones pour lesquelles confirmation a été demandée à l'utilisateur.

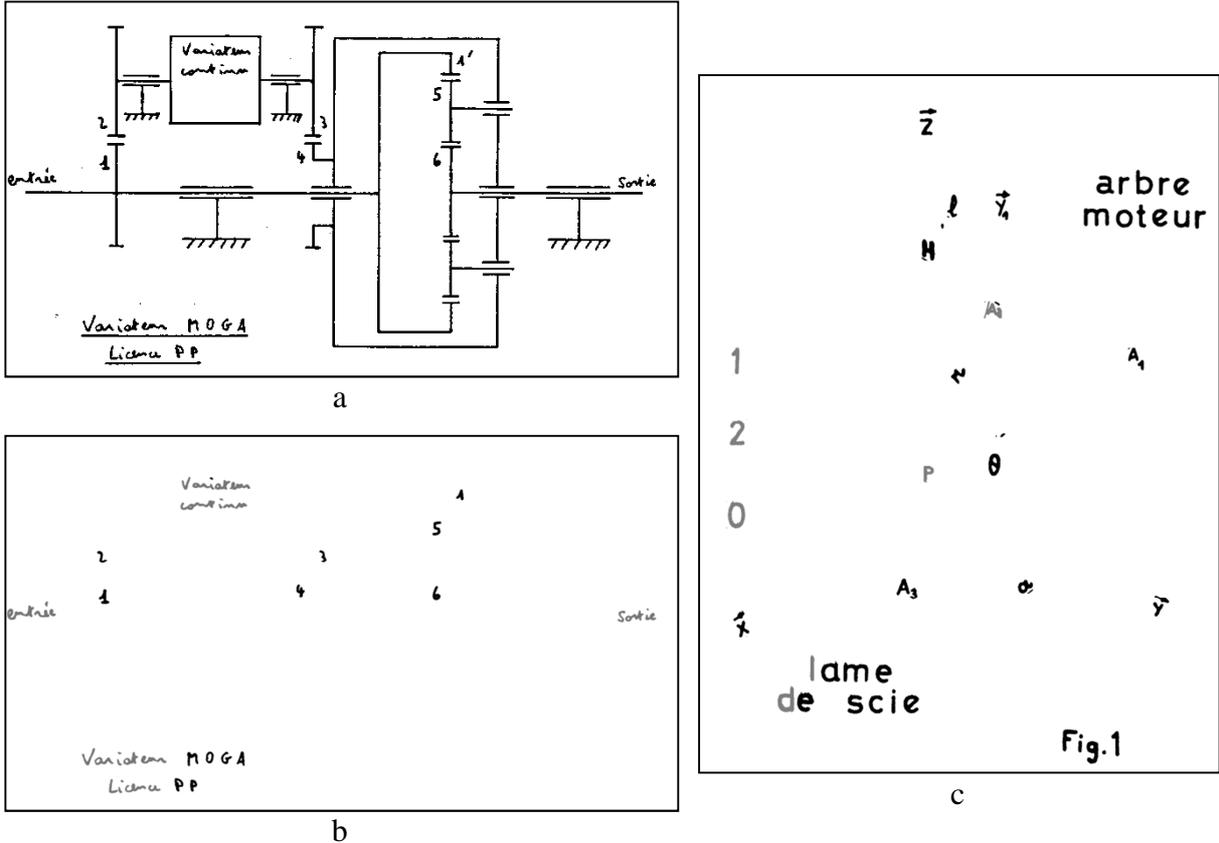


Figure 3.4 : Zones de texte : b de l'image a
 c de l'image présentée figure 3.1b₁

3.2.2. Extraction des lignes

Le spécialiste chargé d'analyser les lignes utilise lui aussi l'ensemble des quadrilatères SQ pour réaliser son travail. Il a pour tâche de déterminer et de caractériser le type des différents traits constituant le document parmi 4 types possibles que nous avons sélectionnés en fonction des documents étudiés :

- Traits continus
- Axes principaux (repère)
- Traits pointillés réguliers (.....)
- Traits pointillés mixtes (.....)

Les traits continus

Dans un premier temps, ce spécialiste devait réaliser l'étape de vectorisation complète c'est à dire résoudre les problèmes cités dans le paragraphe 2.3.1.3 : pour améliorer la détection des lignes droites présentes dans le document [Ramel95].

Dans ce but, chaque quadrilatère du modèle de description est considéré comme un générateur de ligne que l'on essaierait d'allonger pour produire une ligne la plus longue possible. Ce processus appliqué sur chaque quadrilatère se décompose en 5 étapes dont les 4 premières sont illustrées figure 3.5 :

① Sélection d'un *quadrilatère* générateur de la ligne.

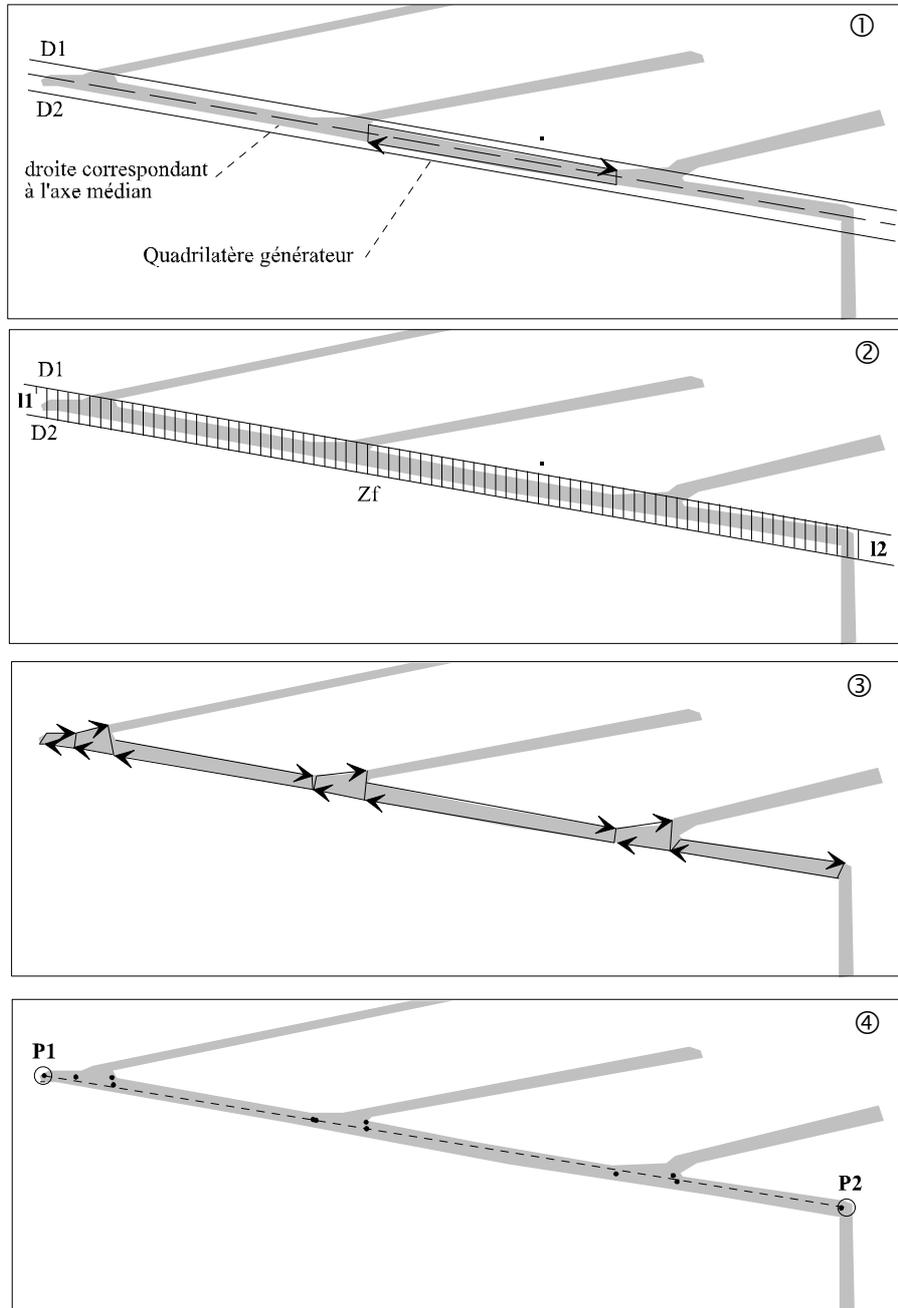
On choisit le quadrilatère de longueur maximum et on construit une zone de focalisation Zf. L'axe médian du quadrilatère définit une pente qui, associée aux attributs d'épaisseur, permet d'établir deux équations de droites. Ces droites délimitent la zone Zf dans laquelle est susceptible de se prolonger la ligne.

② En se référant à l'image initiale, le spécialiste localise les extrémités de la ligne à l'intérieur de la zone d'étude.

③ Tous les quadrilatères appartenant à la zone définie (Zf) sont sélectionnés.

④ L'équation de la droite finale est générée par application de la méthode des moindres carrés appliquée sur les extrémités des axes médians des quadrilatères sélectionnés en ③.

⑤ Les *quadrilatères* sélectionnés en ③ sont enlevés de SQ et les caractéristiques de la ligne globale sont sauvegardées.



(1) Quadrilatère générateur et ses droites associées. (2) Construction de la zone d'étude Zf.

(3) Sélection des quadrilatères. (4) Trait obtenu

Figure 3.5 : Extraction des lignes droites

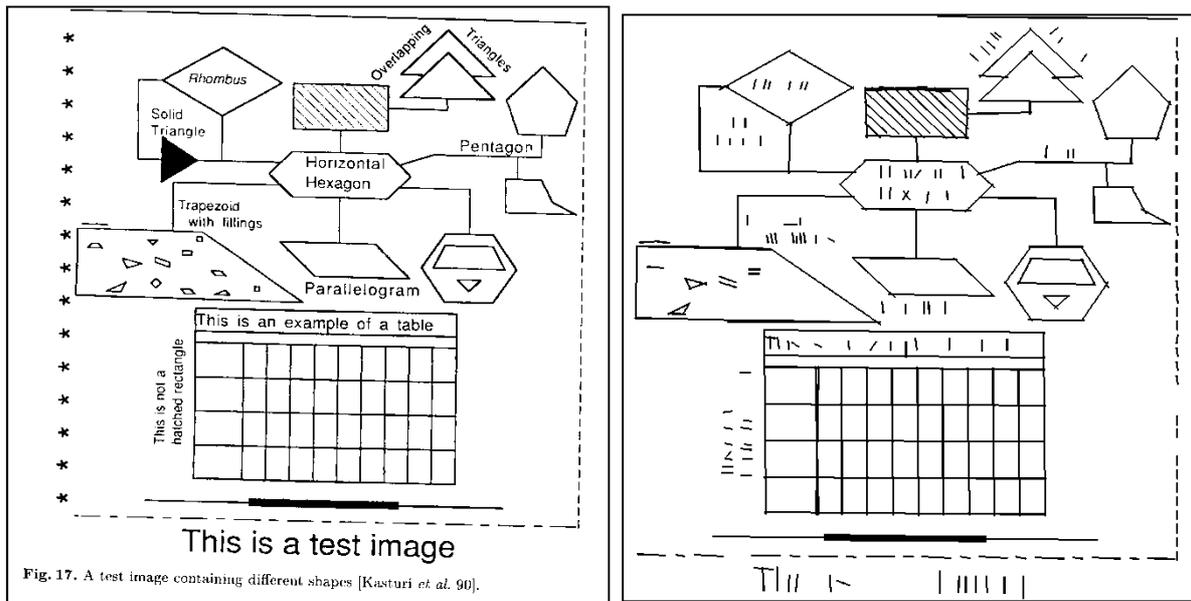


Figure 3.6 : Image des lignes droites

Le résultat final apparaît sur la figure 3.6. Comme pour les autres éléments, l'ensemble des lignes droites détectées est sauvegardé sous forme d'un ensemble d'objets. Chaque objet est défini par les coordonnées de ses 2 points extrémités, sa longueur, sa pente et son épaisseur.

Les axes principaux

Un second procédé a été testé pour, à partir des informations fournies par les quadrilatères, extraire les longues lignes droites contenues dans l'image. Cette fois-ci, la méthode repose sur un changement de représentation (passage des coordonnées cartésiennes aux coordonnées polaires).

A chaque quadrilatère du modèle de description peuvent être associés deux paramètres correspondant aux coordonnées polaires de son axe médian (figure 3.7).

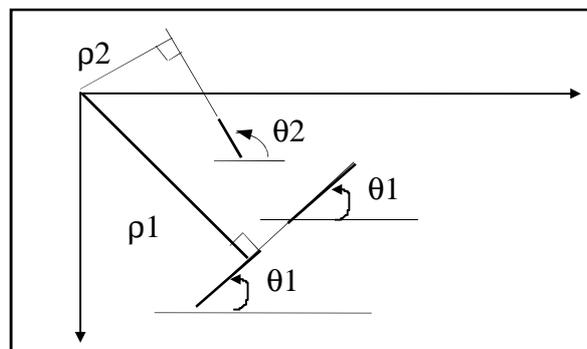


Figure 3.7 : Changement de représentation

Nous avons adapté à notre cas l'algorithme de la transformée de Hough pour l'extraction de droites. L'utilisation d'un accumulateur fournit des renseignements sur les longueurs des lignes droites constituant le document.

Pour chaque quadrilatère (ρ, θ) , l'entrée correspondante de l'accumulateur est incrémentée de la longueur du quadrilatère. Une fois les quadrilatères passés en revue, un pic dans l'accumulateur correspond à une droite dans l'image.

Les 3 axes (Ox, Oy, Oz) correspondent généralement aux 3 droites les plus longues du dessin et constituent le repère dans l'image. Afin de les localiser, nous avons choisi d'analyser le contenu de l'accumulateur.

Nous avons choisi de construire les 3 droites définies par les 3 pics maximums de l'accumulateur. Nous nous sommes vite rendu compte, d'une part que ces pics ne définissaient pas toujours les axes du dessin, et surtout qu'il était très difficile de choisir le pas de discrétisation de l'accumulateur pour faire varier le paramètre ρ . L'application d'un lissage gaussien [Benyacoub95] sur cet accumulateur aurait peut-être réduit cette sensibilité aux variations de ρ (pour des quadrilatères constituant une même droite). Cependant l'information fournie par les axes étant, dans notre cas, non fondamentale, nous avons renoncé à continuer dans cette voie pensant qu'il était peut-être plus judicieux d'utiliser les têtes de flèches pour localiser ces éléments, et que l'utilisation du contexte permettait ainsi de se focaliser sur certaines directions.

La technique décrite ci-dessus semble néanmoins intéressante pour d'autres types de documents comme, par exemple, les partitions musicales où la détection des lignes horizontales (les portées) et la détection des lignes verticales (les queues des notes et les barres de mesure) ont une très grande importance.

Pour l'interprétation des schémas cinématiques, les résultats obtenus avec ces méthodes (traits continus, axes principaux) ne nous ont finalement pas satisfait. Ils n'étaient, en particulier, pas du tout adaptés aux dessins comportant des courbes ; ces dernières étant transformées en une suite de segments difficilement exploitable. En fait, nous avons constaté que les résultats fournis par la suite SQ étaient difficilement améliorables et qu'il semblait donc plus opportun de ne pas s'acharner à tenter d'obtenir une vectorisation parfaite. Il était bien préférable de mettre en place des traitements capables de travailler à partir des informations fournies par le modèle de description initial.

Les traits pointillés

Les techniques précédentes ne sont cependant pas restées totalement inutilisées car les procédés décrits précédemment peuvent être adaptés efficacement à la détection des lignes de traits interrompus.

Les traits interrompus sont constitués d'une suite de quadrilatères de petites tailles (longueurs) alignés. Aussi, l'algorithme de détection des lignes continues par utilisation des paramètres polaires a été repris. Toutefois, la phase 2 a été améliorée : Zf reste délimitée par les 2 droites parallèles à l'hypothétique ligne, mais aucune restriction de longueur n'est envisagée. Les bords de l'image limitent Zf. On a également modifié le critère de choix des quadrilatères générateurs. Ce sont, cette fois, ceux ayant une faible longueur ($0 < L_g < 150$ pixels) qui sont utilisés. Les quadrilatères appartenant à Zf font ensuite l'objet d'une étude plus précise afin de caractériser le type de chaque ligne.

En fonction de la pente du quadrilatère générateur, les quadrilatères sélectionnés (éléments de Zf) sont classés suivant leur projection sur un axe vertical comme dans le cas 1 de la figure 3.8 ou sur un axe horizontal dans le cas 2.

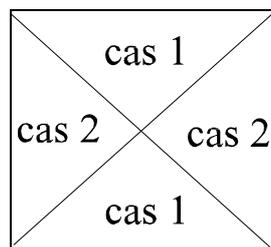


Figure 3.8 : 2 cas suivant la pente du générateur

Une analyse de la succession Trait Noir / Zone Blanche, c'est à dire l'étude de la position des extrémités des quadrilatères successifs et de leur longueur (minimum, maximum, moyenne), permet ensuite de caractériser la ligne en cours d'étude (traits pointillés réguliers, traits pointillés mixtes, traits continus). Un trait pointillé est considéré comme régulier tant que la longueur du tiret de longueur maximum reste inférieure à 2 fois la longueur du tiret le plus court. De plus, un trait pointillé est constitué d'une suite comprenant au moins 4 quadrilatères.

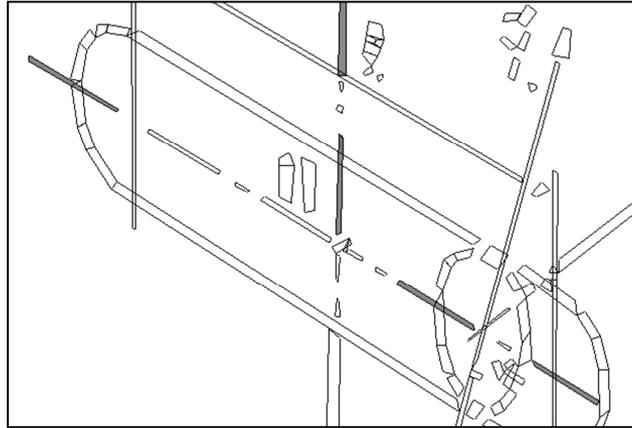


Figure 3.9 : Quadrilatères ambigus

L'algorithme est à même d'extraire les lignes pointillées ayant une partie cachée par un trait plein comme on le voit sur la figure 3.10; en effet, l'accumulation des preuves de présence de trait permet de ne pas tenir compte des positions de chaque quadrilatère.

La position et les caractéristiques des lignes pointillées sont sauvegardées, les quadrilatères qui les constituaient sont soit effacés (élément non ambigu), soit étiquetés comme constituants d'une ligne pointillée (élément ambigu). Un élément est considéré comme ambigu s'il a une intersection non vide avec une autre primitive du modèle de description; ces éléments ont été indiqués en grisés sur la figure 3.9.

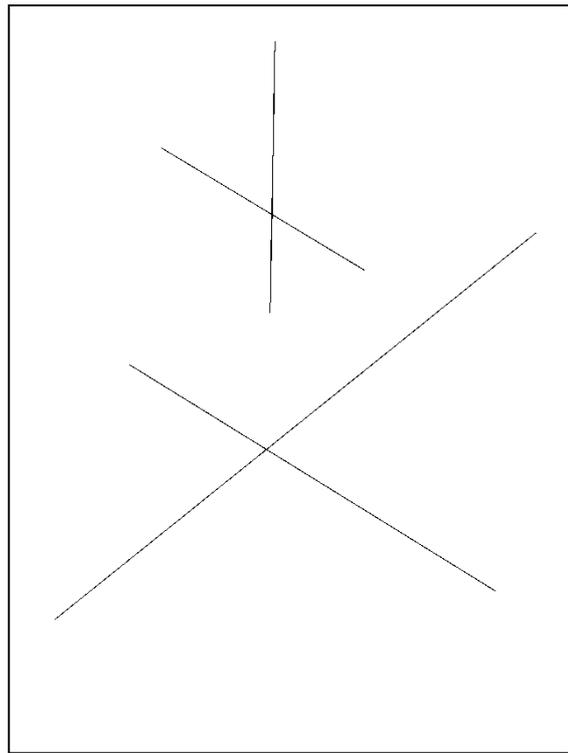
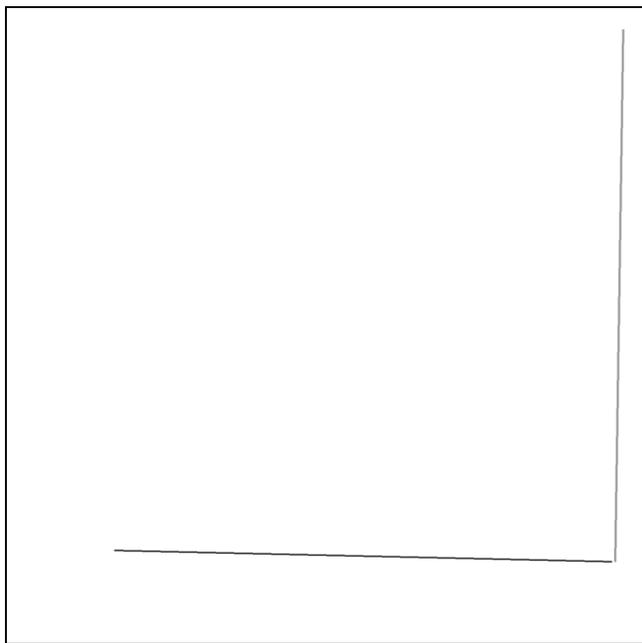
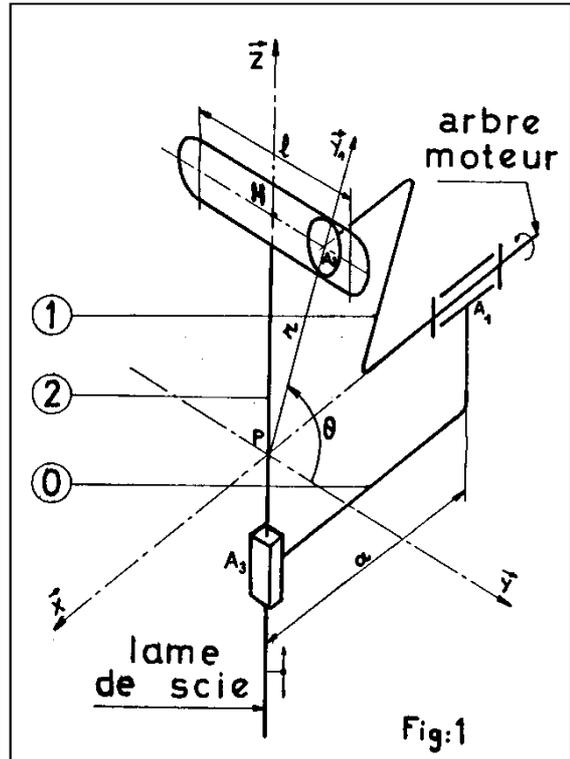
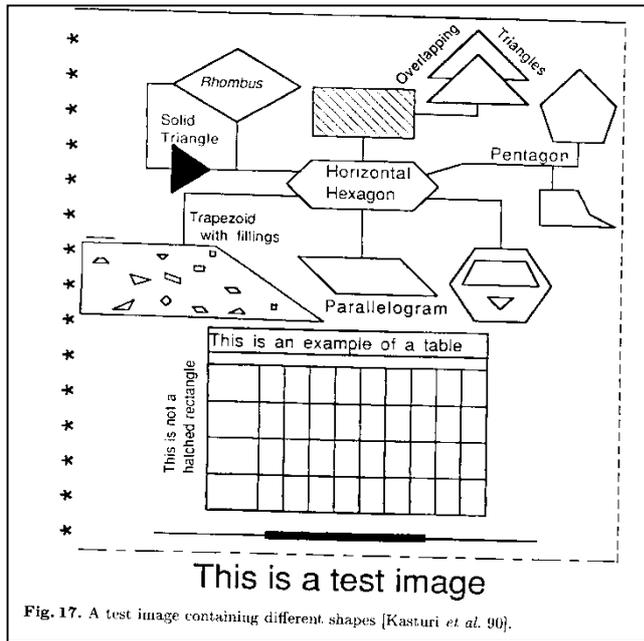


Figure 3.10 : Pointillés détectés (réguliers en gris, mixtes en noir)

La détection des **courbes** (cercles, arcs, ...) en pointillés n'a pas été implémentée puisque aucun des dessins cinématiques fournis n'en comportait. D'autres méthodes de localisation des lignes et courbes en pointillés sont développées dans [Kasturi90, Lai91].

3.2.3. Extraction des courbes

Le travail de ce spécialiste consiste à localiser les courbes contenues dans le schéma. Le problème est de modéliser une courbe connue par un grand nombre de points, à l'aide d'une expression analytique qui la représente le plus fidèlement possible. Parmi toutes les possibilités offertes, nous avons choisi d'utiliser les polynômes de Bézier, d'une part parce qu'ils sont largement utilisés dans le domaine mécanique, et d'autre part à cause de leur simplicité de développement logiciel comparé à d'autres techniques utilisées en CAO.

Quelques Rappels [Besançon88]

Soit une courbe C définie dans un espace rapporté à un repère orthonormé-3D $(O; \vec{i}, \vec{j}, \vec{k})$. Bézier propose de se placer dans un espace à n dimensions et de caractériser la courbe par la suite des a_i ($a_i \in \mathbf{R}$) et les fonctions $f_{i,n}$ suivantes ($f_{i,n} \in \mathcal{C}([0;1], \mathbf{R})$) qui définissent une paramétrisation de C :

$$[0,1] \rightarrow \mathbf{E}_3 \quad t \rightarrow P(t)$$

$$P(t) = a_0 + \sum_{i=1}^n a_i \cdot f_{i,n}(t)$$

On appelle polygone de Bézier, le polygone dont les sommets ont a_i pour coordonnées. C'est le polygone des points-guides.

Les fonctions $f_{i,n}$ pour $t \in [0;1]$ ont les propriétés suivantes :

- $U(0)$ est à l'origine du référentiel considéré si $P(t) = a_0 + U(t)$

- $U(1) = (1,1,1)$

- la tangente au point de coordonnées $U(0)$ est $0x$; au point de coordonnées $U(1)$, la tangente est parallèle à Oz .

Ainsi la courbe de la figure 3.11 ayant $AMNB$ comme polygone caractéristique, admet A comme origine et B pour extrémité. Ses tangentes en A et B sont respectivement AM et NB .

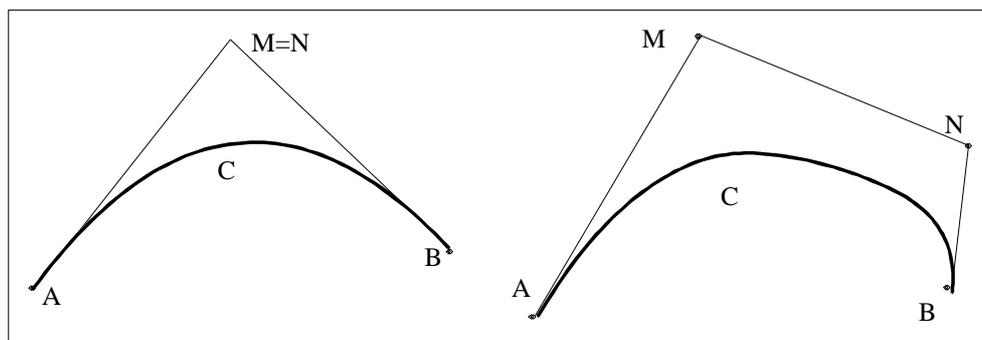


Figure 3.11 : Courbes et polygones caractéristiques

Bézier a choisi comme fonction $f_{i,n}$ les polynômes de Bernstein B_i^n par l'expression :

$$P(t) = a_0 \cdot B_0^n(t) + a_1 \cdot B_1^n(t) + a_2 \cdot B_2^n(t) + \dots + a_n \cdot B_n^n(t) \quad \text{pour le degré } n$$

où les polynômes de Bernstein s'écrivent :

$$B_i^n(t) = \binom{n}{i} \cdot t^i \cdot (1-t)^{n-i} \quad \text{avec } i \in [0, n]$$

Les constantes a_i sont les coordonnées des points-guides (ou pôles) de la portion de courbe C et elles forment les sommets du polygone. Les polynômes de Bernstein sont toujours positifs pour $t \in [0;1]$ ce qui entraîne que la courbe est incluse dans chaque partie convexe du polygone.

Dans les problèmes concernant les images (2D), les polynômes de Bézier sont souvent introduits sous la forme paramétrique suivante :

$$x = P_x(t) = \sum_{i=0}^n \binom{n}{i} t^i \cdot (1-t)^{n-i} \cdot p_i \quad \text{où } \begin{pmatrix} p_i \\ q_i \end{pmatrix} \text{ sont les coordonnées des pôles du polynôme}$$

$$y = P_y(t) = \sum_{i=0}^n \binom{n}{i} t^i \cdot (1-t)^{n-i} \cdot q_i$$

ce qui donne sous forme matricielle :

$$P(t) = \sum_{i=0}^n \binom{n}{i} t^i \cdot (1-t)^{n-i} \cdot P_i$$

Approximation d'une courbe par une courbe de Bézier

Soit $M_k(x_k, y_k)$ $k \in [0; m]$ l'ensemble des points constituant la courbe C à approximer et C_b une courbe de Bézier approximant C . On désire trouver la courbe de Bézier de degré n , telle que l'on puisse trouver $m+1$ points H_k de paramètre t_k sur la courbe de Bézier $(P_x(t_k), P_y(t_k))$, les plus proches possible de M_k :

On note p_i et q_i les coefficients tels que :

$$P_x(t_k) = \sum_{i=0}^n B_i^n(t_k) \cdot p_i$$

$$P_y(t_k) = \sum_{i=0}^n B_i^n(t_k) \cdot q_i$$

Pour cela, on minimise l'erreur cumulée $h = \sum_{k=0}^m \overline{M_k H_k}^2$ qui est alors fonction de p_0, \dots, p_n et q_0, \dots, q_n , les points M_k étant des données.

Pour cela on cherche les meilleurs (p_i, q_i) pour $i \in [0; n]$

$$h(p_0, \dots, p_n, q_0, \dots, q_n) = \sum_{k=0}^m \left[x_k - P_x(t_k) \right]^2 + \left[y_k - P_y(t_k) \right]^2$$

On recherche les minimums parmi les points critiques de la fonction h , c'est à dire les points où la différentielle est nulle, soit à résoudre le système:

$$\frac{\partial h}{\partial p_i}(p_0, \dots, p_n, q_0, \dots, q_n) = \sum_{k=0}^m \left[-2x_k \cdot B_i^n(t_k) + 2 \cdot p_i \cdot (B_i^n(t_k))^2 + \sum_{j \neq i} 2p_j B_i^n(t_k) B_j^n(t_k) \right] = 0$$

i variant de 0 à n

$$\frac{\partial h}{\partial q_i}(p_0, \dots, p_n, q_0, \dots, q_n) = \sum_{k=0}^m \left[-2y_k \cdot B_i^n(t_k) + 2 \cdot q_i \cdot (B_i^n(t_k))^2 + \sum_{j \neq i} 2q_j B_i^n(t_k) B_j^n(t_k) \right] = 0$$

i variant de 0 à n

qui conduit à résoudre 2 systèmes indépendants de $n+1$ équations à $n+1$ inconnues :

$$\sum_{j=0}^n p_j \cdot \left[\sum_{k=0}^m B_j^n(t_k) \cdot B_i^n(t_k) \right] = \sum_{k=0}^m x_k B_i^n(t_k) \text{ pour } i = 0 \text{ à } n$$

et
$$\sum_{j=0}^n q_j \cdot \left[\sum_{k=0}^m B_j^n(t_k) \cdot B_i^n(t_k) \right] = \sum_{k=0}^m y_k B_i^n(t_k) \text{ pour } i = 0 \text{ à } n$$

Il est possible d'écrire ces 2 systèmes sous forme matricielle :

$\mathbf{A} \cdot \mathbf{P}_x = \mathbf{C}_x$ et $\mathbf{A} \cdot \mathbf{P}_y = \mathbf{C}_y$ avec :

$$\mathbf{A} = \mathbf{B} \cdot {}^t\mathbf{B} \quad \mathbf{C}_x = \mathbf{B} \cdot \mathbf{X} \quad \mathbf{C}_y = \mathbf{B} \cdot \mathbf{Y}$$

$$\begin{array}{c}
 n+1 \\
 \left| \right.
 \end{array}
 \mathbf{B} = \overbrace{\begin{pmatrix}
 \mathbf{B}_0^n(t_0) & \mathbf{B}_0^n(t_1) & \dots & \mathbf{B}_0^n(t_m) \\
 \mathbf{B}_1^n(t_0) & \mathbf{B}_1^n(t_1) & \dots & \mathbf{B}_1^n(t_m) \\
 \cdot & & & \cdot \\
 \cdot & \dots & \dots & \cdot \\
 \cdot & & & \cdot \\
 \mathbf{B}_n^n(t_0) & \mathbf{B}_n^n(t_1) & \dots & \mathbf{B}_n^n(t_m)
 \end{pmatrix}}^{m+1}
 \quad
 \mathbf{X} = \begin{pmatrix}
 x_0 \\
 x_1 \\
 \cdot \\
 \cdot \\
 \cdot \\
 x_m
 \end{pmatrix}
 \quad
 \mathbf{P}_x = \begin{pmatrix}
 p_0 \\
 p_1 \\
 \cdot \\
 \cdot \\
 \cdot \\
 p_n
 \end{pmatrix}$$

$$\mathbf{Y} = \begin{pmatrix}
 y_0 \\
 y_1 \\
 \cdot \\
 \cdot \\
 \cdot \\
 y_m
 \end{pmatrix}
 \quad
 \mathbf{P}_y = \begin{pmatrix}
 q_0 \\
 q_1 \\
 \cdot \\
 \cdot \\
 \cdot \\
 q_n
 \end{pmatrix}$$

X et Y représentent les coordonnées des m+1 points successifs connus sur la courbe à approximer.

B est la matrice des polynômes de Bernstein jusqu'au degré n calculés en m+1 points.

\mathbf{P}_x et \mathbf{P}_y sont les inconnues, ces matrices correspondent aux coordonnées des pôles du polynôme de Bézier décrivant la courbe.

Existence d'une solution aux systèmes $\mathbf{A} \cdot \mathbf{P} = \mathbf{C}$ avec $\mathbf{P} \in \mathbf{R}^{n+1}$:

Il existe une solution au système si A est inversible c'est à dire si $\det(\mathbf{A}) \neq 0$

A est symétrique et réelle. On peut remarquer que $\mathbf{A} = \mathbf{B} \cdot {}^t\mathbf{B}$. A est une matrice de Gram de rang n+1.

Nous allons maintenant prouver que l'unique point critique obtenu correspond à un minimum de la fonction h. Pour cela il suffit de constater que la matrice hessienne de h est définie positive.

$$\frac{\partial^2 h}{\partial p_i^2} = 2 \cdot \sum_{k=0}^m (\mathbf{B}_i^n(t_k))^2 \quad \text{et} \quad \frac{\partial^2 h}{\partial p_i \partial p_j} = 2 \cdot \sum_{k=0}^m \mathbf{B}_i^n(t_k) \cdot \mathbf{B}_j^n(t_k)$$

$$\mathcal{H} = \left(\frac{\partial^2 h}{\partial p_i \partial p_j} \right)_{i,j} \quad \text{avec } 0 \leq i \leq n \text{ et } 0 \leq j \leq n$$

$$\text{On voit que } \mathcal{H} = 2 \cdot {}^t B \cdot B = \begin{pmatrix} B_0^n(t_0) & B_0^n(t_1) & \dots & B_0^n(t_m) \\ B_1^n(t_0) & B_1^n(t_1) & \dots & B_1^n(t_m) \\ \cdot & & & \cdot \\ \cdot & B_1^n(t_j) & & \cdot \\ \cdot & & & \cdot \\ B_n^n(t_0) & B_n^n(t_1) & \dots & B_n^n(t_m) \end{pmatrix} \begin{pmatrix} B_0^n(t_0) & B_1^n(t_0) & \dots & B_n^n(t_0) \\ B_0^n(t_1) & B_1^n(t_1) & \dots & B_n^n(t_1) \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ B_0^n(t_m) & B_1^n(t_m) & \dots & B_n^n(t_m) \end{pmatrix}$$

$X^t \cdot \mathcal{H} \cdot X = 2 \cdot {}^t X \cdot B \cdot B \cdot X = 2^t ({}^t B X) \cdot ({}^t B X) > 0$ ($X \neq 0$) donc la forme quadratique associée à \mathcal{H} est définie positive, et h présente un extremum qui est un minimum.

Méthode d'extraction des courbes

Avant d'être approximés à l'aide d'une courbe de Bézier, les ensembles de points susceptibles de constituer des courbes doivent être localisés dans l'image. Les courbes sont des formes fines ; l'ensemble des primitives quadrilatères du modèle de description de l'image traduit donc leur présence. De même que pour les caractères, une courbe dans le document, de par sa nature, correspond à une succession de quadrilatères de faible longueur dans le modèle de représentation. On remarquera que les courbes de Bézier ne sont pas adaptées à la modélisation des courbes contenant de longs segments de droite. Le premier travail du spécialiste est donc de sélectionner toutes les suites de quadrilatères pouvant être représentatives d'une courbe. Pour ne pas alourdir les calculs et de manière empirique, la longueur des quadrilatères constituant la suite a été limitée à 150 pixels.

Pour pouvoir être approximée par une courbe de Bézier, chaque courbe est représentée par une suite de points $M_k (x_k, y_k)$ suivant l'abscisse curviligne croissante. A ces points, on tentera de faire correspondre les points $H_k (P_x(t_k), P_y(t_k))$: points de la courbe de Bézier approximante. Les points M_k sélectionnés sont les extrémités des axes médians des quadrilatères susceptibles de constituer une courbe. Comme précédemment, les quadrilatères sont ordonnés (selon la proximité entre leurs extrémités) de manière à reconstruire la chronologie du tracé de la courbe. Grâce à cette technique, il est possible d'associer un t_k à chaque extrémité des axes médians (figure 3.12).

La courbe est considérée comme un signal continu observable dans le temps à différents instants correspondant aux points M_k précédents. Les temps t_k d'observation associés aux

différents points sont calculés en fonction de la longueur à parcourir pour passer d'un point à l'autre dans un mouvement uniforme. Les temps $t_0 = 0$ et $t_m = 1$ correspondent aux 2 extrémités de la courbe.

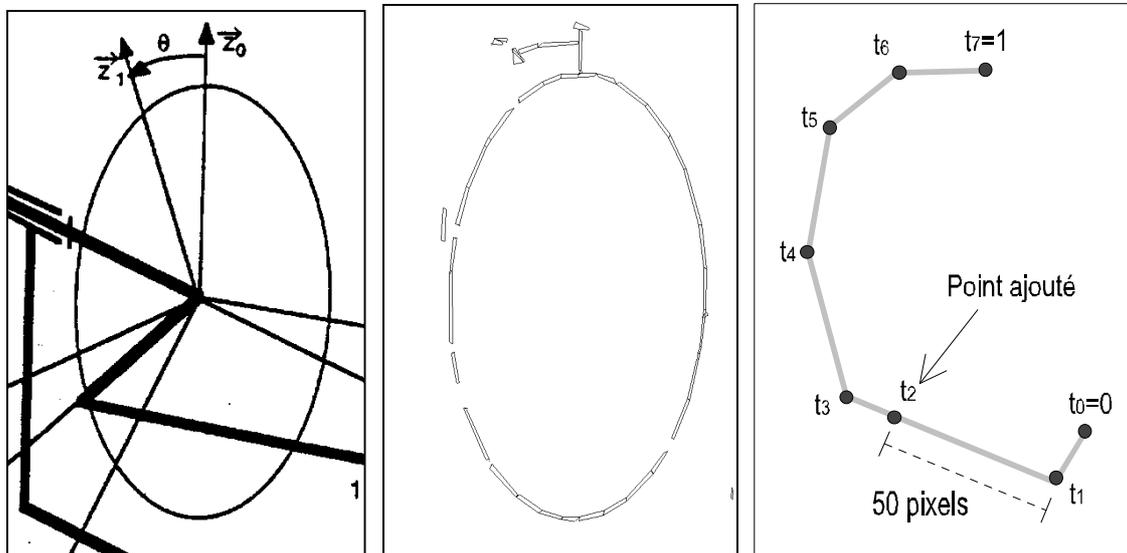


Figure 3.12 : Succession de quadrilatères susceptibles de correspondre à une courbe

Avant la tentative d'approximation, chaque courbe est représentée par une suite ordonnée de points. Il s'agit, soit d'une extrémité d'axe médian, soit d'un point élément de l'axe médian rajouté si la longueur du quadrilatère dépasse 50 pixels. Ainsi, 2 points consécutifs décrivant une courbe ne sont jamais éloignés de plus de 50 pixels.

Afin d'améliorer l'approximation, les courbes ayant des points d'inflexion sont divisées en plusieurs morceaux étudiés séparément. Ce découpage est réalisé par étude du signe du déterminant qu'il est possible d'associer à chaque point de référence formé par deux vecteurs consécutifs (sauf aux 2 extrémités de la courbe comme on peut le voir figure 3.13). Un changement de signe du déterminant dénote la présence d'un point d'inflexion.

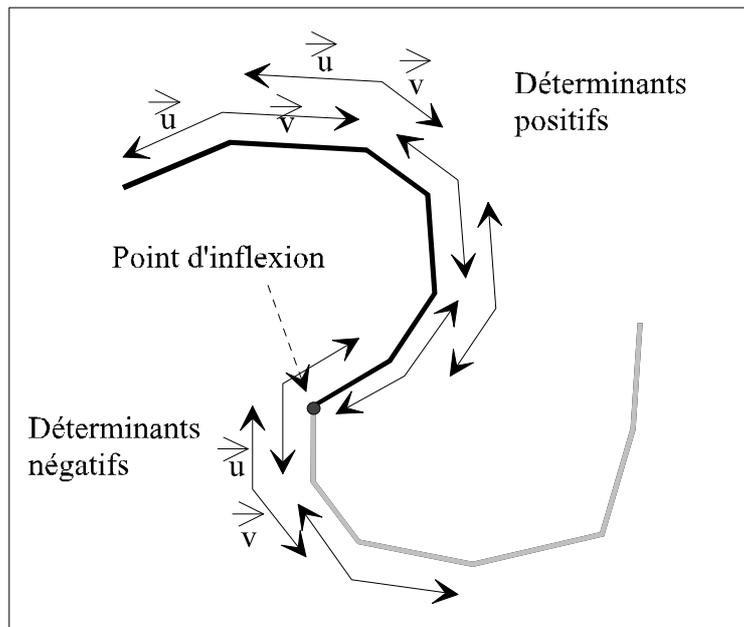


Figure 3.13 : Recherche des points d'inflexion

Ensuite, l'algorithme d'approximation suivant est appliqué :

Courbe = Suite de points à étudier

Degré = 1 (Degré du polynôme de Bézier)

Erreur = $+\infty$

TantQue Erreur diminue Faire

Degré = Degré + 1

Pôles = Approximation_Bézier(Degré, Courbe);

Courbe_Béz = Génère_Courbe_Bézier(Pôles)

Erreur = Calcul_Erreur_Max(Courbe, Courbe_Béz)

FinTantQue

L'erreur commise en considérant la courbe de Bézier trouvée est estimée en évaluant la distance entre les 2 courbes, relative à la norme de la convergence uniforme. Elle est évaluée par un calcul ponctuel à différents instants t_i , et seule l'erreur maximum est prise en compte. Les points de la courbe de Bézier pour chaque t_i sont calculés et sont projetés sur la courbe à approximer qui est une suite de segments. La distance séparant le point de sa projection définit l'erreur à $t = t_i$ (figure 3.14).

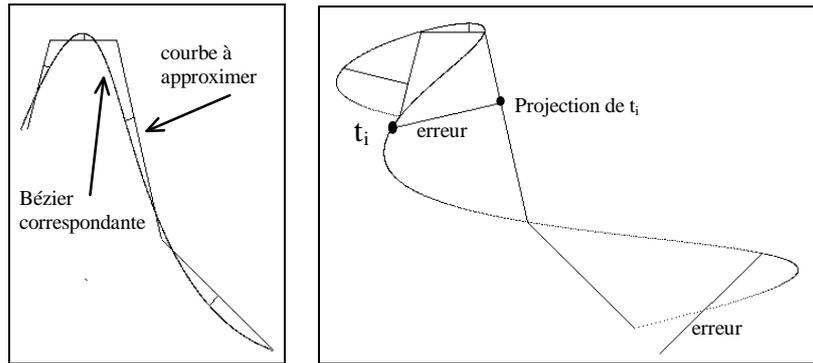


Figure 3.14 : Calcul de l'erreur d'approximation

Etant donnée l'évolution des courbes de Bézier en fonction du degré de leur polynôme caractéristique (figure 3.15), nous avons choisi, pour trouver la meilleure approximation, d'augmenter le degré du polynôme de Bézier tant que le résultat de l'approximation s'améliore, c'est à dire tant que l'erreur d'approximation diminue, ce qui est indiqué dans l'algorithme. Le degré et les pôles correspondant à la meilleure approximation obtenue constituent les paramètres finals de l'approximation. Dans la figure 3.15, la meilleure approximation est obtenue pour le degré 3.

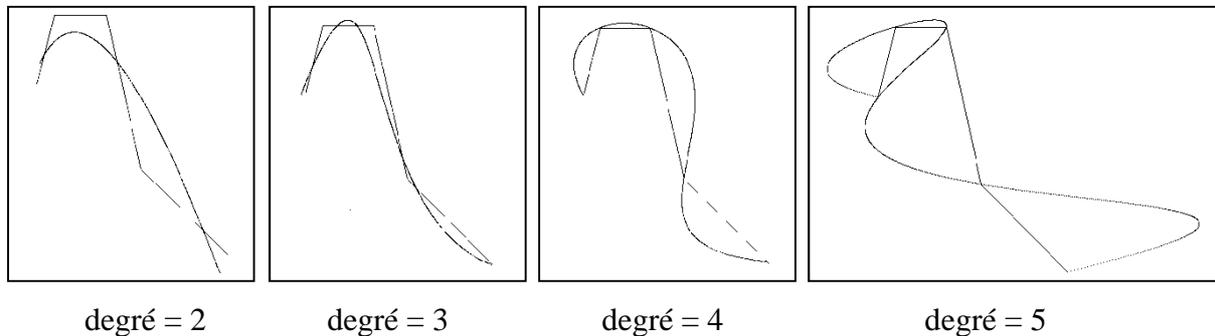


Figure 3.15 : Evolution des courbes de Bézier en fonction du degré du polynôme

Jugement de l'approximation

La méthode permet dans tous les cas de trouver un polynôme de Bézier qui doit alors être contrôlé. Il s'agit de décider si l'ensemble des points étudiés définit une courbe dont la courbe de Bézier approximante traduit bien l'allure. Le premier critère de jugement porte sur la comparaison de l'erreur finale d'approximation avec l'épaisseur moyenne du trait de la courbe.

Il est possible d'utiliser un deuxième critère plus mathématique qui permet de mesurer la proportion de variance qui peut être expliquée par la courbe de Bézier :

Soit $\sigma_m^2 = \frac{1}{m} \sum_{k=0}^m [(x_k - \bar{x}_i)^2 + (y_k - \bar{y}_i)^2]$ la variance de la courbe à approximer

Soit $\sigma_u^2 = \frac{1}{m} \sum_{k=0}^m \left[[(x_k - P_x(t_k)) - \overline{(x_i - P_x(t_i))}]^2 + [(y_k - P_y(t_k)) - \overline{(y_i - P_y(t_i))}]^2 \right]$ la variance résiduelle de l'erreur d'approximation en notant \bar{A} la moyenne des A_k .

Le paramètre $r = 1 - \frac{\sigma_u^2}{\sigma_m^2}$ permet alors de juger l'approximation. L'approximation est bonne

lorsque r est très proche de 1.

Lorsque l'approximation est jugée correcte les paramètres correspondants (degré et pôles du polynôme de Bézier) sont sauvegardés, tandis que les primitives de description ayant servi à la localisation de cette courbe sont supprimées.

Dans le cas contraire, on cherche à réduire la longueur de la courbe étudiée par la suppression du point extrémité qui forme l'angle le plus aigu avec celui qui le suit (figure 3.16).

L'algorithme est alors relancé en considérant la courbe nouvellement obtenue. Les ensembles étudiés doivent comporter au moins 4 points de contrôle, sinon l'hypothèse de présence d'une courbe est réfutée. On passe alors à une autre zone de focalisation.

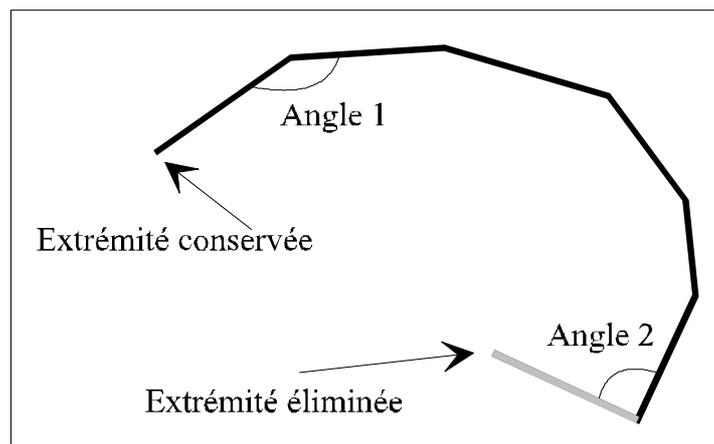


Figure 3.16 : Modification de la courbe étudiée

Des exemples de résultats fournis par ce spécialiste, après plusieurs exécutions successives, sont donnés figure 3.17.

3.2.4. Extraction des éléments mécaniques

Les éléments mécaniques constituant les schémas cinématiques sont composés essentiellement de segments et de symboles pleins dont on voit quelques exemples sur la figure 3.18. Leur assemblage réalise le mécanisme constitué de plusieurs pièces en mouvement les unes par rapport aux autres. Le dessin représente le mécanisme à un instant donné, dans un état donné.

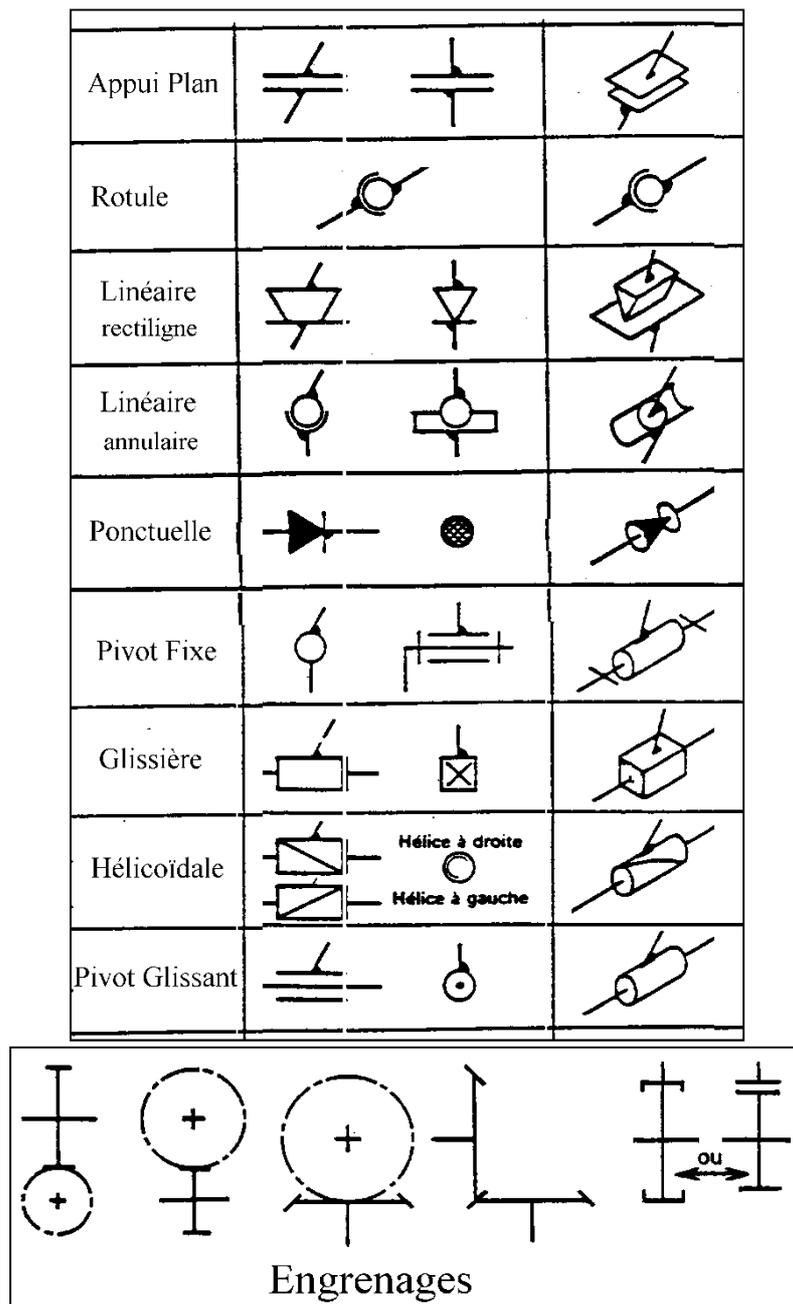


Figure 3.18 : Liaisons cinématiques normalisées

Pour localiser et reconstruire ces entités mécaniques, le spécialiste doit mettre en relation les différentes primitives du modèle de représentation, les regrouper pour former des objets de plus haut niveau.

Comme nous l'avons déjà indiqué, une représentation sous la forme d'un graphe nous est apparue la plus adaptée pour lier les différents objets qui constituent le dessin, et la plus apte à traduire les relations structurelles entre primitives pour obtenir un niveau global de description.

3.2.4.1. Construction du graphe structurel de l'image

Les primitives du modèle de représentation (Quadrilatères, Symboles Pleins) constituent les noeuds du graphe. Les arcs reliant ces noeuds traduisent les relations qui existent entre les primitives. Le graphe représente la structure globale du document.

Pour le construire, on définit, pour chaque primitive, une zone d'influence à laquelle on limite l'étude. Des arcs décriront la nature des liens existant entre la primitive en cours d'étude et celles appartenant à sa zone d'influence (figure 3.19). Nous construirons, dans un premier temps, le graphe correspondant uniquement aux formes fines (suite SQ). Le tableau 3.1 répertorie les différents types de relations qui servent à la construction du graphe. Pour cette primitive, la taille de la zone d'influence est fonction de l'épaisseur maximum des traits détectés dans l'image (figure 3.19).

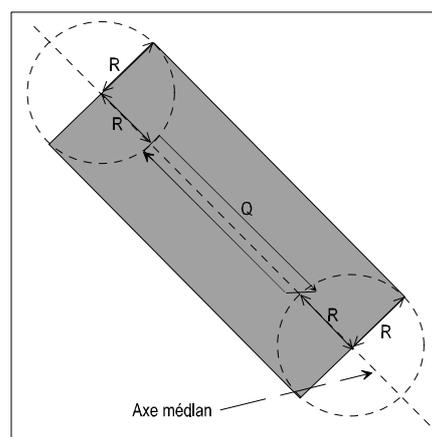


Figure 3.19 : Construction de la zone d'influence de la primitive quadrilatère

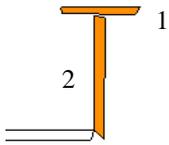
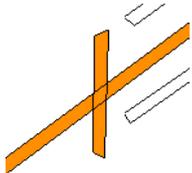
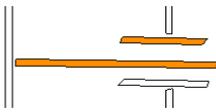
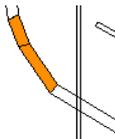
Type de liaison entre deux primitives	Exemple correspondant	Commentaires
T (Liaison en T)		Cette liaison est orientée T(1,2) est différent de T(2,1). Ici on a T(1,2)
X (Intersection)		
P (Parallèles)		Il faut que la projection orthogonale de l'une des primitives intersecte l'autre
L (Liaison en L)		
S (Successifs)		

Tableau 3.1 : Types d'interactions entre primitives

La construction du graphe de l'image est incrémentale. Chaque primitive est passée en revue, la zone d'influence qui lui correspond est construite, elle permet de déterminer ses primitives voisines. Les types d'interactions existant entre les primitives, éléments d'une même zone d'influence, sont étudiés précisément et sauvegardés. Il n'est pas nécessaire qu'il y ait intersection entre les primitives pour qu'un arc les mette en correspondance. Lors de l'étude du type d'interaction, le spécialiste retourne consulter l'image initiale (figure 3.20) pour s'assurer des contacts réels entre traits.

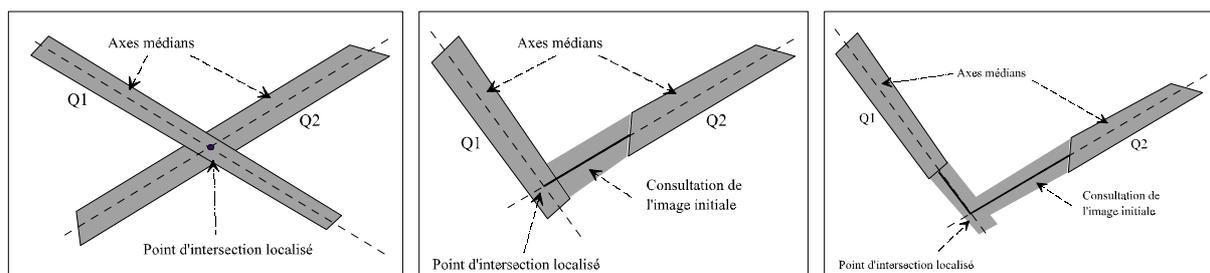


Figure 3.20 : Détermination du type d'interactions entre primitives voisines

Cette étude des interactions entre primitives permet de connaître les points d'intersection entre primitives (ou plutôt entre axes médians). Il est bon de remarquer qu'ainsi il est possible d'achever une étape de vectorisation en résolvant le problème soulevé dans le paragraphe 2.3.1.3 concernant la conservation de la connexité des traits. La figure 3.21 montre d'ailleurs les améliorations obtenues par application de cette méthode sur l'exemple correspondant à la figure 2.32.

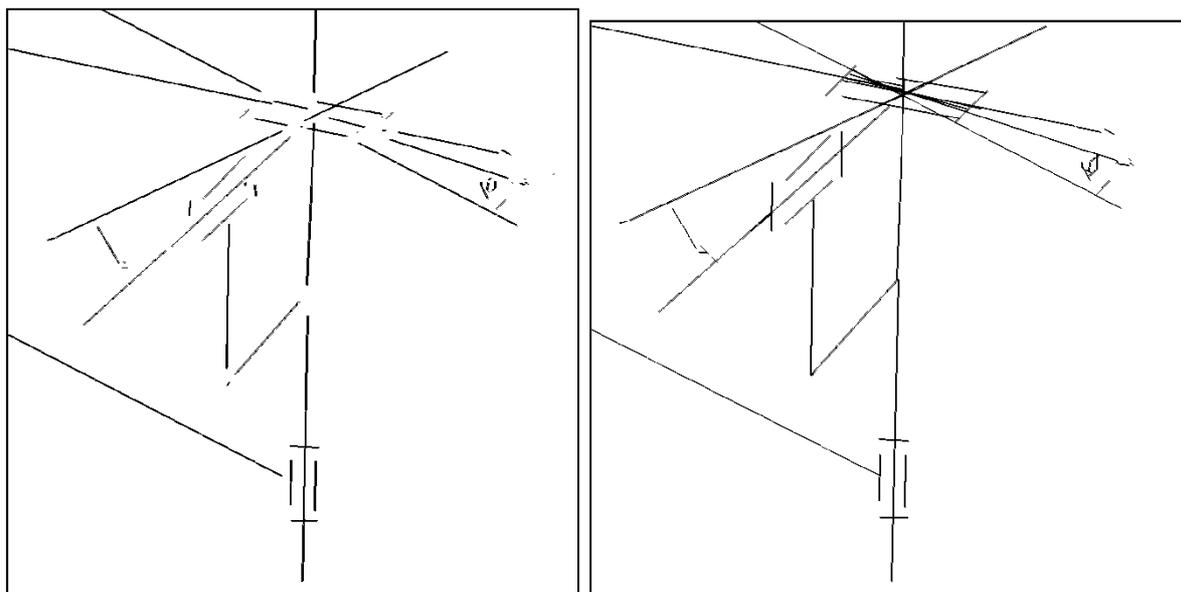
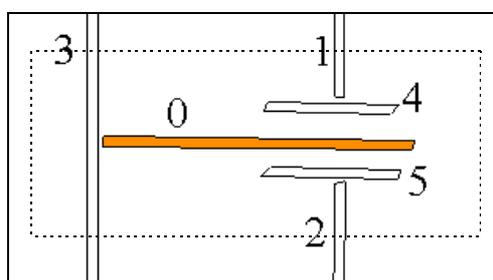


Figure 3.21 : Etude des interactions pour améliorer la vectorisation

Chaque interaction entre primitives se traduit par la création d'un arc mettant en relation 2 primitives et indiquant la nature de l'interaction. La figure 3.22 décrit la partie P_G du graphe global de l'image correspondant au Quadrilatère étiqueté 0 (élément de SQ).



(a) Zone d'influence du Quadrilatère 0

Nb Noeuds=6 - Nb Arcs=6

$L=T(3,0), P(0,4), P(0,5), T(4,1), T(5,2), P(4,5).$

(b) Sous-graphe correspondant à (a)

Figure 3.22 : Construction de la partie P_G du graphe

Quand toutes les primitives du modèle de description sont étudiées, le graphe structurel construit suffit à traduire toutes les relations existant entre les primitives entrant dans la composition du document.

Ce type de description est le point de départ de la phase de reconnaissance que nous allons aborder maintenant.

3.2.4.2. Reconnaissance des éléments mécaniques

Pour reconnaître une entité, il est nécessaire de la comparer à des références. Dans notre cas, la reconnaissance des liaisons consiste à trouver dans un graphe tous les sous graphes qui correspondent à des modèles d'entités mécaniques. Ce problème est connu sous le nom de recherche d'*isomorphisme de sous graphe* ou d'*inclusion de graphes* [Miclet84, Bélaïd92b]. L'intérêt est de pouvoir comparer deux représentations pour extraire les sous graphes caractéristiques de formes élémentaires (éléments mécaniques appartenant à une base d'apprentissage) du grand graphe modélisant une forme complexe (le document).

Notre algorithme, permettant de déterminer si 2 graphes sont isomorphes, est combinatoire ; il cherche toutes les correspondances licites, et teste pour chacune d'elles s'il y a isomorphisme. Par licite on entend que les **noeuds associés** aient même degré et qu'il y ait correspondance entre les attributs des arcs mis en correspondance. Pour chaque graphe élément de la base des modèles, l'algorithme construit progressivement le sous-graphe candidat en appariant un noeud du modèle avec un noeud du graphe de l'image tant que la correspondance reste licite. Lorsque tous les noeuds du modèle ont un correspondant dans le graphe de l'image, l'isomorphisme est trouvé ce qui traduit la présence de la forme recherchée (correspondant au modèle) dans l'image (figure 3.23).

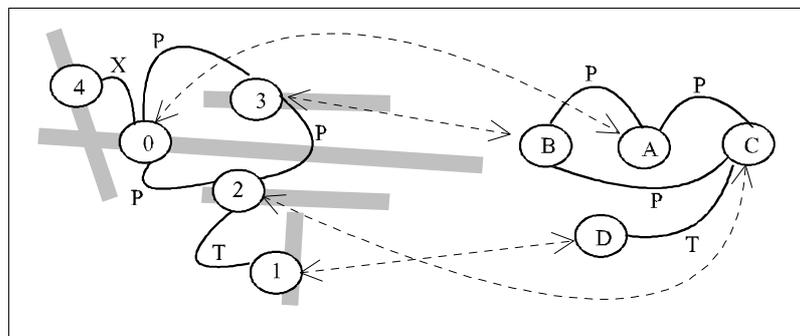
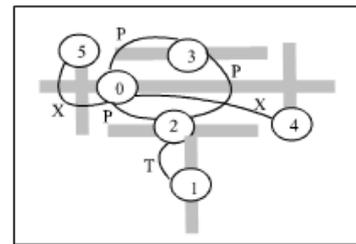


Figure 3.23 : Isomorphisme entre graphes

Les graphes décrivant chaque liaison cinématique normalisée (les modèles) ont été construits manuellement et stockés dans un dictionnaire (base d'apprentissage). Les modèles doivent être classés suivant un ordre précis puisque le graphe correspondant à certains modèles peut correspondre à une partie du graphe correspondant à une entité plus complexe (par exemple pivot glissant \subset pivot fixe). Il faut donc commencer par rechercher les entités les plus complexes et terminer par les entités simples. Nous avons choisi de classer les modèles en fonction du nombre de noeuds constituant le graphe les décrivant.

Il est possible d'inclure un mécanisme d'apprentissage automatique dans le système, dans ce cas, l'utilisateur indique le nom à associer à la configuration spéciale des primitives du modèle qui lui est présenté. Sur la figure 3.4, on a indiqué quelques graphes de liaisons cinématiques usuelles.

Nb Noeuds=6 - Nb Arcs=7
 $L=P(0,2),P(0,3),X(0,4),X(0,5),P(2,3),P(2,5),T(2,1)$.
 Nom=Liaison Pivot Fixe.



Nb Noeuds=6 - Nb Arcs=5
 $L=T(2,0),T(5,0),T(0,1),P(2,3),T(3,4)$.
 Nom=Engrenage (type 1).

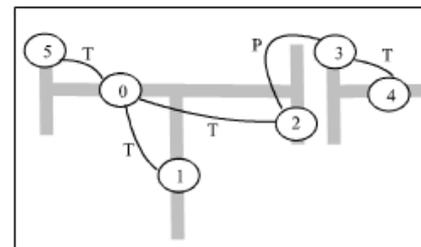


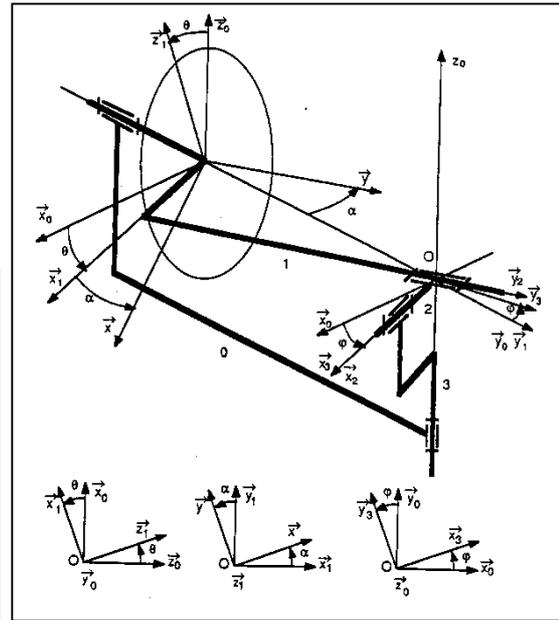
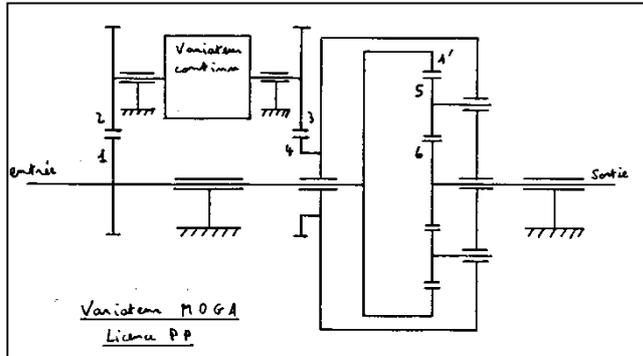
Figure 3.24 : Extrait de la base des modèles : Graphes de 2 modèles d'entités mécaniques

On peut remarquer que durant la construction même du graphe, il est possible de localiser certains éléments mécaniques (certaines liaisons cinématiques normalisées). Les éléments du dictionnaire sont recherchés dans chacun des graphes obtenus pour les différentes zones d'influence construites.

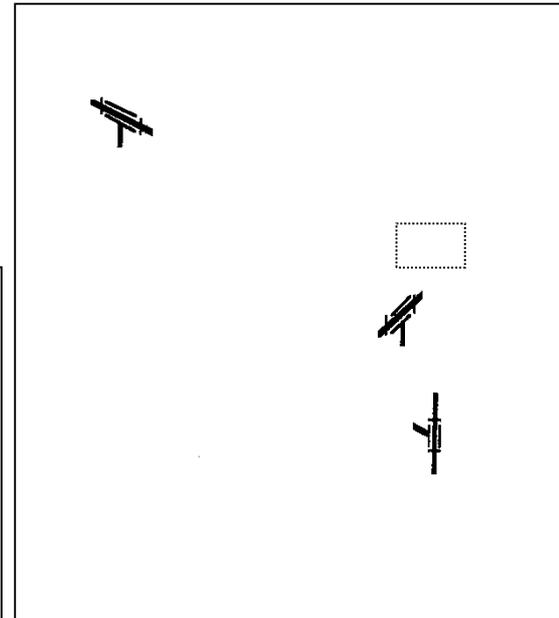
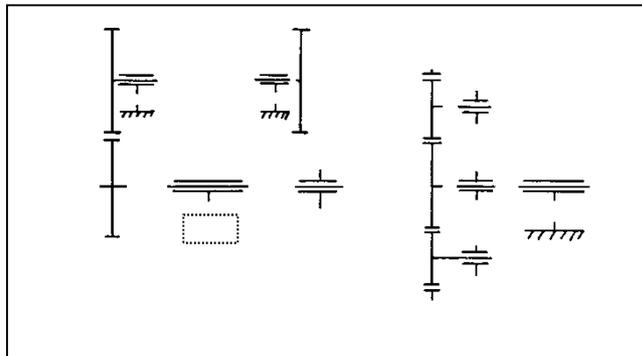
Reprenons l'exemple correspondant à la partie P_G , l'existence d'un sous-graphe correspondant au graphe modèle d'une *liaison pivot glissant* est constatée dans le graphe correspondant à P_G (figure 3.22b). La présence de cette entité est détectée, sa **position** ainsi que les **orientations** des 2 pièces mises en relation sont sauvegardées. Les primitives entrant

dans la composition de la liaison sont étiquetées de manière à prévenir les autres spécialistes ; ce mécanisme d'étiquetage sera décrit plus en détail dans la suite de ce mémoire.

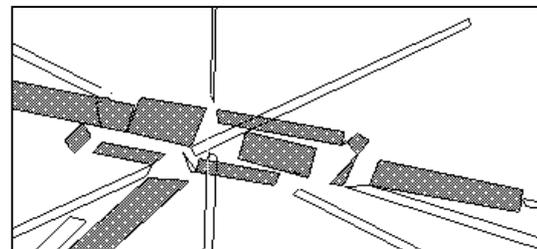
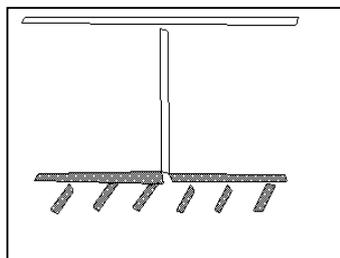
Le graphe global est ainsi construit et les entités mécaniques non ambiguës sont reconnues et étiquetées durant cette étude. La figure 3.25 fournit des exemples de résultats. Dans chacun des 2 cas décrits, cependant une entité mécanique, pourtant élément de la base des modèles, a été oubliée. Les grossissements des images des quadrilatères correspondants expliquent ces oublis dus au manque de précision dans le dessin initial.



(a)



(b)



(c)

(a) Images initiales, (b) Entités reconnues, (c) Entités non reconnues

Figure 3.25 : Résultats de l'extraction des entités mécaniques

Validation de l'hypothèse par vérification de contraintes

Afin de ne pas alourdir le graphe de l'image, les différents types de relations (attributs des arcs du graphe) liant les primitives de description sont peu nombreux. Ils servent tous à décrire la disposition relative des primitives mais pas leur taille. Les entités normalisées que l'on désire localiser sont, elles, constituées d'éléments pour lesquels les dimensions importent (longueurs comparables, ...). Il est donc nécessaire, lorsqu'un isomorphisme a été trouvé, d'effectuer une étude plus précise des primitives concernées.

En fait, chaque isomorphisme trouvé provoque la génération d'une hypothèse (présence de l'entité mécanique correspondant au graphe détecté). Des vérifications supplémentaires sont alors réalisées pour valider ou rejeter cette hypothèse.

Ces contraintes dépendent du type de la liaison supposée être présente. Ainsi à chaque modèle, il est également possible d'associer une liste de contraintes devant être vérifiées par le spécialiste avant qu'il ne puisse vraiment affirmer la présence du modèle.

Ces contraintes sont stockées dans une base faisant correspondre une liste de contraintes à chaque graphe modèle. Les contraintes utilisées jusqu'à présent sont résumées dans le tableau suivant.

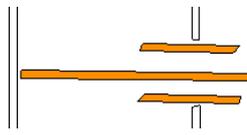
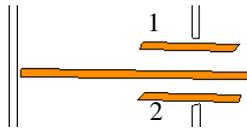
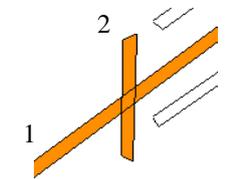
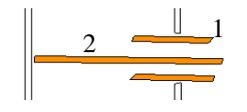
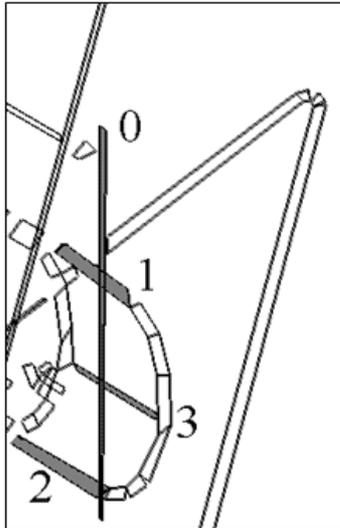
Type de contraintes	Exemple correspondant	Commentaires
E (Entre)		La contrainte E(0,1,2) implique que la primitive 0 soit entre les primitives 1 et 2.
= (Longueur comparable)		=(1,2,ratio) implique que le quotient q de la longueur de la primitive la plus longue par celle de l'autre primitive soit inférieur à $ratio$ ($ratio \geq 1$).
> (Plus long)		>(1,2,facteur) implique que la longueur de la primitive 1 soit au moins <i>facteur</i> fois plus grande que celle de la primitive 2.
D (Distinct)		D(1,2) implique qu'il n'existe pas de chemin entre le noeud 1 (primitive 1) et le noeud 2 (primitive 2) dans le graphe correspondant à la zone d'influence.

Tableau 3.2 : Les différentes contraintes

La figure 3.26 illustre la nécessité d'utiliser des contraintes en montrant qu'il est possible d'obtenir un isomorphisme entre le graphe correspondant à un modèle et une partie du graphe de l'image ne comportant pourtant aucune liaison mécanique normalisée. Seules les contraintes permettent au spécialiste de rejeter l'hypothèse.



Grphe de la partie en grisé de l'image :

$L = X(0,1) , T(0,2) , X(0,3) , P(1,2) , P(1,3) , P(2,3) , \dots$

Grphe du modèle Engrenage (type 2) et contraintes associées :

$L = X(A,B) , T(A,C) , X(A,D) , P(B,C) , P(C,D) , P(B,D)$

$C = E(C,B,D) , >(A,B,2) , >(A,D,2) , \downarrow=(B,D,1.6) , >(C,D,2) , >(C,B,2)$

Isomorphisme trouvé :

$A = 0, B = 1, C = 2, D = 3$

Contraintes non respectées :

$E(2,1,3) , >(3,1,2) , >(2,1,2)$

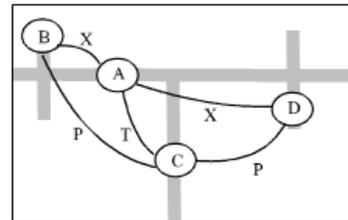


Figure 3.26 : Fausse détection d'entité normalisée

Etiquetage des pièces du mécanisme

Une liaison cinématique met en relation 2 pièces mécaniques différentes et décrit le type de mouvement de l'une par rapport à l'autre. Lorsqu'une liaison est reconnue, les primitives de description entrant dans la composition sont d'abord étiquetées en tant que constituants de la liaison. Mais, il est aussi possible de leur attribuer un autre type d'étiquette, des étiquettes qui permettront de faire la distinction entre les 2 pièces mécaniques mises en rapport par la liaison détectée (figure 3.27). Les primitives entrant dans la composition d'une pièce 1 (au niveau de la liaison) se voient attribuer un même numéro. Le numéro attribué est le numéro d'étiquette courante + 1 si aucune de ces primitives ne possède déjà d'étiquette. Dans le cas contraire, le numéro déjà attribué à une primitive est propagé sur les autres constituants de la pièce. Le même mécanisme est exécuté sur les primitives constituant la deuxième pièce.

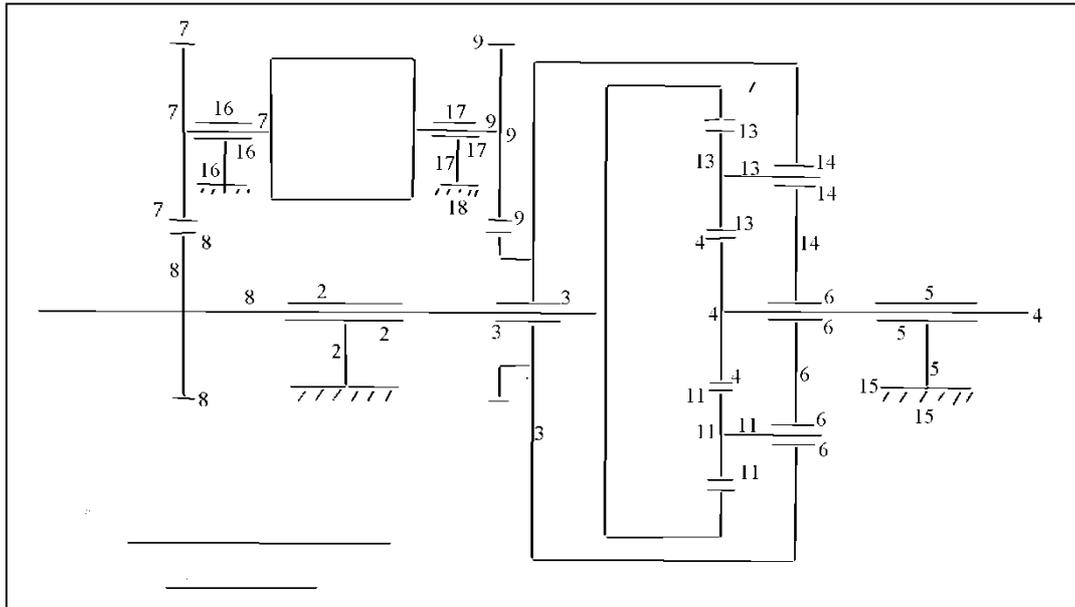


Figure 3.27 : Etiquetage des pièces mécaniques

Cette propagation d'étiquettes permet de dissocier les différentes formes entrant dans la composition du mécanisme décrit par le schéma. Il doit aussi permettre de lever certaines ambiguïtés et de vérifier la cohérence des données extraites. Des corrections peuvent être réalisées et de nouvelles reconnaissances peuvent avoir lieu puisque le modèle de description du document est maintenu à jour durant toutes ces phases.

Quelques remarques

Tout d'abord, comme le montre la figure 3.25, il faut signaler que les résultats obtenus par la méthode choisie (recherche d'isomorphismes), dépendent beaucoup de la représentation de l'image obtenue à l'aide des primitives de description. La correspondance entre le graphe de l'image et le graphe modèle doit être parfaite. Aucun calcul de distance entre graphes n'a été implémenté pour l'instant. Les entités se trouvant dans des zones complexes (superposition de traits) sont donc mal détectées. Néanmoins, il semble possible de mettre en place des règles de transformation du graphe image permettant de modifier une configuration particulière en une configuration plus simple (figure 3.28) pour résoudre ce problème. Une autre solution, moins performante, consiste à multiplier les modèles pour une même entité dans la base d'apprentissage.

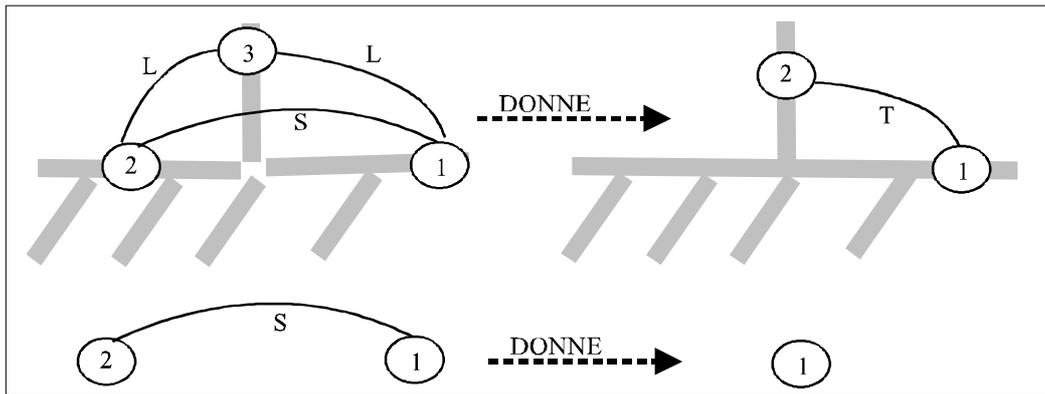


Figure 3.28 : Exemple de mutation

L'utilisation de zones d'influence limite la détection à des formes de taille comparable à celles des différentes zones générées. Il n'est, par exemple, pas possible de détecter un carré constitué de 4 quadrilatères pendant la phase de construction du graphe (figure 3.29). Pour le localiser, il est nécessaire d'attendre que le graphe de l'image soit totalement construit pour rechercher ensuite de nouvelles dispositions particulières des primitives.

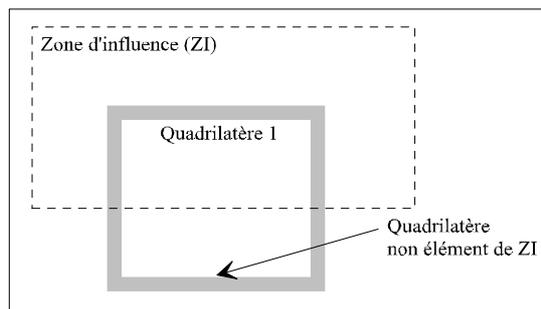


Figure 3.29 : Limite de l'utilisation des zones d'influence

L'étiquetage des pièces constituant le mécanisme ne peut pas être réalisé entièrement par le spécialiste chargé de localiser les entités normalisées. Il ne possède, en effet, pas les connaissances suffisantes pour traiter les quadrilatères n'entrant pas dans la composition des éléments mécaniques (cotations, axes, traits encore non identifiés, ...).

3.2.5. Reconstruction 3D du mécanisme

Ce spécialiste poursuit le travail d'interprétation mécanique commencé lors de l'extraction des symboles normalisés. Son but est d'analyser les relations existant entre les différentes liaisons dans le but de leur attribuer des coordonnées dans l'espace 3D et de poursuivre le repérage des différentes pièces constituant le mécanisme décrit par le schéma cinématique.

Détermination des directions des axes du repère

Il est possible que le dessin ne comporte que 2 directions différentes (0° et 90° pour un dessin en 2D) ; dans ce cas aucun problème de reconstruction 3D ne se pose, les positions des liaisons dans l'espace (en 3D) correspondent aux coordonnées de ces symboles dans l'image du document. Celles-ci sont stockées dès la localisation et la reconnaissance des entités. Il ne reste, dans ce cas, qu'à réaliser la phase de reconstruction des différentes pièces du mécanisme.

La détermination de la direction des trois axes formant le repère par rapport auquel seront, par la suite, positionnées les différentes entités reconnues est un préalable nécessaire. Pour cela, le spécialiste consulte l'ensemble des liaisons cinématiques déjà reconnues (par le spécialiste décrit précédemment et normalement chargé de cette tâche) et étudie les 2 orientations associées à chacun des objets (figure 3.30). Les 3 directions les plus fréquemment rencontrées (d_1, d_2, d_3) constitueront les directions des axes du repère. Lors de cette étude, le nombre de valeurs différentes que peuvent prendre les orientations des quadrilatères est ramené à 8. Celles-ci sont précisées sur la figure 3.31.

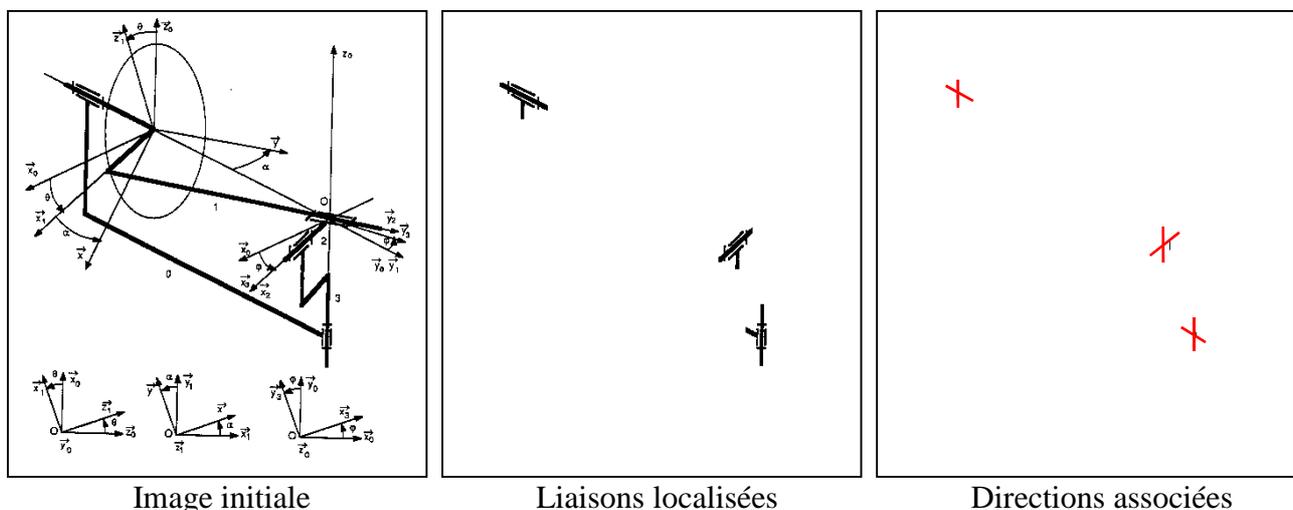


Figure 3.30 : Directions associées aux liaisons

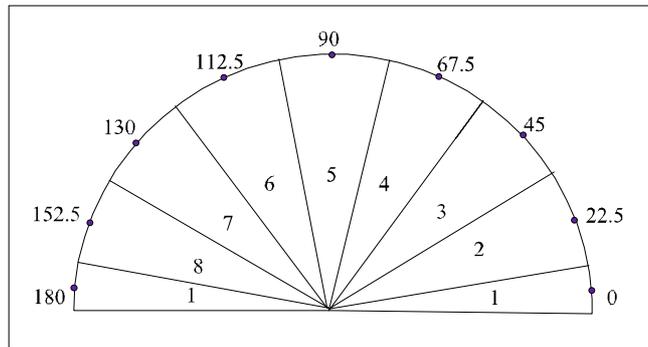


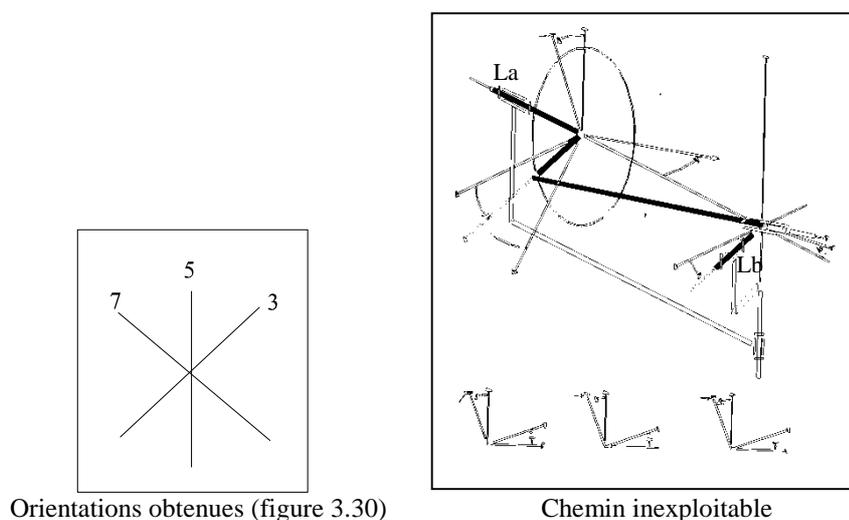
Figure 3.31 : Les différentes orientations possibles

Reconstruction des pièces

Dans les schémas cinématiques, les différentes liaisons sont reliées entre elles par des pièces dont l'ensemble constitue un mécanisme. Chaque pièce peut comporter une ou plusieurs liaisons qui déterminent son mouvement par rapport aux autres pièces.

Notre système utilise ce fait pour, d'une part, différencier les traits, éléments du mécanisme, des autres traits, par exemple ceux faisant partie des annotations. D'autre part, ceci permet de positionner chaque liaison par rapport à une ou plusieurs autres concernant la même pièce.

Par ailleurs, il n'est pas toujours possible d'obtenir les coordonnées d'une liaison en fonction des coordonnées d'une autre, notamment si la pièce du mécanisme qui les relie a une orientation (quadrilatères) qui ne correspond pas aux orientations des axes précédemment définis. On constate que dans la figure 3.32, il existe un chemin de La à Lb, néanmoins les traits ne respectent pas les orientations du repère ce qui ne permet pas de positionner une liaison par rapport à l'autre dans l'espace à 3 dimensions. C'est pourquoi le concepteur a jugé indispensable de faire figurer, en bas du dessin, des projections des différents axes de manière à rendre possible la compréhension de son dessin par le lecteur.



Orientations obtenues (figure 3.30)

Chemin inexploitable

Figure 3.32 : Chemin inexploitable de la figure 3.30

Pour interpréter la projection sur un plan d'une scène en 3D, il est nécessaire de recourir à certaines hypothèses. Voici celles utilisées par ce spécialiste :

- si les projections 2D de 2 droites sont parallèles, alors ces droites sont considérées comme effectivement parallèles dans la scène

- 2 segments, ayant une extrémité commune sur l'image (espace 2D), sont considérés sécants dans l'espace 3D

- 2 droites parallèles, «perpendiculaires» à une même droite sur l'image au sens du repère 3D déjà détecté, appartiendront toutes à un même plan de l'espace.

Le spécialiste utilise ces postulats pour, partant de l'élément d'une liaison, que l'on sait appartenir à une pièce P_1 (figure 3.33), essayer de reconstruire cette pièce jusqu'à rencontrer un élément d'une autre liaison mécanique (fusion de 2 pièces pour en constituer une seule plus complète). La reconstruction des pièces est ainsi réalisée, par parcours du modèle de représentation, par recherche du plus court chemin entre 2 liaisons mécaniques L_i et L_j (figure 3.33). Durant cette recherche, seulement certains arcs peuvent être utilisés (suivant la valeur de l'attribut décrivant le type d'interaction qui existe entre les 2 primitives) afin de respecter les hypothèses : les liaisons de type 'P' et 'X' ne peuvent pas être utilisées. De même, les orientations des quadrilatères utilisés pour passer d'une liaison à l'autre doivent toutes correspondre à l'une des directions du repère préalablement choisi.

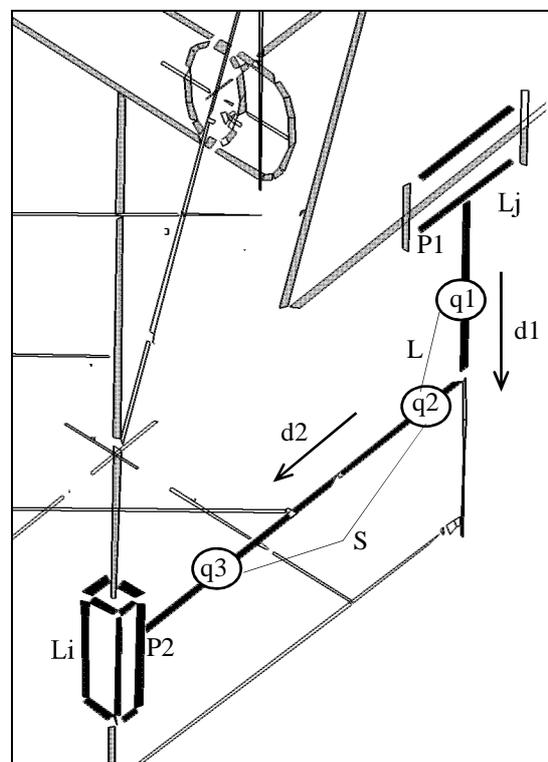


Figure 3.33 : Reconstruction d'une pièce

Cette recherche du plus court chemin dans le graphe (minimisation du nombre d'arcs utilisés par étiquetage des noeuds - figure 3.34) est lancée entre chacun des quadrilatères utilisés pour fixer les 2 orientations associées aux liaisons préalablement détectées; et ceci pour chacune des liaisons. Dans certains cas, aucun chemin n'est trouvé par l'algorithme, il faut alors espérer qu'un chemin existera à partir d'une autre liaison si l'on veut réussir à positionner chacune des liaisons du dessin.

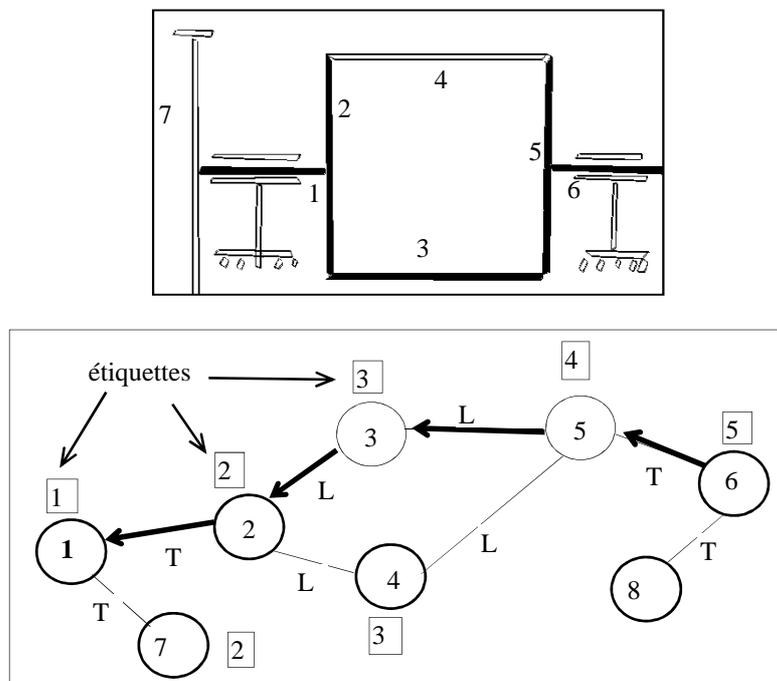
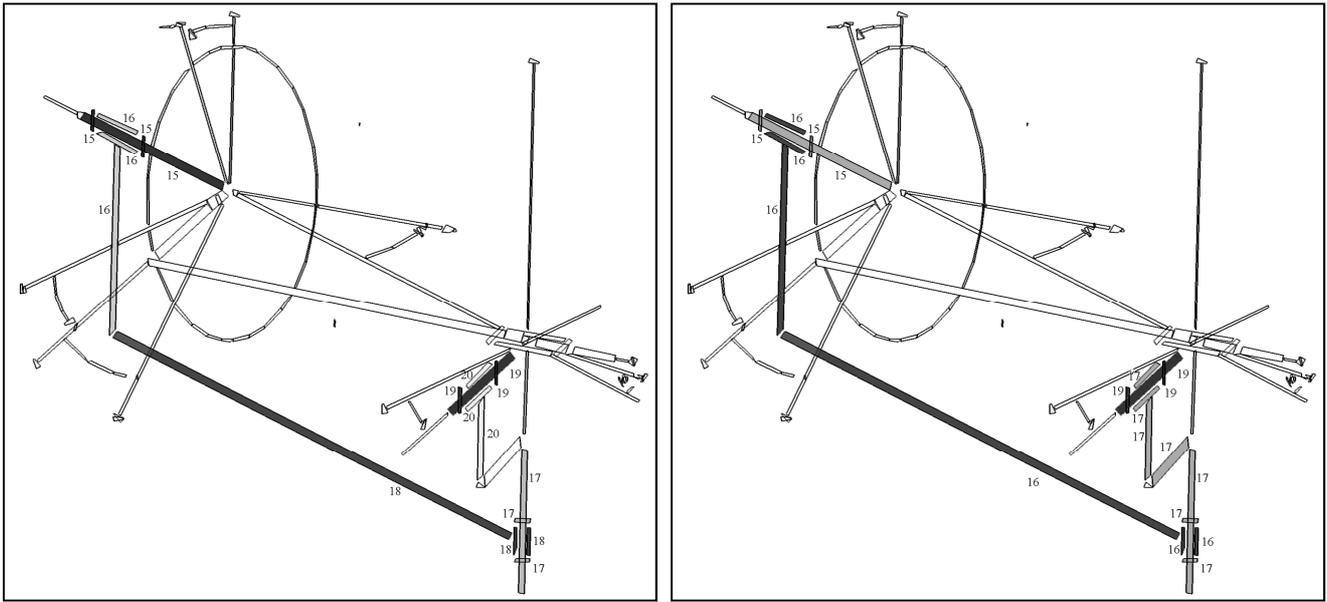


Figure 3.34 : Algorithme de recherche du plus court chemin

Lorsqu'un chemin est trouvé, il peut ne pas être accepté si le nombre de primitives nécessaires pour passer d'une liaison à l'autre est supérieur à 10.

Si au contraire un chemin respectant toutes les contraintes citées ci-dessus a été reconstitué, les étiquettes correspondant aux pièces sont propagées sur toute la longueur du chemin (sur chacune des primitives concernées) pour traduire leur appartenance à une même pièce. On voit sur la figure 3.35 la transformation opérée.



Avant la reconstruction

Après la reconstruction

Figure 3.35 : Propagation des étiquettes

Localisation 3D

La prise en compte des distances séparant les points d'intersection des quadrilatères successivement utilisés durant le parcours (figure 3.36), mises en rapport avec chacune des orientations correspondantes, permet d'attribuer des positions relatives aux différentes liaisons selon des équations du type :

$$\begin{aligned} \vec{OL}_2 &= \vec{OL}_1 + a \cdot \vec{d}_{a1} + b \cdot \vec{d}_{a2} + c \cdot \vec{d}_{a3} && \text{avec } (a,b,c) \in \mathfrak{R}^3 \\ \vec{OL}_3 &= \vec{OL}_2 + a \cdot \vec{d}_{b1} + b \cdot \vec{d}_{b2} + c \cdot \vec{d}_{b3} && \text{où } L_i \text{ désigne les coordonnées de la liaison } i \\ &\dots && \end{aligned}$$

La figure 3.36 donne un exemple de positionnement des liaisons obtenu grâce à cette reconstruction des pièces du mécanisme.

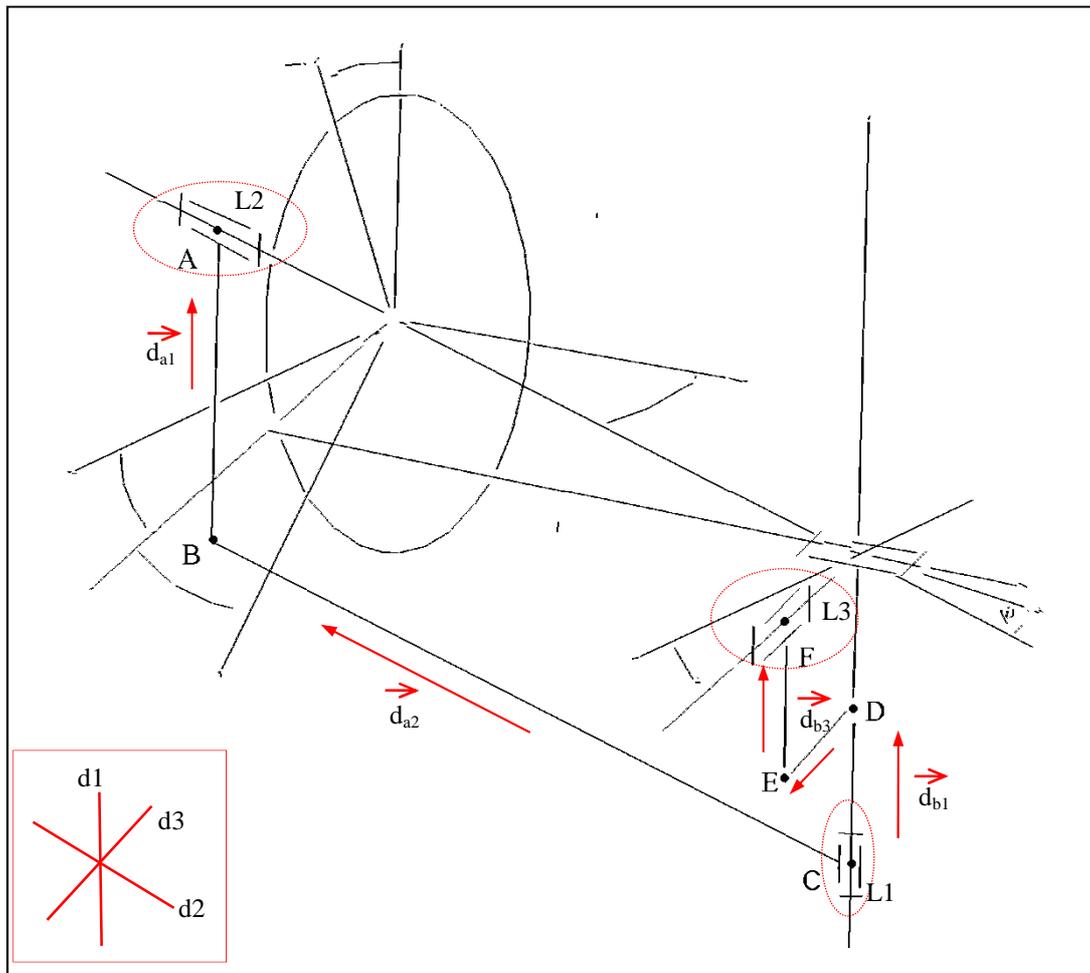


Figure 3.36 : Positionnement des liaisons

Le sens associé à chacun des axes est déterminé lors de chaque parcours du plus court chemin trouvé et correspond à celui du premier couple de points (vecteur) de même orientation rencontré. Il est, de cette manière, possible d'associer une valeur algébrique à chaque couple de points (quadrilatère). Ce qui donne toujours sur la figure 3.36 :

- Lors de l'étude du chemin menant de L1 à L2, le vecteur \overrightarrow{CB} détermine le sens à associer à \vec{d}_{a2} et le vecteur \overrightarrow{BA} celui à associer à la direction \vec{d}_{a1} . On obtient alors :

$$\overrightarrow{L_1L_2} = \overline{CB} \cdot \vec{d}_{a2} + \overline{BA} \cdot \vec{d}_{a1}$$

- Lors de l'étude du chemin menant de L1 à L3, le vecteur \overrightarrow{CD} détermine le sens de \vec{d}_{b1} et le vecteur \overrightarrow{DE} celui de \vec{d}_{b3} . On obtient alors :

$$\overrightarrow{L_1L_3} = (\overline{CD} + \overline{EF}) \cdot \vec{d}_{b1} + \overline{DE} \cdot \vec{d}_{b3}$$

Il est possible que certaines liaisons soient positionnées par rapport à plusieurs autres, on peut alors, dans le cas d'obtention d'une boucle fermée, vérifier la cohérence de l'interprétation (selon le même procédé que celui utilisé pour les chaînes de cotations dans les projections orthogonales [Dori92]).

La figure 3.37 montre les résultats de l'étape de reconstruction 3D et de l'étiquetage des pièces obtenus sur l'image correspondant à la figure 1 tandis que la figure 3.38 décrit l'apport fourni par ce spécialiste sur une image en 2 dimensions.

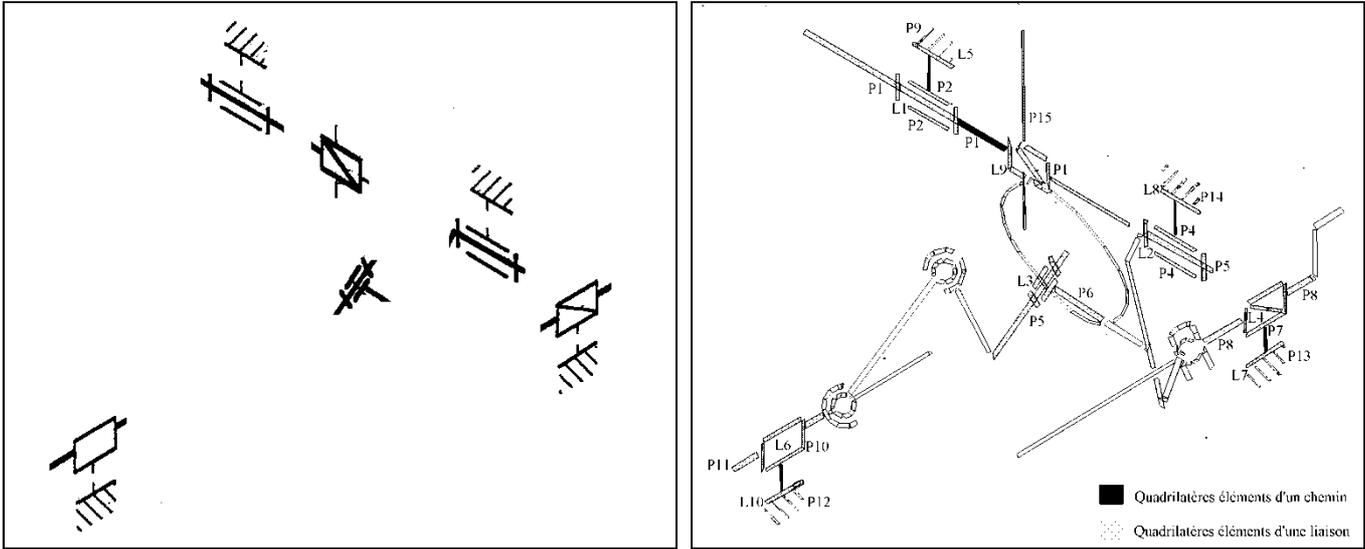


Figure 3.37 : Résultats obtenus sur une image 3D

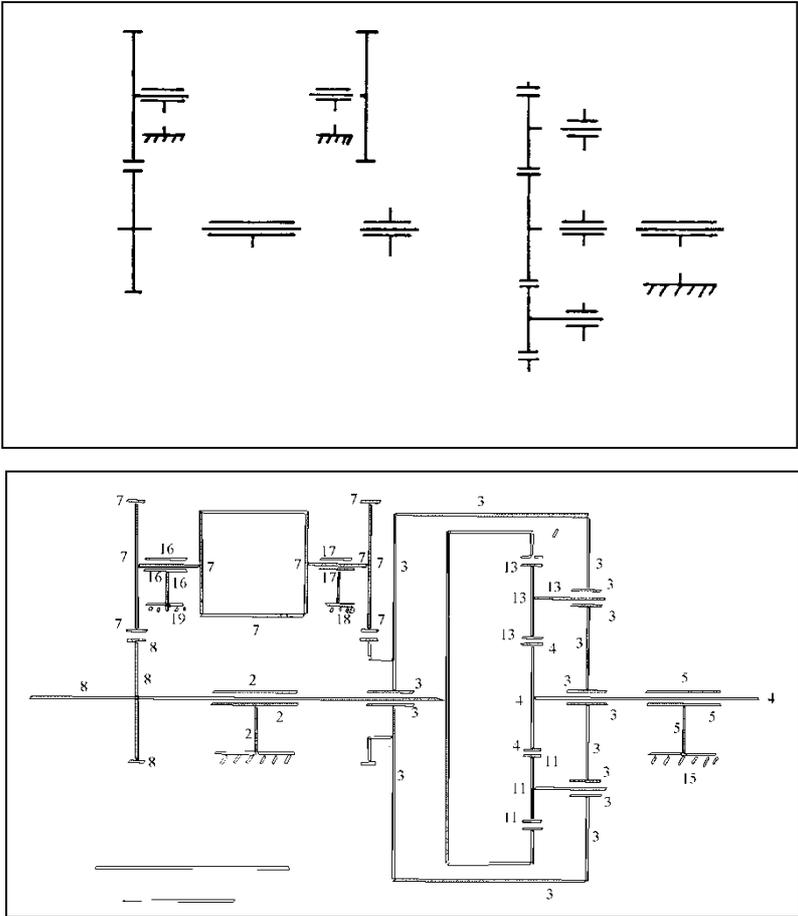


Figure 3.38 : Résultats obtenus sur une image 2D

3.3. Coopération et cycles perceptifs

Cycles perceptifs

Les **spécialistes** possèdent des connaissances élémentaires, précises et suffisantes dans des domaines particuliers. Tous utilisent le modèle de représentation (la vision globale) pour focaliser leur attention en différents endroits de l'image et vérifier les hypothèses qu'ils ont émises lors de l'exploration.

La mise en place informatique du mécanisme hypothèses-validation implique une certaine séquentialisation. Nous avons défini un mécanisme de **cycles perceptifs** qui permet de faire coopérer les différents processus. Les spécialistes sont ainsi activés de manière séquentielle jusqu'à ce que la **compréhension totale** du document soit traduite par le modèle de représentation. Chaque activation d'un spécialiste provoque l'évolution du modèle de représentation (figure 3.39).

La vision globale est à la base des hypothèses générées, tandis que les analyses locales permettent la validation ou la réfutation de ces hypothèses.

Ogier parle lui aussi de cycles perceptifs [Ogier95], cependant la méthodologie qu'il associe à ce concept diffère de la nôtre par plusieurs points. Tout d'abord, un cycle perceptif désigne, dans son système, l'enchaînement des traitements permettant d'obtenir une entité cadastrale cohérente : « Dans notre cas, parcourir le cycle revient à effectuer des communications entre différents points de vue issus d'images réduites rappelant la vision périphérique (approche globale) et de l'image haute définition rappelant la vision fovéale (approche analytique) ». Il s'agit donc de cycles de construction d'objets cohérents qui peuvent être effectués par chacun des spécialistes (dans leur coin) tandis que, dans notre cas, un cycle perceptif correspond à l'activation séquentielle de l'ensemble des spécialistes permettant la coopération.

Ogier parle également de vision globale et locale; dans son système, cela consiste à « émettre des hypothèses sur le contenu du document, à partir d'images basse résolution simulée par des réductions ». Ainsi, il propose une argumentation différente de la nôtre selon laquelle les hypothèses se basent sur la vision globale et les analyses locales, et sont validées par confrontation avec le contexte. Les hypothèses sont validées par analyse de la cohérence alors que notre système est plutôt basé sur l'extraction des **informations évidentes** à un instant donné et sur la minimisation de la prise de risque durant les cycles perceptifs.

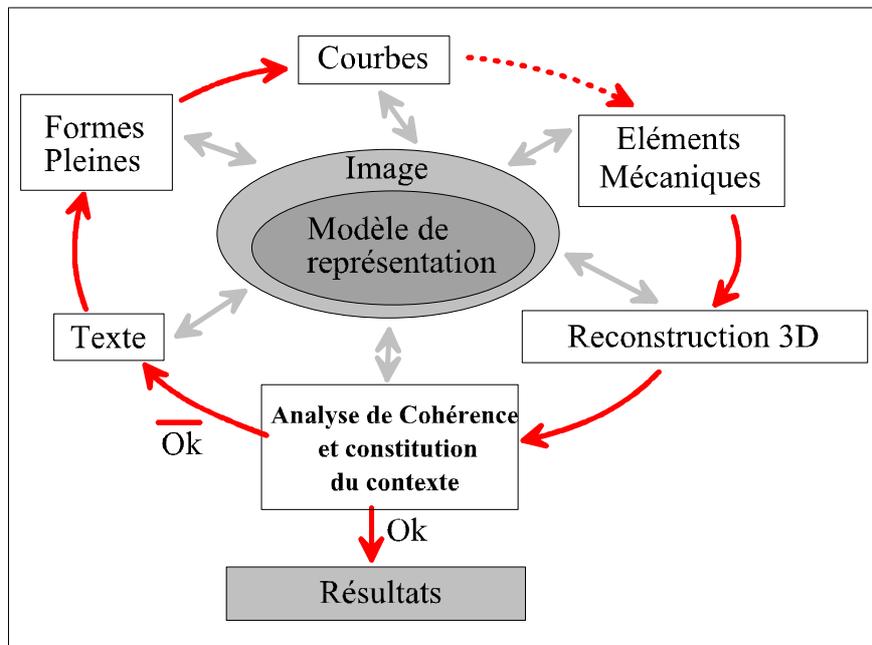


Figure 3.39 : Cycle perceptif

Le choix de l'ordre d'activation des différents spécialistes est un problème qui caractérise cette méthode, comme nous l'avons déjà souligné. Ce choix de l'ordre d'activation se base, pour l'instant, sur les connaissances a priori tirées de l'étude de la structure des documents. La stratégie exploratoire de capture de l'information est ainsi définie par l'utilisateur, l'expert qui doit mettre en place le cycle.

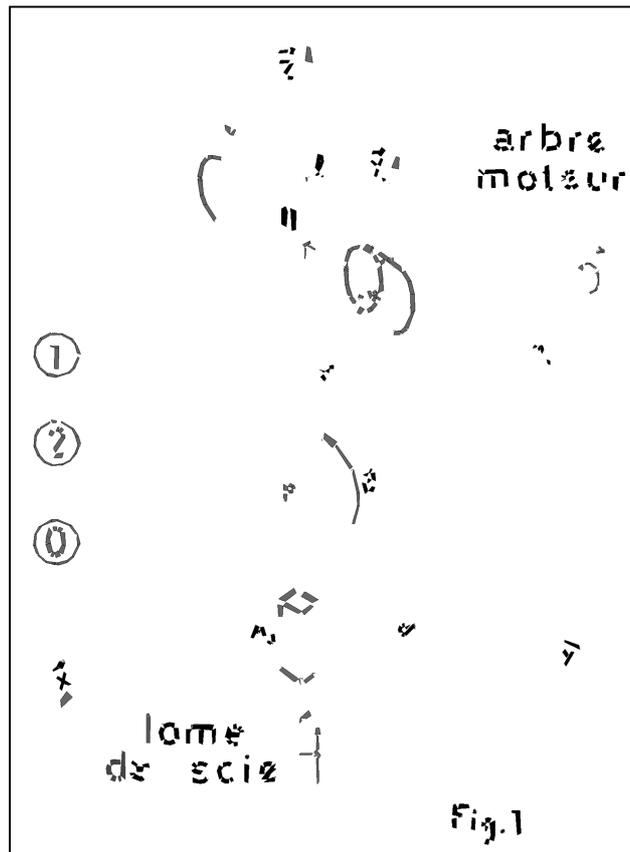
Au premier abord, le 1^{er} cycle perceptif peut sembler différent des autres puisqu'il n'existe encore aucun objet du type recherché par le spécialiste activé à un instant donné. Une hiérarchie est ainsi associée au modèle de représentation. En fait, par la méthode choisie : cycle perceptif + minimisation de la prise de risque, nous voulons que le changement de l'ordre d'activation des processus n'influe pas sur le résultat final, mais seulement sur la rapidité de l'interprétation. Certains objets sont plus facilement localisables que d'autres ou sont un préalable à la localisation d'entités de plus haut niveau mais tous sont détectés de la même façon à plus ou moins longue échéance. Si on lance le spécialiste chargé de réaliser la reconstruction 3D alors qu'aucune liaison n'a encore été détectée, il ne pourra fournir aucun résultat.

Chaque spécialiste contribue à simplifier et à enrichir le modèle de représentation. Des données ambiguës à un instant donné, pour un spécialiste donné, peuvent constituer des évidences pour

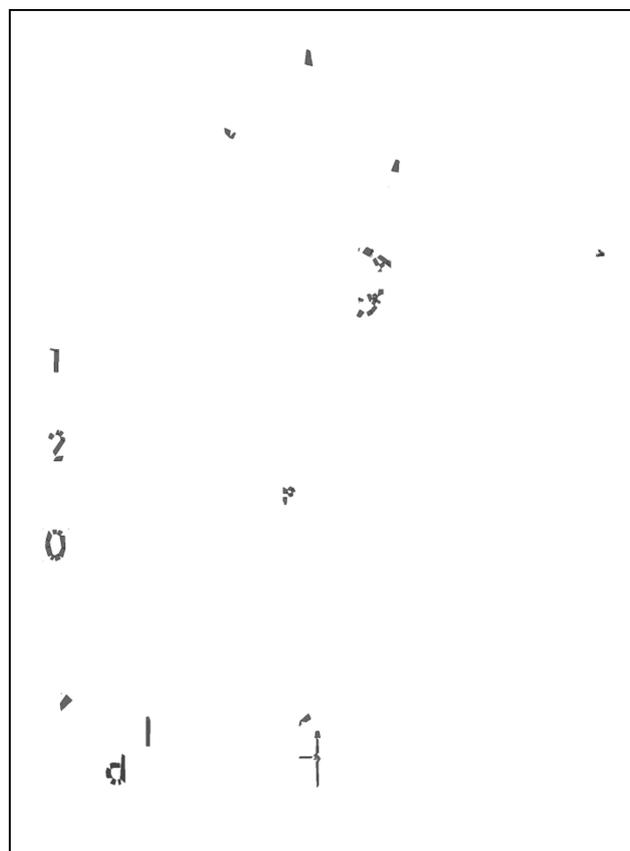
un autre spécialiste. Ainsi lors d'une prochaine activation d'un spécialiste le modèle aura été simplifié et le nombre d'ambiguïtés réduit par rapport à l'état courant d'interprétation.

La figure 3.40a illustre le phénomène pour le spécialiste de l'extraction du texte : lors de sa première activation, ce dernier reconnaît de nombreuses zones de texte sans difficulté (zones de couleur noire) mais relève aussi quelques zones d'indécision (zones de couleur grise).

L'activation des autres spécialistes durant le cycle courant va permettre de résoudre bon nombre de ces ambiguïtés. La figure 3.40b, décrivant les zones restant indéterminées avant la deuxième activation de ce spécialiste, montre qu'un nombre conséquent d'indécisions ont été levées par les autres processus d'analyse.



(a) 1^{er} cycle perceptif



(b) 2^e cycle perceptif

Figure 3.40 : Evolution du modèle de représentation

L'interprétation est terminée lorsqu'un cycle perceptif entier n'apporte plus aucune modification au modèle de représentation. La compréhension est complète si les données fournies par le modèle de représentation ne présentent aucune zone d'indécision. Dans le cas contraire, le système interrogera l'utilisateur qui devra l'aider à comprendre et à résoudre les problèmes insolubles automatiquement.

Coopération

Les problèmes parfois complexes de choix de la répartition des connaissances a priori et du raisonnement dans le système sont résolus à l'aide du modèle de représentation et du mécanisme d'activation cyclique de spécialistes.

Le second travail des spécialistes consiste à fournir le résultat de leurs investigations aux autres spécialistes. Pour cela, ils vont maintenir à jour ce qui correspond à l'état actuel de l'interprétation du document dans le système : c'est à dire modifier et faire évoluer le modèle de représentation du document vers la solution finale. Un spécialiste ne perçoit que ce à quoi il prête attention, et ne reconnaît que ce qu'il cherche. Les cycles permettent de relier les différentes perceptions, de combiner des opérations pour atteindre un but commun.

Pour l'instant les spécialistes n'exploitent que rarement le contexte d'un point de vue sémantique, ils utilisent le fait que le modèle a été simplifié mais pas son enrichissement, il serait très intéressant d'ajouter de telles connaissances aux différents spécialistes, peut-être sous forme de règles traduisant des relations entre les entités de haut niveau [Habacha93b] : liens entre zones de texte et liaisons mécaniques, liens entre courbes et zones de texte, liens entre courbes et courbes,

Analyse de la cohérence

Chaque spécialiste, à son niveau, utilise les connaissances dont il dispose pour renvoyer des données cohérentes. Il analyse la cohérence du modèle de représentation de son point de vue par confrontation avec ses propres connaissances. Suivant le travail qu'il a à réaliser, un spécialiste gère des entités de plus ou moins haut niveau sémantique. Pour l'instant, un spécialiste ne peut pas remettre en cause les données fournies par ses confrères et les conflits entre spécialistes sont interdits : un spécialiste ne peut pas utiliser une entité du modèle de représentation déjà conquise par un autre spécialiste. Ceci est cependant envisageable, sans grande modification du système. De récents travaux réalisés en analyse d'images [Lefèvre96]

[Charroux96] montrent que cette voie est prometteuse mais mentionnent aussi de nombreux problèmes ne pouvant être traités dans le cadre de cette thèse; on peut mentionner :

- la difficulté de définir exhaustivement les conditions d'utilisation des traitements (spécialistes) en fonction du contexte (appels d'offre),
- la difficulté d'évaluer les résultats des traitements,
- la difficulté d'élaborer des stratégies fusionnant des résultats fournis par des traitements différents.

3.4. Résultats

3.4.1. Un exemple complet

Ce dernier exemple de résultats, obtenu sur un schéma fait à main levée (figure 3.41), montre les avantages et les lacunes de notre système d'interprétation. La difficulté de traitement de tels dessins est importante et dépend beaucoup de la qualité avec laquelle ceux-ci ont été réalisés. Généralement, les dessinateurs en mécanique sont entraînés et habiles et leurs dessins sont de qualité suffisante pour que l'interprétation soit réalisée correctement par notre système.

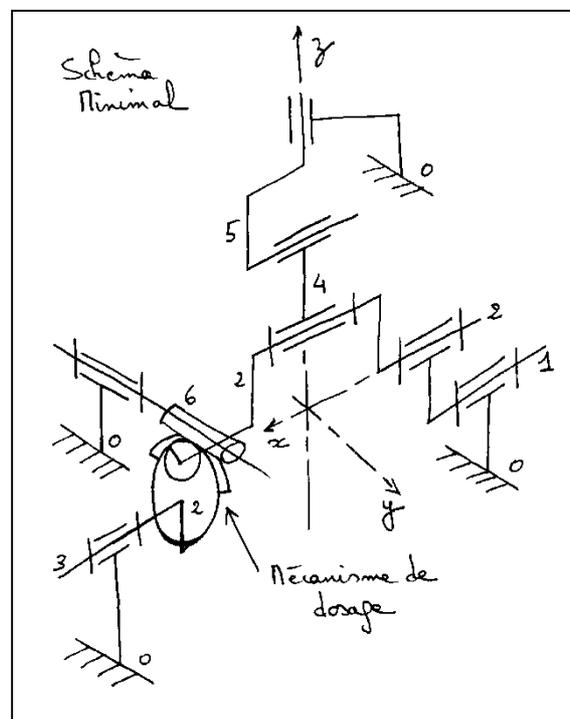


Figure 3.41 : Image initiale

Il existe de grandes différences entre un dessin réalisé avec l'aide d'outils informatiques et un dessin fait à main levée ; le système que nous avons développé est capable d'analyser les dessins manuscrits, cependant, afin que ces derniers soient mieux appréhendés, il est nécessaire d'adapter certains paramètres utilisés durant la phase de construction de la vision globale du document. Les traits dessinés à main levée sont moins rectilignes que ceux obtenus à l'aide de logiciels, les seuils régissant les approximations polygonales effectuées successivement (construction des vecteurs et des quadrilatères) ont donc été modifiés afin que de légers écarts dans les tracés soient tolérés. D'autres paramètres ont été fixés et sont utilisés par les différents spécialistes. Ils pourraient eux aussi être modifiés suivant la qualité et le type du document à analyser ; nous avons néanmoins préféré ne pas changer ces derniers pour toutes les expérimentations présentées afin que les résultats fournis soient aisément comparables.

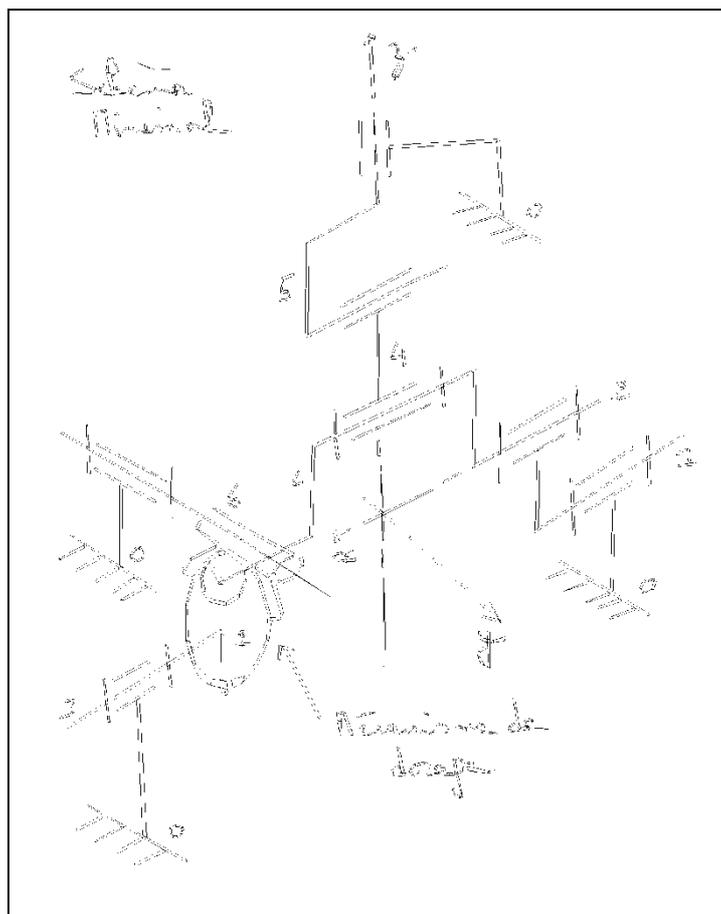


Figure 3.42 :Image des quadrilatères

Le premier résultat fourni figure 3.42 montre la représentation de l'image, obtenue après la phase de vision globale (c'est à dire la mise en place des primitives du modèle de

représentation). Le dessin ne comportant pas de formes pleines, l'intégralité de l'information se trouve contenue dans l'ensemble des quadrilatères.

Le premier cycle perceptif débute par l'activation du spécialiste chargé d'extraire les zones de texte. On peut remarquer, qu'à cet instant, le nombre des zones d'indécision auxquelles est confronté le spécialiste est assez important (figure 3.43). Néanmoins les caractères isolés et les caractères non attachés sont détectés sans problème (figure 3.44). On peut remarquer que les têtes de flèches des axes x et y sont considérées comme des caractères.

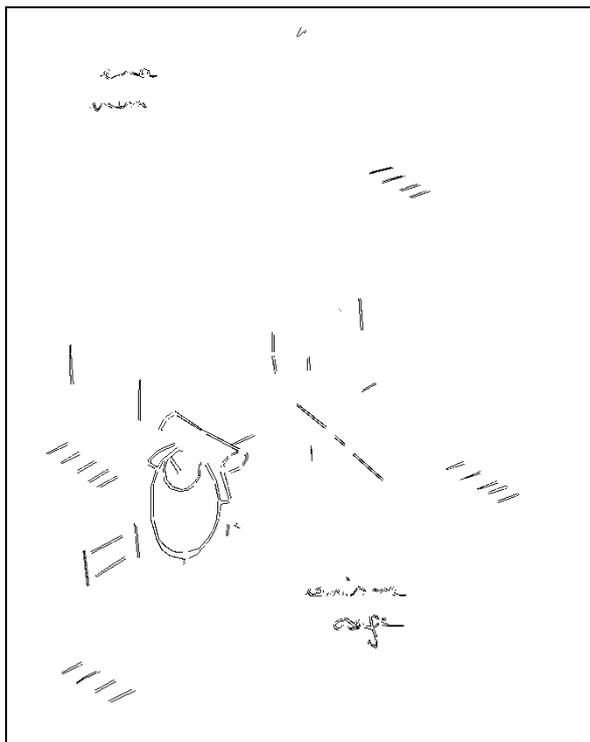


Figure 3.43 : Zones d'indécision
(1^{er} cycle - spécialiste du texte)

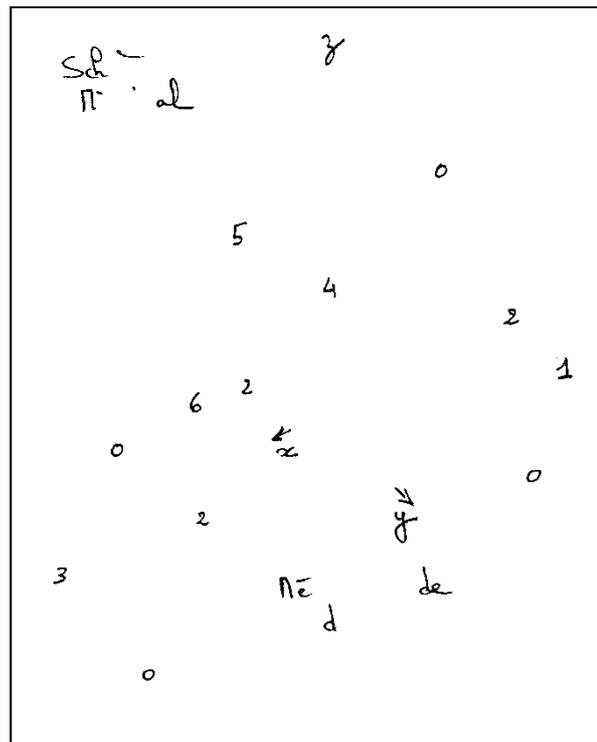


Figure 3.44 : Zones de texte localisées

Le nombre de tirets successifs constituant chaque trait pointillé (présent au centre de l'image) n'est jamais suffisant pour que leur localisation constitue une évidence (4 tirets, au minimum, sont nécessaires), le spécialiste «Pointillés» ne fournit donc aucun résultat supplémentaire.

La figure 3.45 décrit les résultats obtenus par le processus chargé de localiser les courbes présentes dans l'image. Celles-ci sont bien localisées, néanmoins une succession de quadrilatères traduisant un trait non rectiligne a été prise pour une courbe et approximée par une courbe de Bézier. Bien que cette courbe ne corresponde pas vraiment à un tracé non

rectiligne (à une véritable courbe), ce résultat permet de regrouper en une entité unique plusieurs quadrilatères.

Les entités mécaniques sont, quant à elles, localisées correctement (figure 3.46) ; seule une liaison pivot fixe a été confondue avec une liaison pivot glissant. L'erreur est due, comme nous l'avons déjà souligné précédemment, à la mauvaise représentation d'un trait : un trait mal dessiné (non rectiligne) s'est traduit par la génération de 2 quadrilatères. Une solution de ce problème est donnée paragraphe 3.2.4.2.

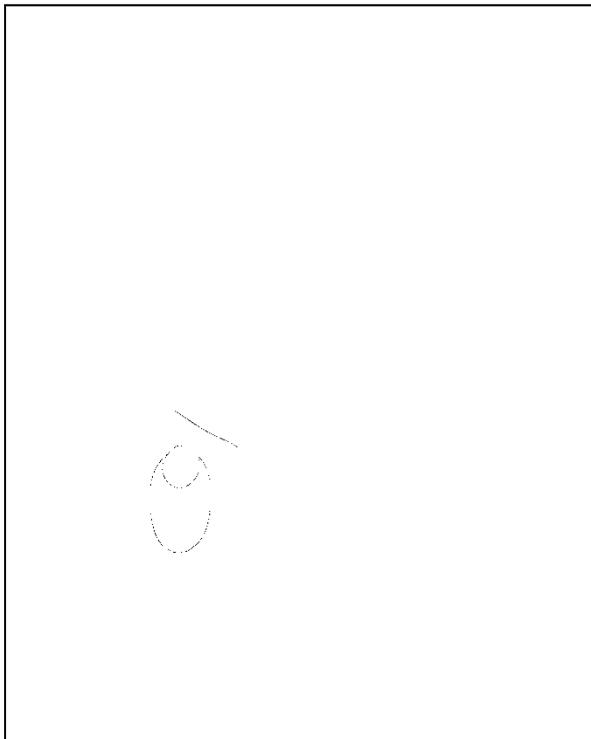


Figure 3.45 : Courbes localisées

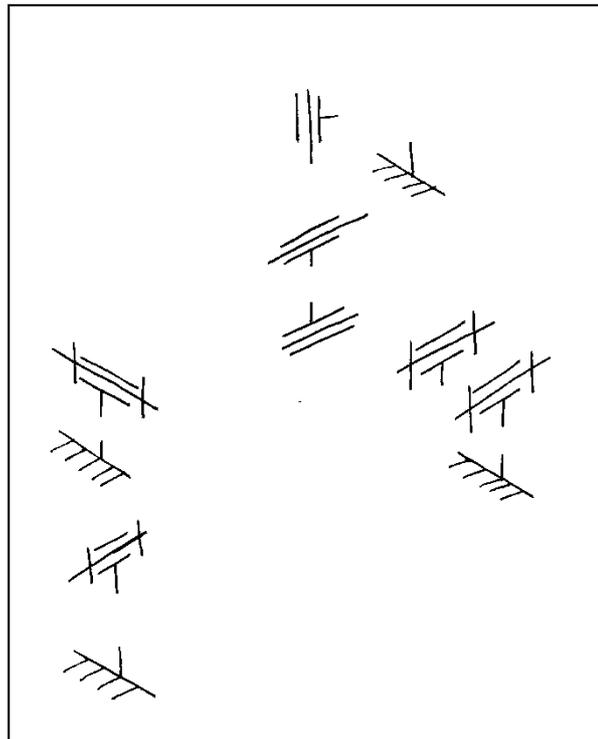


Figure 3.46 : Entités mécaniques localisées

La reconstruction 3D a permis de déterminer les 3 orientations principales du dessin (figure 3.47), de positionner les différentes liaisons les unes par rapport aux autres et de poursuivre la reconstruction des pièces constituant le mécanisme représenté après étiquetage sur la figure 3.48. Le positionnement des liaisons les unes par rapport aux autres n'est pas réalisé complètement car certaines liaisons n'ont pas pu être mises en correspondance. Le résultat de l'interprétation mécanique du schéma est stocké sous forme de fichier (Format MécaMaster) afin de pouvoir être exploité par des logiciels de CAO ou CFAO. Ce fichier permet de connaître précisément le positionnement relatif des différentes liaisons (le fichier correspondant à cet exemple est fourni en fin de paragraphe).

Il est à noter que la représentation du mécanisme sur le dessin de l'auteur comporte une ambiguïté dans l'orientation donnée à une pièce ; une représentation correcte de la partie correspondante est donnée figure 3.51. Parce qu'aucune des 3 orientations principales déterminées lors de la reconstruction 3D n'est respectée, le positionnement de la liaison carter par rapport aux autres ne peut être réalisé. Les zones comportant de telles ambiguïtés nécessitent une réelle interaction entre le système d'interprétation et le dessinateur afin que les données finales reflètent véritablement les intentions initiales du concepteur. Le système joue alors plus le rôle d'un logiciel d'aide à la conception que d'un simple logiciel de saisie automatique.

Le second cycle perceptif n'apportera ici que peu d'informations supplémentaires : les zones de texte manuscrites seront localisées par présentation des zones d'indécisions à l'utilisateur (figure 3.50). Le nombre de ces zones a cependant considérablement diminué (figure 3.49).

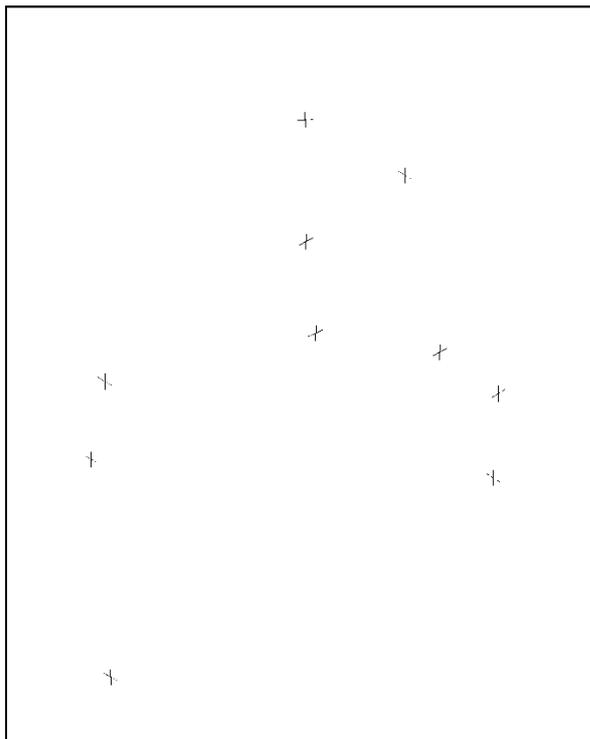


Figure 3.47 : Orientations des liaisons

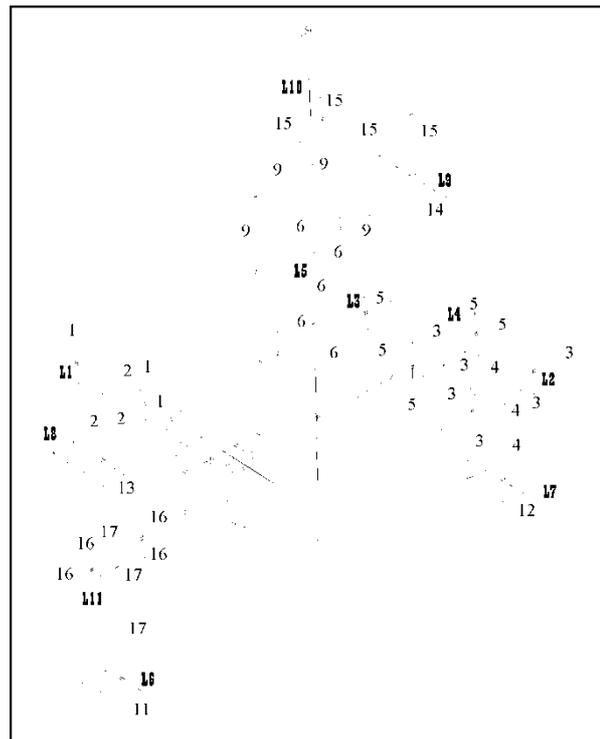


Figure 3.48 : Reconstruction 3D

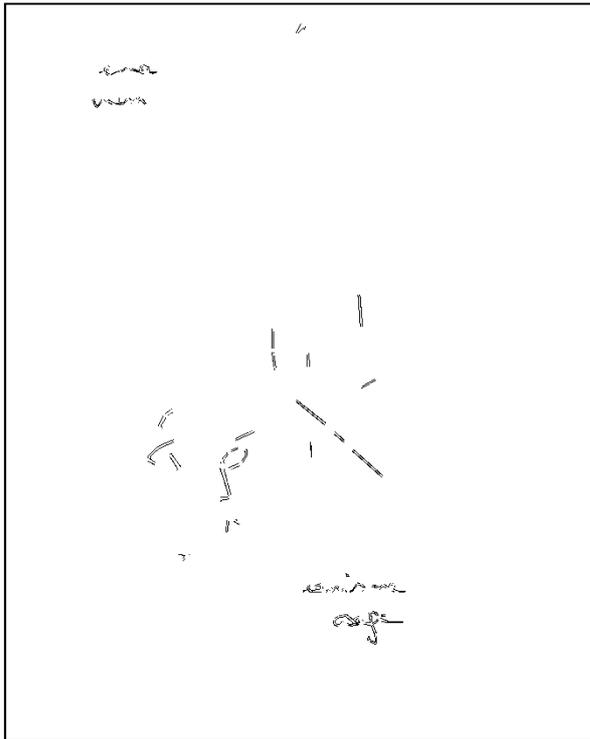


Figure 3.49 : Zones d'indécision (2^e cycle)

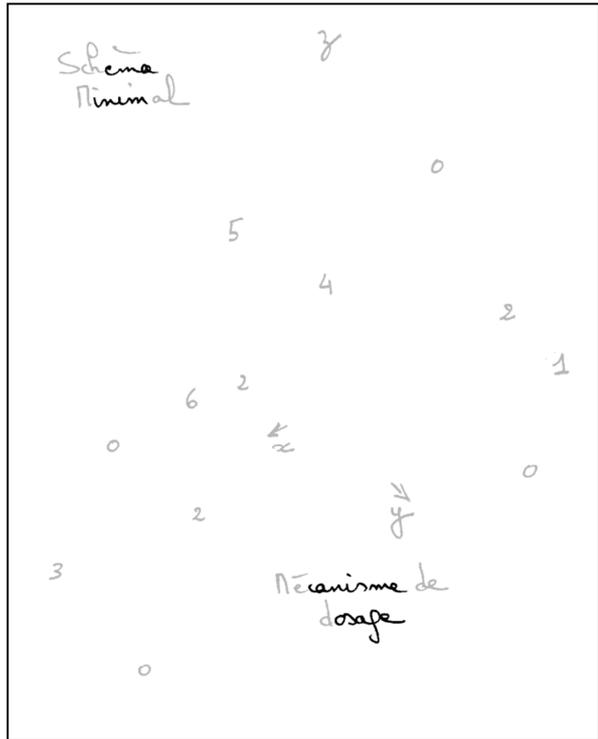
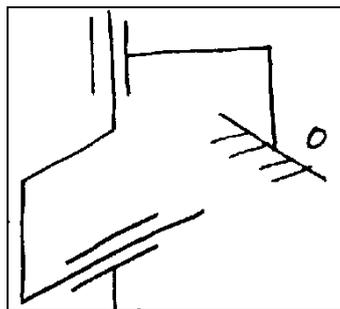
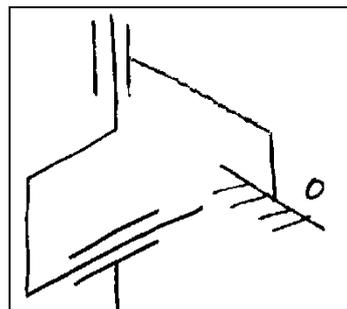


Figure 3.50 : Zones de textes
(en grisé : zones détectées automatiquement,
en noir : par présentation à l'utilisateur)



(a)



(b)

Figure 3.51 : (a)Partie de l'image initiale (b)Représentation correcte

Exemple de fichier de sortie

Fichier Résultats : mld.lad - (Image Mldoseur.tif) -

*** Informations Générales :

Schéma 3D
Orientations..... 22.500 90.000 130.000

**** DONNEE numéro 1 ****

Liaison PIVOT FIXE.
sur : Piece 1.
Piece 2.

Au point 235.000 905.000 0.000
Direction -0.857 0.515 0.000
Référence 0
Position 0.000 0.000 0.000

**** DONNEE numéro 2 ****

Liaison PIVOT FIXE.
sur : Piece 3.
Piece 4.

Au point 1172.000 934.000 0.000
Direction 0.839 0.545 0.000
Référence 0
Position 0.000 0.000 0.000

**** DONNEE numéro 3 ****

Liaison PIVOT GLISSANT.
sur : Piece 5.
Piece 6.

Au point 737.000 788.000 0.000
Direction 0.906 0.423 0.000
Référence D2
Position -502.000 399.000 0.000

FIN

```

**** DONNEE numéro 4 *** -----
Liaison ..... PIVOT FIXE.
      sur : Piece 5
           Piece 3
Au point ..... 1032.000  834.000  0.000
Direction ..... 0.891  0.454  0.000
Référence ..... D2
Position ..... -190.000  196.000  0.000

```

```

**** DONNEE numéro 5 *** -----
Liaison ..... PIVOT GLISSANT.
      sur : Piece 9.
           Piece 6.
Au point ..... 714.000  566.000  0.000
Direction ..... 0.891  0.454  0.000
Référence ..... D2
Position ..... -502.000  627.000  0.000

```

```

**** DONNEE numéro 6 *** -----
Liaison ..... CARTER.
      sur : Piece 11.
           Piece 17
Au point ..... 249.000  1620.000  0.000
Direction ..... -0.866  0.500  0.000
Référence ..... 0
Position ..... 0.000  0.000  0.000

```

```

**** DONNEE numéro 7 *** -----
Liaison ..... CARTER.
      sur : Piece 12.
           Piece 4.
Au point ..... 1159.000  1138.000  0.000
Direction ..... -0.848  0.530  0.000
Référence ..... D2
Position ..... 0.000 -240.000  0.000

```

```

**** DONNEE numéro 8 *** -----
Liaison ..... CARTER.
      sur : Piece 13.
           Piece 2.
Au point ..... 202.000  1093.000  0.000
Direction ..... -0.839  0.545  0.000
Référence ..... D1
Position ..... 0.000  196.000  0.000

```

```

**** DONNEE numéro 9 *** -----
Liaison ..... CARTER.
      sur : Piece 14.
           Piece 15.
Au point ..... 950.000  406.000  0.000
Direction ..... -0.848  0.530  0.000
Référence ..... 0
Position ..... 0.000  0.000  0.000

```

```

**** DONNEE numéro 10 *** -----
Liaison ..... PIVOT GLISSANT.
      sur : Piece 9.
           Piece 15.
Au point ..... 712.000  272.000  0.000
Direction ..... -0.052  0.999  0.000
Référence ..... D2
Position ..... -497.000  832.000  0.000

```

```

**** DONNEE numéro 11 *** -----
Liaison ..... PIVOT FIXE.
      sur : Piece 16.
           Piece 17.
Au point ..... 246.000  1320.000  0.000
Direction ..... 0.906  0.421  0.000
Référence ..... D6
Position ..... 0.000 -306.000  0.000

```

```

##### FIN #####

```

3.4.2. Images tests

Les méthodes d'analyse présentées ont été testées sur un ensemble d'images provenant de sources diverses. En ce qui concerne la première phase du traitement (perception globale du document), les images choisies correspondent à différents types de documents afin de prouver la généralité des techniques employées. La seconde phase : lecture et interprétation du document, nécessite des connaissances sur le domaine et a donc été développée pour un type précis de documents: les schémas cinématiques mécaniques. Les résultats obtenus sur ce type de dessins sont très encourageants et les techniques utilisées semblent, de plus, facilement généralisables à d'autres types de documents. Nous n'avons retenu, pour la rédaction de ce mémoire, que les images démontrant les résultats intéressants ou les lacunes des méthodes employées. On retrouve notamment les résultats commentés correspondant aux images tests suivantes :

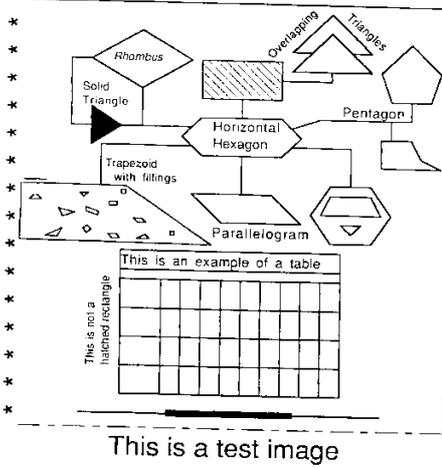
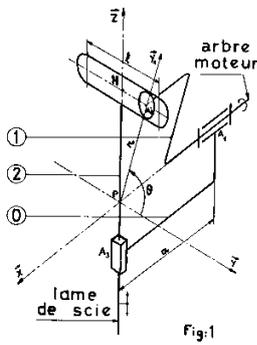


Fig. 17. A test image containing different shapes [Kasturi et al. 90].

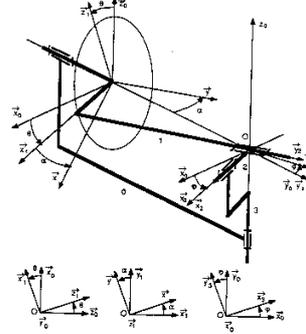
voir pages: 58, 61, 76, 78, 93, 98, 102



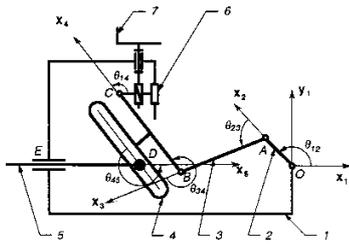
voir pages: 81, 82



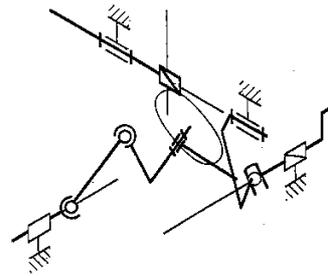
voir pages: 71, 93, 95, 102, 112, 135



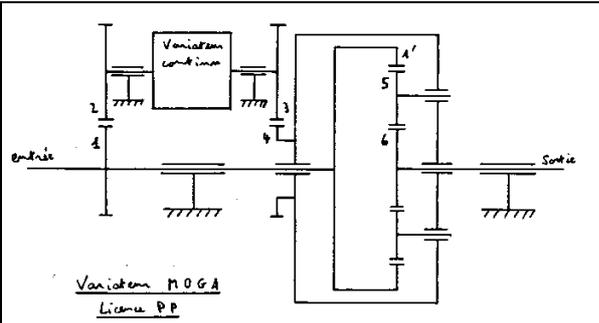
voir pages: 67, 72, 73, 108, 116, 120, 125, 129, 130



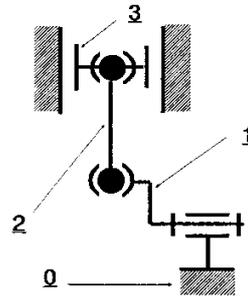
voir pages: 58, 68



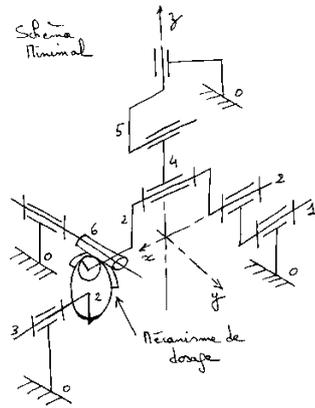
voir pages: 112, 131



voir pages: 62, 95, 120, 123, 131



voir pages: 65, 75, 78



voir pages: 137 à 143

Conclusion

Tout au long de ce mémoire nous avons présenté une nouvelle méthode d'interprétation automatique de documents techniques et plus particulièrement des schémas cinématiques mécaniques. La méthode proposée a pu être élaborée grâce à une longue concertation avec des « mécaniciens », en particulier ceux de l'Université de Grenoble que nous remercions ici.

Le premier chapitre, en décrivant l'évolution des recherches dans ce domaine, montre que lors des premiers temps, les systèmes se contentaient de reconnaître des formes géométriques dans des images binaires (hachures, cercles, caractères, ...); ceci a permis le développement de nombreuses méthodes réalisant ces tâches de bas niveau ; ces méthodes ont donné des résultats jugés parfois satisfaisants. Cependant, les résultats finals obtenus par les systèmes commerciaux de l'époque ont vite montré qu'il n'était pas suffisant de récupérer des formes géométriques dans un dessin pour répondre aux souhaits des industriels. De plus, erreurs et oublis étaient assez fréquents dès que l'on se hasardait à prendre des schémas un peu complexes.

Les recherches ont donc évolué et provoqué l'apparition de nouveaux types de problèmes auxquels cette thèse tente de répondre tout en proposant aussi de nouvelles techniques d'extraction des formes.

Le premier problème, fondamental, se situe au niveau de la récupération d'entités ayant un sens vis à vis du domaine concerné. Pour atteindre ce but, il a été nécessaire de mettre au point de nouvelles méthodes mettant en jeu des problèmes du ressort de l'intelligence artificielle distribuée tels que la mise en commun des données extraites et des résultats, l'utilisation conjointe de différentes méthodes d'extraction et techniques d'analyse, ... De nouvelles méthodes, utilisant au maximum les aspects sémantiques d'un domaine et le contexte obtenu au fur et à mesure de l'interprétation, ont aussi été développées pour extraire directement des entités de haut niveau.

Originalité de la méthode

Ainsi, dans cette thèse, nous avons choisi d'utiliser une nouvelle approche pour analyser les images de documents [Ramel96b]. Nous nous sommes constamment efforcé de nous inspirer le plus possible des enseignements que nous avons obtenus par l'étude des résultats récents sur les mécanismes de perception et de reconnaissance utilisés par le cerveau humain. Sans doute sommes-nous encore infiniment éloigné de la compréhension de son fonctionnement en raison de l'insuffisance de précision de nos connaissances biologiques, toutefois nous pensons que les

remarques que nous en avons tiré nous ont permis de réaliser une légère avancée dans le domaine de l'interprétation automatique de documents.

Pour interpréter, l'homme extrait d'abord un ensemble de caractéristiques du document pour établir un contexte qu'il utilise ensuite pour valider les informations qu'il obtient durant la lecture. Pour cela, le lecteur enchaîne sans cesse des activités complexes telles que la prise d'indices (graphiques, lexicaux, syntaxiques, sémantiques), l'élaboration d'hypothèses permettant d'anticiper, et la réalisation de corrections éventuelles.

Ainsi, nous avons choisi de commencer le traitement par une phase d'extraction d'informations sur l'aspect global du document initial. Les méthodes utilisées durant cette phase doivent permettre l'extraction d'un maximum d'indices et de données sur le document initial. Cette étape d'extraction des caractéristiques globales permet d'acquérir les éléments indispensables à la réalisation d'une analyse locale plus adaptée et donc plus précise. Ces connaissances sont obtenues et gérées à l'aide d'un modèle de représentation performant, à même de décrire toutes les formes présentes dans le document initial. Les primitives utilisées sont le *vecteur* pour les formes pleines et le *quadrilatère* pour les formes fines.

Les méthodes utilisées pour obtenir cette vision globale sont applicables sur la plupart des dessins techniques (constitués de courbes, traits, textes, formes pleines, ...). Cette phase permet de plus, de résoudre les problèmes souvent complexes dus à l'amorçage du processus d'interprétation, résolus ici par la vision globale qui fournit un contexte initial pour l'interprétation.

La prise en compte du contexte durant l'analyse n'est cependant pas une chose triviale; elle nécessite la modification des techniques traditionnelles. Il est nécessaire de leur adjoindre des mécanismes les rendant capables de tirer parti du contexte. Nous avons opté pour l'utilisation de zones de focalisation générées à l'aide du modèle de représentation issu de l'image (vision globale) et la mise en place de liens contextuels définissant des relations entre entités déjà reconnues et entités à localiser.

La plate-forme développée permet d'obtenir une *perception globale d'un document* et fournit une approche intéressante en apportant une abstraction organisée de l'image tout en préservant la capacité du système à accéder aux informations de bas niveau. Cette représentation n'est pas rigide mais au contraire évolutive, ce qui facilite le passage de la phase de récupération d'information à la phase de reconnaissance.

Le trop grand nombre d'erreurs encore commises par les systèmes automatiques est aussi la cause du succès très mitigé des systèmes réalisés jusqu'à présent. La solution se situe, là aussi, dans l'utilisation de la sémantique et du contexte pour réduire le nombre d'erreurs. La mise en commun des résultats permet de vérifier la cohérence de l'interprétation mais aussi la génération des hypothèses se trouvant généralement à la base de l'évolution du mécanisme d'interprétation.

La gestion des incohérences est une tâche ardue, il nous a semblé plus simple d'empêcher au maximum les erreurs d'interprétation plutôt que de mettre en place des méthodes complexes de retour-arrière. La correction d'une incohérence peut d'ailleurs provoquer l'apparition de nouvelles incohérences dont la gestion devient quasiment impossible : effet « boule de neige » pouvant remettre la totalité de l'analyse en cause. Un système doit gérer des incohérences s'il prend en compte le contexte a posteriori, c'est à dire pour vérifier la cohérence des analyses. Il est à notre avis plus judicieux d'utiliser le contexte lors de la génération des hypothèses, pour vérifier celles-ci, d'effectuer des analyses locales ou même de retourner consulter l'image initiale, de confronter plusieurs points de vue et de gérer les éventuels conflits,

Pour continuer, il nous a fallu remarquer que l'homme possède une habileté très forte à extraire les évidences en faisant appel aux vastes connaissances qu'il a acquises. Lorsqu'il lit, des échanges s'effectuent constamment entre l'information visuelle (image du document) et l'information non visuelle (arrière plan conceptuel : ce que l'on sait déjà sur le document). Il semble alors normal qu'un système artificiel s'attache, lui aussi, à extraire les évidences en premier puis, au fur et à mesure de la simplification de la scène à interpréter, et de l'enrichissement du contexte, tente d'interpréter les éléments plus ambigus et plus complexes.

Pour mettre en place un tel fonctionnement, nous avons choisi de tirer parti des systèmes multi-spécialistes. La « coopération intelligente » d'un nombre assez grand de spécialistes et du modèle de représentation (pour le partage de l'information) nous permet, au cours de la reconnaissance et de la reconstruction des objets, d'utiliser de façon duale vision globale et vision locale.

Les problèmes liés à la répartition des connaissances dans le système et à l'adaptabilité du système à d'autres types de document sont, par là même, résolus par l'utilisation de spécialistes indépendants, réalisant chacun un travail précis. De plus, les résultats fournis par chacun de ces experts seront englobés dans le modèle de représentation de manière à les rendre visibles et exploitables par tous. Le modèle de représentation constitue le contexte et doit servir à la génération d'hypothèses durant les futures analyses réalisées.

Il a été nécessaire de trouver une technique permettant le partage de l'information et surtout l'extraction des évidences. Pour cela, il est obligatoire que chaque spécialiste puisse intervenir à différents instants du traitement, qu'il puisse consulter l'état d'avancement de l'interprétation et apporter, si cela est possible, et sans risque d'erreur, sa contribution à l'interprétation. Une exécution séquentielle ordonnée de l'ensemble des spécialistes durant ce que l'on a décidé d'appeler un cycle perceptif, constitue une méthode simple et efficace de gestion de la coopération entre les différents spécialistes tout en leur permettant de n'extraire que ce qui constitue pour eux des évidences à un instant donné. Ce mécanisme de cycles perceptifs permet aux spécialistes de coopérer pour limiter les erreurs d'analyse. Chaque spécialiste réalise une tâche précise telle que par exemple la détection du texte, des courbes, des éléments mécaniques ou la reconstruction 3D du mécanisme. En fin de chaque cycle, la complétude de la compréhension du document est vérifiée et un nouveau cycle d'activation des spécialistes approfondit, si nécessaire, l'interprétation. Ainsi, le contexte évolue et fait apparaître, sans cesse, de nouvelles évidences tout en augmentant la structuration des connaissances acquises sur le document au cours de chacun des cycles perceptifs jusqu'à obtenir la compréhension globale du document.

La localisation des courbes (par le spécialiste chargé d'extraire les éléments non linéaires) montre l'intérêt d'une représentation utilisant les courbes de Bézier pour l'étude des plans cinématiques. Pour la reconnaissance des entités mécaniques et l'analyse mécanique du mécanisme dessiné (reconstruction 3D), nous proposons une méthode originale de mise en correspondance de graphes attribués à base d'étiquetages structurels.

La description détaillée des techniques utilisées par les différents spécialistes montre les adaptations et améliorations apportées aux techniques classiques d'extraction des formes pour permettre la prise en compte du contexte. C'est par l'utilisation de la vision globale, permettant la génération de zones de focalisation, que chaque spécialiste effectue les études poussées de chaque zone susceptible de comporter des informations intéressantes. Les phases d'analyses locales permettent dans chaque cas de vérifier la pertinence de l'hypothèse et ainsi de limiter le nombre d'erreurs commises.

Grâce à cette démarche, nos travaux se distinguent de ceux employant la chaîne de traitements habituelle : pré-traitements - segmentation - analyse - reconnaissance. Le système s'efforce d'extraire aussi bien les formes que les connaissances, de manière hiérarchique et réfléchie, pour

les transcrire dans un modèle de représentation évolutif durant chaque cycle perceptif et en utilisant la vision globale du document préalablement construite.

Perspectives

Nous ne nous dissimulons pas pourtant, à la fin de ce travail, tout ce qui reste à faire. Nous n'avons effectué qu'un premier pas vers la solution. Nous souhaitons que notre travail puisse être étendu, de façon approfondie. Certains points nous semblent particulièrement prometteurs :

- L'amélioration des traitements réalisés par les différents spécialistes. Il est notamment nécessaire de leur donner la capacité, d'apprendre (au cours des cycles perceptifs successifs) et d'améliorer leur coopération en leur ajoutant la capacité de gérer des conflits d'appropriation d'entités ou de primitives (concurrence, gestion de conflits, ...), de leur adjoindre la possibilité d'utiliser au maximum les résultats fournis par les analyses de leurs confrères : ceci nécessite la mise en place de liens contextuels traduisant les relations pouvant exister entre les positions relatives des différentes entités constituant le document.

- L'adjonction d'artifices permettant d'adapter les paramètres des extracteurs en fonction de l'état du contexte, pour améliorer la performance de l'interprétation durant chaque cycle perceptif.

- La mise en place de nouveaux outils d'extraction ou l'amélioration de ceux existant permettant, par exemple, la localisation des têtes de flèches, des cotations, ... afin d'enrichir encore le contexte et d'améliorer l'analyse.

Pour ce qui concerne la perception globale du document, la généralité de la méthode a été testée avec succès notamment sur des images de partitions musicales et sur des organigrammes. Cependant, certains seuils et paramètres ont été fixés de manière empirique, ces choix devront donc être validés. La partie lecture et interprétation du document a été plus spécialement conçue pour traiter les schémas cinématiques mais son adaptation à d'autres types de documents ne semble poser aucun problème étant donné la philosophie utilisée. La seule difficulté pourrait résider dans la conception et l'ordonnancement des nouveaux spécialistes. La meilleure façon de vérifier est certainement d'employer cette méthode pour interpréter un autre type de documents. Des essais sur les partitions musicales et les croquis à main levée nous ont donné des résultats encourageants. Nous souhaitons enfin que notre travail puisse apporter une petite contribution à la résolution du problème général de la reconnaissance des formes et nous espérons, en tout cas, avoir provoqué, chez certains d'entre vous, l'envie d'utiliser une méthode similaire et ainsi d'améliorer encore ses performances.

Références Bibliographiques

- [Abe86] **Abe, K., Azumatani, Y., Mukouda, M. and Suzuky, S.** Discrimination of symbols, lines, and characters in flowchart recognition. In : *Proceedings of the 8th International Conference on Pattern Recognition, Paris (France), october 27-31, 1986.* Vol. 2, p. 1071-1074.
- [AhSoon95] **Ah-Soon, C. and Tombre, K.** A step towards reconstruction of 3D CAD models from engineering drawings. In : *Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal (Canada), august 14-16, 1995.* Vol. 1, p. 331-334.
- [Antoine91] **Antoine, D.** CIPLAN : A ruled-based system with original features for understanding French plats. In : *Proceedings of the First International Conference on Document Analysis and Recognition, Saint-Malo (France), september 30- october 2, 1991.* p. 647-655.
- [Antoine92] **Antoine, D.** Techniques spécialisées et représentation de la connaissance pour l'interprétation des plans cadastraux. *Bigre*, 1992, N° 80, p. 65-73.
- [Arias93] **Arias, J.F., Lai, C.P., Chandran, S., Kasturi, R. and Chhabra, A.** Interpretation of telephone manhole drawings. In : *Proceedings of the 2nd International Conference on Document Analysis and Recognition, Tsukuba (Japon), october 20-22, 1993.* p. 365-368.
- [Arias94] **Arias, J.F., Prasad, A., Kasturi, R. and Chhabra, A.** Interpretation of telephone central office equipment drawings. In : *Proceedings of the 12th International Conference on Pattern Recognition, Jérusalem (Israël), october 9-13, 1994.* Vol. 2, p. 310-314.
- [Baumann95] **Baumann, S.** A simplified attributed graph grammar for high level music recognition. In : *Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal (Canada), august 14-16, 1995.* Vol. 2, p. 1080-1083.
- [Belaïd92a] **Belaïd, A. et Tombre, K.** Analyse de documents : de l'image à la sémantique. *Bigre*, 1992, N° 80, p. 3-28.
- [Belaïd92b] **Belaïd, A. et Belaïd, Y.** *Reconnaissance de formes : méthodes et applications.* Paris : InterEdition, 1992. 429 p.
- [BenYacoub95] **Ben Yacoub, S.** *Extraction hiérarchique robuste d'attributs géométriques dans une image numérique.* Thèse de doctorat : Université Claude Bernard Lyon I, Lyon (France), 1995. 143 p.

- [Besançon88] **Besançon, J.E.** *Vision par ordinateur en deux et trois dimensions*. Paris : Eyrolles, 1988. 411 p.
- [Bley84] **Bley, H.** Segmentation and preprocessing of electrical schematics using picture graphs. *Computer Vision, Graphics and Image Processing*, 1984, Vol. 28, p. 271-288.
- [Boatto92] **Boatto, L. and al.** Detection and separation of symbols connected to graphics in line drawings. In : *Proceedings of the 11th International Conference on Pattern Recognition, The Hague (Netherlands), august 30-september 3, 1992*. Vol. 2, p. 545-548.
- [Bottoni92] **Bottoni, P., Mussio, P. and Protti, M.** Metareasoning as tool for pattern recognition. In : *Proceedings of the 11th International Conference on Pattern Recognition, The Hague (Netherlands), august 30-september 3, 1992*. p. 285-289.
- [Bruce93] **Bruce, V., Green, P.R.** *La perception visuelle : Physiologie, psychologie et écologie*. Grenoble : Presses universitaire de Grenoble, 1993. 511 p.
- [Brun94] **Brun, J.M.** From characteristic shapes to form features : a recognition strategy. In : *Proceedings of the the International Conference on Feature Modelling and Recognition in advanced CAD/CAM systems, Valenciennes (France), 1994*. Valenciennes : IFIP Eds, 1994. p. 315-326.
- [Charroux96] **Charroux, B., Philipp, S. et Cocquerez, J.P.** Système de vision mettant en oeuvre une coopération d'opérateurs de segmentation guidée par l'interprétation. In : *Proceedings du 10^e Congrès Reconnaissance de Formes et Intelligence Artificielle, Rennes (France), 16-18 juin, 1996*. Vol. 2, p. 527-537.
- [Chassery91] **Chassery, J.M. et Montanvert, A.** *Géométrie discrète en analyse d'images*. Paris : Hermès, 1991. 358 p.
- [Collin91] **Collin, S. and Vaxivière, P.** Recognition and use of dimensioning in digitized industrial drawings. In : *Proceedings of the First International Conference on Document Analysis and Recognition, Saint-Malo (France), september 30-october 2, 1991*. p. 161-169.
- [Collin92] **Collin, S.** Grammaire pour la reconnaissance de la cotation dans les dessins techniques. *Bigre*. 1992, N° 80, p. 57-64.
- [Collin93] **Collin, S., Tombre, K. and Vaxivière, P.** Don't tell Mom I'm doing document analysis; she believes I'm in the computer vision field. In : *Proceedings of the 2nd International Conference on Document Analysis and Recognition, Tsukuba (Japan), october 20-22, 1993*. p. 619-622.
- [Cugini84] **Cugini, U., Ferri, G., Mussio, P. and Protti, M.** Pattern directed restoration and vectorization of digitized engineering drawings. *Computer and Graphics*, 1984, Vol. 8, N° 4, p. 337-350.
- [Cugini88] **Cugini, U., Falcidieno, B., Mussio, P. and Protti, M.** Towards automatic indexing of product models in a CIM environment. In : *Proceedings of the*

IEEE Workshop on Languages for Automation, Maryland (USA), august 29-31, 1988. p. 106-114.

- [Cugini90] **Cugini, U., Falcidieno, B. and Mussio, P.** Exploiting knowledge in CAD/CAM architecture. *Intelligent CAD II*, H. Yoshikawa Editor. Amsterdam : Elsevier Science Publishers, 1990. p. 133-148.
- [Denhière92] **Denhière, G., et Baudet, S.** *Lecture, compréhension de texte et science cognitive*. Paris : PUF, 1992. 317 p.
- [Desseilligny94] **Pierrot-Desseilligny, M., Le Men, H. et Stamon, G.** Lecture automatique des écritures sur cartes scannées. In : *Proceedings of the 3^e Colloque National sur l'Écrit et le Document, Rouen (France), 6-8 juillet, 1994*. p. 195-204.
- [Desseilligny95] **Pierrot-Desseilligny, M., Le Men, H. and Stamon, G.** Characters string recognition on maps, a method for high level reconstruction. In : *Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal (Canada), august 14-16, 1995*. Vol. 1, p. 249-252.
- [Dori91] **Dori, D.** Symbolic representation of dimensioning in engineering drawings. In : *Proceedings of the First International Conference on Document Analysis and Recognition, Saint-Malo (France), september 30-october 2, 1991*. p. 1000-1010.
- [Dori92] **Dori, D.** Dimensioning analysis : Towards automatic understanding of engineering drawings. *Communication of the A.C.M.*, 1992, Vol. 35, N°10, p. 92-103.
- [Dori93] **Dori, D. and Tombre, K.** Paper drawing to 3D CAD : a proposed agenda. In : *Proceedings of the 2nd International Conference on Document Analysis and Recognition, Tsukuba (Japon), october 20-22, 1993*. p. 866-869.
- [Dori95] **Dori, D. and Tombre, K.** Form engineering drawings to 3D CAD models : are we ready now ? *Computer Aided Design*, 1995, Vol. 29, N° 4, p. 243-254.
- [Elliman90] **Elliman, D.G.** A review of segmentation and contextual analysis techniques for text recognition. *Pattern Recognition*, 1990, Vol. 23, N°3, p. 337-346.
- [Fahmy91] **Fahmy, H. et Blostein, D.** Une grammaire de graphe pour la reconnaissance de partitions musicales à un niveau élevé. In : *Proceedings of the First International Conference on Document Analysis and Recognition, Saint- Malo (France), september 30-october 2, 1991*. p. 70-78.
- [Fletsher88] **Fletsher, L.A. and Kasturi, R.** A robust algorithm for text string separation from mixed text/graphics images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1988, Vol. 10, N° 6, p. 910-918.
- [Fortin92] **Fortin, C., Rousseau, R.** *Psychologie cognitive : une approche du traitement de l'information*. Québec : Presses de l'université du Québec,

1992. 434 p.

- [Gloger92] **Gloger, J.M.** Use of the Hough Transform to separate merged Text/Graphics in Forms. In : *Proceedings of the 11th International Conference on Pattern Recognition, The Hague (Netherlands), august 30-september 3, 1992*. Vol. 2, p. 268-271.
- [Gordon89] **Gordon, I.E.** *Theories of visual perception*. New-York : John Wiley, 1989. 267 p.
- [Habacha93a] **Habacha-Hamada, A.** A new system for the analysis of schematic diagrams. In : *Proceedings of the 2nd International Conference on Document Analysis and Recognition, Tsukuba (Japon), october 20-22, 1993*. p. 369-372.
- [Habacha93b] **Habacha-Hamada, A.** *Reconnaissance de symboles techniques et analyse contextuelle de schémas*. Thèse de doctorat : Institut National Polytechnique de Lorraine : Nancy (France), 1993. 169 p.
- [Habacha94] **Habacha-Hamada, A.** Un système d'analyse contextuelle de schémas. In : *Proceedings du 9^e Congrès Reconnaissance de Formes et Intelligence Artificielle, Paris (France), 11-14 janvier, 1994*. Vol. 1, p. 647-653.
- [Hori93] **Hori, O. and Tanigawa, S.** Raster to vector conversion by line fitting based on contours and skeletons. In : *Proceedings of the 2nd International Conference on Document Analysis and Recognition, Tsukuba (Japon), october 20-22, 1993*. p. 353-358.
- [Joseph89] **Joseph, S.H.** Processing of engineering line drawings for automatic input to CAD. *Pattern Recognition*, 1989, Vol. 22, N° 1, p. 1-11.
- [Joseph91] **Joseph, S.H.** On the extraction of text connected to linework in document images. In : *Proceedings of the First International Conference on Document Analysis and Recognition, Saint-Malo (France), september 30-october 2, 1991*. p. 993-999.
- [Joseph92] **Joseph, S.H. and Pridmore, T.P.** Knowledge directed interpretation of mechanical engineering drawings. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992, Vol. 14, N°9, p. 928-940.
- [Kasturi90] **Kasturi, R.** A system for interpretation of line drawings. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, Vol. 12, N°10, p. 978-991.
- [Kasturi92] **Kasturi, R.** An Overview of techniques for graphics recognition. In : *Structured Document Analysis*. Baird, H.S., Bunke, H. and Yamamoto, K. Eds. Berlin (Allemagne) : Springer Verlag, 1992. P. 285-324.
- [Kuner86] **Kuner, P. and Ubbereiter, B.** Knowledge-based pattern recognition in disturbed line images using graph theory, optimization and predicate calculus. In : *Proceedings of the 8th International Conference on Pattern Recognition, Paris (France), 27-31 octobre, 1986*. Vol. 1, p. 240-243.

- [Lai91] **Lai, C.P. and Kasturi, R.** Detection of dashed lines in engineering drawings and maps. In : *Proceedings of the First International Conference on Document Analysis and Recognition, Saint-Malo (France), september 30-october 2, 1991*. Vol. 2, p. 507-515.
- [Lai93] **Lai, C.P. and Kasturi, R.** Detection of dimension sets in engineering drawings. In : *Proceedings of the 2nd International Conference on Document Analysis and Recognition, Tsukuba (Japon), october 20-22, 1993*. p. 606-613.
- [Lecas92] **Lecas, J.C.** *L'attention visuelle : de la conscience aux neurosciences*. Bruxelles (Belgique) : Mardaga Eds, 1992. 300 p.
- [Lee92] **Lee, S.W.** Recognizing hand-written electrical circuit symbols with attributed graph matching. In : *Structured Document Analysis*. Baird, H.S., Bunke, H. and Yamamoto, K. Eds. Berlin (Allemagne) : Springer Verlag, 1992. P. 340-358.
- [Lee93] **Lee, I., Dan, S., Kitahashi, T. and Abe, K.** A study of dividing machine-parts into functional groups for technical illustrations. In : *Proceedings of the 2nd International Conference on Document Analysis and Recognition, Tsukuba (Japon), october 20-22, 1993*. p. 886-889.
- [Lefèvre96] **Lefèvre, V., Pollet, Y., Philipp, S. et Brunesseaux, S.** Un système multi-agents pour la fusion de données en analyse d'images. *Traitement du Signal*, 1996, Vol. 13, N° 1, p. 99-111.
- [Leplumey93] **Leplumey, I., Camillerapp, J. and Lorette, G.** A robust detector for music staves. In : *Proceedings of the 2nd International Conference on Document Analysis and Recognition, Tsukuba (Japon), october 20-22, 1993*. p. 902-905.
- [Leplumey95] **Leplumey, I., Camillerapp, J. and Queguiner, C.** Kalman filter contributions towards document segmentation. In : *Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal (Canada), august 14-16, 1995*. Vol. 2, p. 765-769.
- [Lin85] **Lin, X., Shimotsuji, S., Minoh, M. and Sakai, T.** Efficient diagram understanding with characteristic pattern detection. *Computer Vision, Graphics and Image Processing*, 1985, Vol. 30, p. 84-106.
- [Lorette92] **Lorette, G. et Lecourtier, Y.** Reconnaissance et interprétation de textes manuscrits hors-ligne : un problème d'analyse de scène ? *Bigre*, 1992, N° 80, p. 109-135.
- [Luo95] **Luo, H., Agam, G., and Dinstein I.** Directional mathematical morphology approach for line thinning and extraction of character strings from maps and line drawings. In : *Proceedings of the 3rd International Conference on Document Analysis and Recognition. Montreal (Canada), august 14-16, 1995*. Vol. 1, p. 257-260.
- [Lysac91] **Lysac, B.D. and Kasturi, R.** Interpretation of engineering drawings of polyhedral and non-polyhedral objects. In : *Proceedings of the First*

International Conference on Document Analysis and Recognition, Saint-Malo (France), september 30-october 2, 1991. p. 70-78.

- [Madej91] **Madej, D.** An intelligent map to CAD conversion system. In : *Proceedings of the First International Conference on Document Analysis and Recognition, Saint-Malo (France), september 30-october 2, 1991. p. 603-610.*
- [Marr82] **Marr, D.** *Vision : A computational investigation into human representation and processing of visual information.* New-York (USA) : W.H. Freeman and company, 1982. 397 p.
- [Marston95] **Marston, R.E. and Kuo, M.H.** Interpretation of 3-view engineering drawings as central quadric surface mechanical parts. In : *Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal (Canada), august 14-16, 1995. Vol. 1, p. 343-346.*
- [Marti91] **Marti, E., Regincos, J., Lopez-Krahe, J. and Villanueva, J.J.** Interprétation de dessins faits à la main comme scènes tridimensionnelles pour l'entrée de données dans les systèmes CAO. In : *Proceedings of the First International Conference on Document Analysis and Recognition, Saint-Malo (France), september 30-october 2, 1991. p. 472-480.*
- [Miclet84] **Miclet, L.** *Méthodes structurelles pour la reconnaissance des formes.* Paris (France) : Eyrolles, 1984. 184 p.
- [Nagy95] **Nagy, G.** Document image analysis: what is missing ? In : *Proceedings of the 8th International Conference on Image Analysis and Processing, San-Remo (Italy), september, 1995. p. 577-587.*
- [Ogier93] **Ogier, J.M., Labiche, J. and Lecourtier, Y.** Attributes extraction for french map interpretation. In : *Proceedings of the 2nd International Conference on Document Analysis and Recognition, Tsukuba (Japon), october 20-22, 1993. p. 672-675.*
- [Ogier94a] **Ogier, J.M., Mullot, R., Labiche, J. et Lecourtier, Y.** Intégration d'outils bas niveau dans une stratégie d'interprétation de documents. In : *Proceedings du 9^e Congrès Reconnaissance de Formes et Intelligence Artificielle, Paris (France), 11-14 janvier, 1994. Vol. 1, p. 533-544.*
- [Ogier94b] **Ogier, J.M.** *Contribution à l'analyse automatique de documents cartographiques : Interprétation de données cadastrales.* Thèse de doctorat : Université de Rouen, 1994. 374 p.
- [Ogier95] **Ogier, J.M., Mullot, R., Labiche, J. et Lecourtier, Y.** Interprétation de documents par cycles perceptifs de construction d'objets cohérents. Application aux données cadastrales. *Traitement du Signal*, 1995, Vol. 12, N° 6, , p. 627-638.
- [Okazaki88] **Okazaki, A. and al.** An automatic circuit diagram reader with loop-structure-based symbol recognition. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 1988, Vol. 10, N° 3, p. 331-340.

- [Pao91] **Pao, D., Li, H.F. and Jayakumar, R.** Graphic features extraction for automatic conversion of engineering drawings. In : *Proceedings of the First International Conference on Document Analysis and Recognition, Saint-Malo (France), september 30-october 2, 1991*. Vol. 2, p. 533-541.
- [Paquet90] **Paquet, T., Vallee, T. et Lecourtier, Y.** Extraction de primitives par suivi de traits dans l'image binarisée d'un mot manuscrit. *Bigre*. 1990, N° 68, p. 179-188.
- [Paquet91] **Paquet, T. et Lecourtier, Y.** Reconnaissance de l'écriture manuscrite sur des chèques. In : *Proceedings du 8^e Congrès Reconnaissance des Formes et Intelligence Artificielle, Lyon-Villeurbanne (France), 25-29 novembre, 1991*. Vol. 2, p. 687-693.
- [Pavlidis82] **Pavlidis, T.** *Algorithms for graphic and image processing*. New York (USA) : Computer Sciences Press, 1982. 235 p.
- [Pavlidis86] **Pavlidis, T.** A vectorizer and feature extractor for document recognition. *Computer Vision, Graphics and Image Processing*, 1986, Vol. 35, p. 111-127.
- [Poulain94] **Poulain d'Andecy, V., Camillerapp, J. et Leplumey, I.** Analyse de partitions musicales. In : *Proceedings du 3^e Colloque National sur l'Ecrit et le Document, Rouen (France), 6-8 juillet, 1994*. p. 223-232.
- [Ramachandra80] **Ramachandran, K.** Coding method for vector representation of engineering drawings. In : *Proceedings of the IEEE*, 1980, Vol. 68, N° 7, p. 813-817.
- [Ramel93] **Ramel, J.Y.** *Lecture automatique de partitions musicales*. Rapport de DEA : Laboratoire RFV - INSA de Lyon (France), 1993. 65 p.
- [Ramel94] **Ramel, J.Y., Vincent, N. et Emptoz, H.** Reconnaissance de partitions musicales. In : *Proceedings du 3^e Colloque National sur l'Ecrit et le Document, Rouen (France), 6-8 juillet, 1994*. p. 325-334.
- [Ramel95] **Ramel, J.Y., Vincent, N. and Emptoz., H.** A method for the hierarchical extraction and description of lines and solid regions in technical drawings. In : *Proceedings of the 5th International Conference on Image Processing and its Applications, Edinburgh (Ecosse), july 4-6, 1995*. p. 379-383.
- [Ramel96a] **Ramel, J.Y., Vincent, N. et Emptoz, H.** Coopération entre vision globale et analyse locale pour une meilleure interprétation des documents techniques. In : *Proceedings du 4^e Colloque National sur l'Ecrit et le Document, Nantes (France), 3-5 juillet, 1996*. p. 117-124.
- [Ramel96b] **Ramel, J.Y., Vincent, N. et Emptoz, H.** Combining global and local vision for technical document understanding. In : *Proceedings the International Conference on Pattern Recognition, Viennas (Austria), 25-30 août, 1996*. Vol. C, p. 173-177.
- [Roosli95] **Roosli, M. and Monagan, G.** A high quality vectorisation combining local quality measures and global constraints. In : *Proceedings of the 3rd*

International Conference on Document Analysis and Recognition, Montreal (Canada), august 14-16, 1995. Vol. 1, p. 243-248.

- [Rosenfeld70] **Rosenfeld, A.** Connectivity in digital picture. *Journal of the Association for Computing Machinery*, 1970, Vol. 17, N°1, p. 146-160.
- [Rosenfeld86] **Rosenfeld, A.** Axial representation of shape. *Computer Vision, Graphics and Image Processing*, 1986, Vol. 93, p. 156-173.
- [Tanigawa94] **Tanigawa, S., Hori, O. and Shimotsuji, S.** Precise line detection from an engineering drawing using a figure fitting method based on contours and skeletons. In : *Proceedings of the 12th International Conference on Pattern Recognition, Jérusalem (Israël), 9-13 octobre, 1994.* Vol. 2, p. 356-360.
- [Tombre91] **Tombre, K. and Vaxivière, P.** Structure, syntax et semantics in technical document recognition. In : *Proceedings of the First International Conference on Document Analysis and Recognition, Saint-Malo (France), september 30-october 2, 1991.* Vol. 1, p. 61-69.
- [Tombre92] **Tombre, K.** Technical drawing Recognition and understanding : From Pixel to semantics. In : *Proceedings of the IAPR Workshop on Machine Vision and Application. Tokyo (Japon), 7-9 decembre, 1992.* p. 393-401.
- [Vaxivière92] **Vaxivière, P.** Structuration des niveaux de connaissance pour la conversion des plans mécaniques vers une CAO. *Bigre*, 1992, N°80, p. 75-82.
- [Vaxivière94a] **Vaxivière, P.** Une vectorisation par maillage du dessin et séparation en calques, adaptée aux plans mécaniques. In : *Proceedings of the 3^e Colloque National sur l'Ecrit et le Document, Rouen (France), 6-8 juillet 1994.* p. 205-214.
- [Vaxivière94b] **Vaxivière, P. and Tombre, K.** Knowledge organisation and interpretation process in engineering drawing interpretation. In : *Proceedings of the International Workshop on Document Analysis System, Kaiserslautern (Allemagne), 1994.* P. 313-321.
- [Vaxivière94c] **Vaxivière, P. and Tombre, K.** Subsampling : a structural approach to technical document vectorization. In : *Proceedings of the International Workshop on Syntactic and Structural Pattern Recognition, Nahariya (Israël), octobre 1994.*
- [Wall84] **Wall, K. and Danielsson, P.** A fast sequential method for polygonal approximation of digitized curves. *Computer Vision, Graphics and Image Processing*, 1984, Vol. 28, p. 220-227.
- [Wang92] **Wang, P. and Dori, D.** Line drawing, feature extraction, and symbol recognition. In : *Structured Document Analysis.* Baird, H.S., Bunke, H. and Yamamoto, K. Eds. Berlin (Allemagne) : Springer Verlag. 1992. p. 568-569.

- [Wang93] **Wang, P.** Machine visualization, understanding and interpretation of polyhedral line-drawings in document analysis. In : *Proceedings of the 2nd International Conference on Document Analysis and Recognition, Tsukuba (Japan), october 20-22, 1993.* p. 882-885.
- [Weiss95] **Weiss, M. and Dori, D.** A scheme for 3D object reconstruction from dimensioned orthographic views. In : *Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal (Canada), august 14-16, 1995.* Vol. 1, p. 335-338.
- [Ylä Jääski92] **Ylä Jääski, A. and Ade, F.** Line Segment Ribbons and their grouping. In : *Proceedings of the 11th International Conference on Pattern Recognition, The Hague (Netherlands), august 30-september 3, 1992.* Vol. 1, p. 750-754.

FOLIO ADMINISTRATIF

THESE SOUTENUE DEVANT L'INSTITUT NATIONAL DES SCIENCES APPLIQUEES DE LYON

NOM : RAMEL	DATE de SOUTENANCE 26 novembre 1996
Prénom : Jean-Yves	
TITRE : INTERPRETATION AUTOMATIQUE DE DESSINS METHODES D'ANALYSE ET DE RECONNAISSANCE APPLICATION AUX PLANS CINEMATIQUES	
NATURE : Doctorat	Numéro d'ordre : 96 ISAL 0105
Formation doctorale : DEA d'INFORMATIQUE de LYON	
Cote B.I.U. - Lyon : T 50/210/19	CLASSE :
/ et bis	
RESUME : L'objectif de l'interprétation automatique de documents est de faciliter l'exploitation des dessins en proposant un ensemble d'étapes permettant la transformation du schéma papier en un stockage numérique interprété. Il s'agit alors d'extraire l'information graphique et alphanumérique du document pour reconstituer la sémantique associée aux différents objets du dessin et à leurs relations. Dans le but de construire un tel système, nous avons choisi de nous inspirer du fonctionnement du système visuel humain. Durant une première phase de perception globale, une première représentation du document est obtenue à l'aide d'un nombre limité de primitives élémentaires : le Pixel, le Vecteur, et le Quadrilatère. Cette représentation fournira une description précise de toutes les formes du document initial et cela aussi longtemps que possible dans les différents niveaux d'analyse. Ensuite, l'utilisation duale d'un système multi-spécialistes et de la représentation préalablement construite permet, au cours d'une phase de lecture et d'interprétation, d'utiliser les données fournies par vision globale pour améliorer les analyses locales. Un mécanisme de cycles perceptifs permet aux différents spécialistes de coopérer pour extraire seulement les évidences et donc de limiter les erreurs d'analyse. Chaque spécialiste réalise une tâche précise telle que, par exemple, la détection du texte, des courbes, des éléments mécaniques ou la reconstruction 3D du mécanisme. Tous utilisent la vision globale pour focaliser leur attention sur des zones précises de l'image qu'ils étudient en détails pour faire avancer l'analyse. En fin de chaque cycle, la complétude de la compréhension du document est vérifiée et un nouveau cycle d'activation des spécialistes approfondit, si nécessaire, l'interprétation. Grâce à cette démarche, nos travaux se distinguent de ceux employant la chaîne de traitements habituelle : pré-traitements - segmentation - analyse - reconnaissance. Notre système s'efforce d'extraire aussi bien les formes que les connaissances, de manière hiérarchique et réfléchie, et les transcrit dans un modèle de représentation évolutif.	
MOTS-CLES : reconnaissance forme, interprétation image, documents, intelligence artificielle, perception, coopération.	
Laboratoire (s) de recherches : Laboratoire de Reconnaissance de Formes et Vision - INSA de LYON.	
Directeur de Thèse : N. Vincent	