

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et
de la Recherche Scientifique
Université Saâd Dahlab de Blida

FACULTE DE SCIENCES
DEPARTEMENT D'INFORMATIQUE

Mémoire de projet de fin d'études
Pour l'obtention du diplôme
D'ingénieur d'état en génie informatique

Spécialité : Informatique
Option : Système Information

Présenté par :
M^r BOUNOUNI Abderahmane
M^{elle} MERABET Nadjia

Thème :

**OPTIMISATION DE LA BASE DES ATTRIBUTS
QUANTITATIFS DES IMAGES DES DEFAUTS DE
SOUDURE EN RADIOGRAPHIE INDUSTRIELLE
PAR L'ANALYSE DE DONNEES**

Proposé par :
M^r NACEREDDINE NAFAA

Setenu publiquement le : 09/12/2004
devant le jury composé de :
Mr Dj.BOUNOUAR
Mr M.BOUGHARA
Melle N.BOUSTIA

Président
Examinateur
Examinatrice

MIG-004-38-1

Promotion 2003-2004



Dédicace



Ce mémoire est dédié :

A la mémoire de mon père que dieu le pardonne et l'accueille en son Vaste Paradis, qui malgré son absence son souvenir et son amour reste gravé pour toujours dans mon cœur

A ma très cher Mère pour son soutien durant toute ma carrière, qui a été toujours pour moi une source d'affection, d'amour l'aide de Dieu et puis sans elle je ne serais jamais devenu ce que je suis aujourd'hui.

Une dédicace tout spéciale pour mon fiancé Rédha qui n'hésitait pas de se priver du nécessaire pour me venir en aide et me combler de bonheur, qui ne cessait jamais de m'offrir le soutien moral.

A mes frères kadar, Boutouchent, Hamid Abderezak et mes sœurs Zoulikha, Saâda, Fatiha, Amina, Ghania, Zakia, Rym.

A toute la famille BENSMAIL ainsi que Houria, Mohamed, Hayat

A mon binôme Abderahmane et à toute sa famille

A mes amis Akila Nadjet, Naima, Sara, Salima, Souhila, Ichrak,

Amel, Nihed, Ghada, Rachida, Chahra, Byba, Darine, Souad

O.Ibrahim, Ahmed, Hamza, M.Ibrahim.

A tous mes copines à la cité Universitaire Benboulaid que j'aime et qui m'aiment.

M. Nadjia

Dédicace

*Je dédie ce mémoire avec toute ma gratitude
A ma grand-mère Fatma A ceux qui ne cessent de m'encourager
durant mes longues années d'études et qui étaient toujours
présent dans les moments difficiles,
A mes parents pour leur soutien, leurs efforts durant
toute ma carrière
A ma femme Malika et ma petite princesse Zahra
A mes frères Hamid, Ali, Hamza, Chamessedine,
A mes sœurs Farida, Leila, Ferial, Kenza sans oublier
mes beaux frères Abdelkader, Mourad, Raouf, Smail,
Hamid et sa femme Bahia
A mes belles sœurs Nacéra, Assia.
A mes adorables petits neveux et nièces Zahra, Manal,
Zaki, Zineb, Meriém, Baya, Kamilia, Rania,
Achraf et Akram
A mon binôme Nadja et à toute sa famille ainsi son fiancé
Rédha A tous mes amis Benyoucef, Lyes, Fodil, Mourad,
Brahim, Farid, Salima, Hocine, Sid Ali et Mohamed.
Et enfin à mes beaux frères Farid et Zahra.
A tous mes amis que j'aime et qui m'aiment.*

B. Abderahmane

Remerciements

Nous remercions avant tout le bon dieu qui nous a aidé à réaliser ce modeste travail.

Nous tenons à remercier M^r. NAFAA Nacereddine et qui a bien voulu proposer le sujet pour son aide, sa patience, sa disponibilité et sa compréhensibilité.

Nous remercions M^{me} le chef du département d'Informatique, tous les enseignants de la faculté des sciences de BLIDA et surtout ceux du département informatique.

Nous remercions les membres du jury pour nous avoir fait l'honneur de juger notre travail.

Les plus sincères remerciements sont destinés au laboratoire CSC pour leur impeccable accueil

Un remerciement tout spécial à M^{me} MOKHTARI pour sa disponibilité et sa direction

Nous remercions, de tout coeur, tous ceux qui ont contribué de près ou de loin à la réalisation de ce travail.

Et enfin, Nous tenons à remercier tous nos amis, B.Redha, ZREF Nadjet, Hassaine Salima, Mezaourou Ibrahim pour l'aide qu'il nous a apportée pour l'édition de la thèse.

B.Abderahmane et M.Nadjia

Liste des Equations:

Equation III.1	33
Equation III.2	34
Equation III.3	35
Equation III.4	35
Equation III.5	35
Equation III.6	36
Equation III.7	36
Equation III.8	37
Equation III.9	37
Equation III.10	37
Equation III.11	38
Equation III.12	38
Equation III.13	39
Equation III.14	39
Equation III.15	39
Equation III.16	39
Equation III.17	39
Equation III.19	39
Equation III.20	40
Equation III.21	59
Equation V.1	60
Equation V.2	65
Equation V.3	66
Equation V.4	66
Equation V.5	66
Equation V.6	66
Equation V.7	66
Equation V.8	67
Equation V.9	67
Equation V.10	67

Liste des figures:

Figure I.1	Schéma représentatif d'un joint soudé	10
Figure I.2	Les différents types des défauts de soudage	13
Figure I.3	Montage de base pour la radiographie sur film	16
Figure II.1	les types de connexité	22
Figure II.2	Description générale d'une chaîne de vision par ordinateur	25
Figure II.3	Principe de recadrage dynamique	35
Figure II.4	Schéma de principe du processus de reconnaissance automatique de défaut de soudure	27
Figure III.1	Extraction de contours	33
Figure III.2	Illustration de la notion de la connexité	34
Figure III.3	Orientation de l'objet	36
Figure III.4	Longueur et largeur du rectangle exinscrit	36
Figure III.5	Cercle maximal inscrit	37
Figure III.7	Illustration de la notion des surfaces partielles	38
Figure IV.1	Représentation des individus dans le plan principal	51
Figure IV.2	Représentation dans l'espace des caractères	57
Figure V.1	Modèle d'un réseau de neurone, perceptron	62
Figure V.2	Réseau de neurones multicouches	63
Figure V.3	Configuration du RNA utilisé	69
Figure VI.1	Diagramme de haut niveau du système	72
Figure VI.2	Cas d'utilisation traitement sur la matrice initiale	73
Figure VI.3	Cas d'utilisation réduction la matrice par l'ACP	74
Figure VI.4	Cas d'utilisation classification des défauts par réseaux de neurones	75
Figure VI.5	Diagramme de collaboration	76
Figure VI.6	Diagramme de classe globale	78
Figure VI.7	Paquetage de diagramme des classes : Mesures des paramètres	79
Figure VI.8	Paquetage de diagramme des classes : ACP	81
Figure VI.10	Diagramme d'activité pour le processus extraction des primitives	83
Figure VI.11	Diagramme d'activité pour le processus ACP	85
Figure VI.12	Diagramme d'activité de la méthode Jacobi	87
Figure VII.1	Présentation du Menu Principale	92
Figure VII.2	Présentation quelques Images de contour associés aux défauts de soudures	93
Figure VII.3	Illustration des paramètres géométriques aux images des défauts utilisées	94
Figure VII.4	Illustration graphique des 1 ers Axes principaux aux images des défauts utilisés	95
Figure VII.5	Illustration graphique du Rectangle exinscrit aux images des défauts utilisés	96
Figure VII.6	Illustration graphique du cercle inscrit aux images des défauts utilisés	96
Figure VII.7	La matrice initiale	97
Figure VII.8	A propos de l'application	98

TABLE DES MATIERES

	Résumé	1
	Abstract	1
INTRODUCTION		3
GENERALE		
	I. Objectif	5
	II. Cahier de charge	7
CHAPITRE I	SOUDAGE ET RADIOGRAPHIE	8
	INDUSTRIELLE	
	I. Introduction	9
	II. Définition du soudage	9
	II.1 Description d'un joint soudé	10
	II.2 Source et types des défauts de soudage	10
	III. Aperçu sur les méthodes de contrôle non destructif	13
	IV. Principe de la radiographie	15
	IV.1 Qualité du film radiographique	16
	IV.2 Le contraste d'image	17
	IV.3 La netteté de l'image	17
	IV.4 La granulation	17
CHAPITRE II	INTRODUCTION AU TRAITEMENT	19
	D'IMAGES	
	I.1 Introduction	20
	II. Généralité	20
	III. Prétraitement	23
	IV. Segmentation	26
CHAPITRE III	Analyse Quantitative des Images des défauts de soudure	31
	I. Introduction	32
	II. Extraction de contours	33
	III. Détermination des paramètres	33
	III.1 Mesure des paramètres géométriques	33
	IV. Relation entre les attributs invariants proposés et les types des défauts de soudures	40
	V. Conclusion	41
CHAPITRE IV	Analyse en Composante Principale (ACP)	42
	I. Généralités	42
	II. Choix d'une métrique en ACP	44
	III. Analyse dans l'espace des individus R_p	46
	III-1. Coordonnées centrées	46
	III-2. Coordonnées centrées réduites	47
	III.3. Axes principaux d'inertie	48
	IV. Analyse dans l'espace des caractères R^n	53

CHAPITRE V	Les Réseaux de Neurones et la Classification des défauts de soudures	59
	I. Introduction	60
	II. La Présentation des réseaux de neurones	60
	II.1 Perceptron multicouches	63
	III. Travaux Futur	67
CHAPITRE VI	La Conception et la Réalisation	70
	I La Vue Cas d'utilisation (Diagramme de cas d'utilisation)	71
	I.1 Définition des acteurs	72
	I.2 Cas d'utilisation de haut niveau	72
	I.3 Cas d'utilisation de bas niveau	72
	I.4 Diagramme de collaboration	76
	II La Vue Logique (Diagramme de classes)	77
	II.1 Diagramme de classe	78
	III La Vue Dynamique (Diagramme d'activités)	83
CHAPITRE VII	Implémentation et Test	90
	I. Introduction	91
	II .Choix de langage de programmation	91
	III. Test	91
	III.1 Menu Principale	91
Conclusion Générale		100
ANNEXE A	UML et la Méthode de Conception	101
ANNEXE B	Illustration graphique de quelques paramètres géométriques	112

Résumé :

La radiographie s'avère être une technique de plus en plus prometteuse dans le contrôle non destructif. La radiation transmise détecte sur un film radiographique. Ce dernier est interprété pour détecter des discontinuités dans la soudure. On rencontre divers types de défauts de soudure tels que les porosités, les inclusions, les fissures, le manque de pénétration, le manque de fusion etc. La tâche d'interprétation est assurée par des interpréteurs en radiographie, elle est soumise au jugement humain telles que l'expérience de l'interpréteur, rendant l'interprétation de la qualité de la soudure très subjective et parfois inconsistante. Il est ainsi souhaitable de développer des techniques par ordinateur pour assister l'interpréteur à évaluer la qualité des joints soudés.

Ce travail est réalisé dans le but est la classification passant par l'analyse quantitative des images de défaut basée sur la recherche d'attributs invariants (Compacité, Elongation, Rectangularité... etc) qui découlent du calcul des paramètres géométriques (surface, périmètre, etc.). Les valeurs de ces attributs et vu d'autres considérations imposées par l'analyse en composante principale (ACP) dans le but de réduire la base d'attributs. Par la suite cette nouvelle base d'attributs sera l'entrée de réseau de neurones artificiel pour la classification des défauts de soudure.

Abstract

There are various types of weld defects such as porosity, inclusion, crack, lack of penetration, lack of fusion, etc. The task of interpretation is ensured by interpreters in radiography. Consequently, it is submitted to the human judgment such as the aptitude and the experience of the interpreter, making the weld quality interpretation subjective and sometimes biased. It is thus desirable to develop the computer-aided techniques to assist the human interpreter in evaluating the quality of the welded joints.

The aim of this work is to develop a classification which consists in the quantitative image analysis based on the features invariant calculation and which result from the calculation of the geometrical parameters (surface, perimeter, etc), in values of the raw features and taking into account other considerations imposed by ACP. Thereafter, an artificial neural network for the weld defect classification task was used.



INTRODUCTION GENERALE

La qualité des produits, des composants ou des pièces dépend de plusieurs facteurs qui sont la conception, les caractéristiques des matériaux et les techniques de fabrication. La qualité peut être définie en termes des défauts et imperfections présents dans les matières employées pour la conception du produit ou en termes de la présence de tels défauts et imperfections dans le produit fini lui-même. Beaucoup de défauts peuvent également être formés dans les produits en fonctionnement. La nature de ces défauts diffère selon les processus de sa conception et sa fabrication aussi bien que les conditions de service dans lesquels il doit fonctionner. La connaissance de ces défauts dans un produit en vue de les déterminer et puis, de les réduire au minimum est essentielle pour réaliser un meilleur niveau de qualité.

Le soudage est la façon la plus appropriée d'assembler deux ou plusieurs pièces en métal. Le soudage est utilisé pour fabriquer et réparer des produits métalliques. La soudure est obtenue en chauffant le métal à la température de soudage avec ou sans application de pression, et avec ou sans métal d'apport. Des soudures sont largement utilisées dans beaucoup de secteurs tels que la pétrochimie, l'aéronaval, la construction automobile, les réacteurs d'énergie nucléaire et l'électronique.

Le contrôle non destructif joue un rôle important dans le contrôle de qualité non seulement du produit fini mais également des produits semi finis aussi bien des matières premières initiales. Le CND peut être employé à toutes les étapes de la production. Dans l'industrie, la radiographie qui est l'une des méthodes du contrôle non destructif, peut être utilisée pour révéler la structure interne des matériaux, spécialement des assemblages soudés. Dans cette technique, on utilise des radiations pénétrantes : rayons X et rayons gamma. Les variations d'épaisseur et de densité modifient le passage des rayons à travers l'échantillon. L'intensité de la radiation transmise peut être détectée sur un film radiographique. Ce film est interprété pour détecter d'éventuelles imperfections dans la soudure. Il y a divers types de défauts de soudure tels que les porosités, les inclusions, les fissures, le manque de pénétration, le manque de fusion etc. Ceux-ci peuvent être classés initialement en défauts plans ou volumiques avant leur identification finale. Cette tâche est exécutée par des interpréteurs en radiographie. Donc, elle est soumise au jugement humain et à des considérations subjectives, telles que, l'aptitude et l'expérience de l'interpréteur. En plus, l'interprétation humaine de la qualité de la soudure est très subjective, fastidieuse, et parfois inconsistante. Il est ainsi souhaitable de développer certaines techniques par ordinateur pour assister l'interpréteur à évaluer la qualité des joints soudés.

Le traitement numérique des images recouvre l'ensemble des procédés d'extraction d'informations qualitative des images numériques, allant du prétraitement jusqu'à l'extraction de la zone de l'objet à identifier. On assiste aujourd'hui à l'utilisation d'algorithmes de traitement d'image qui sont de plus en plus performants et spécifiques. Dans l'analyse des images, l'intérêt primordial que portent les attributs géométriques invariants (Compacité, Elongation, Rectangularité...etc) qui découlent du calcul des paramètres géométriques (surface, périmètre, etc.). Les valeurs de ces attributs et vu d'autres considérations imposées par l'analyse en composante principale (ACP) dans le but de réduire la base d'attributs cette nouvelle base d'attributs sera l'entrée de réseau de neurone est qu'on arrive à identifier dans un premier lieu les différents types des défauts de soudure contenus dans les radiogrammes.

Notre travail s'articule sur Sept chapitres. Dans le premier chapitre, nous donnons un aperçu sur le soudage, les défauts de soudure et le contrôle non destructif par la radiographie industrielle. Dans le deuxième chapitre, notamment, le traitement numérique des images, nous présentons le schéma général du système de vision par ordinateur dédié aux images des films radiographiques, allant de la numérisation des images des films radiographiques jusqu'à la segmentation, passant par le filtrage du bruit et l'amélioration du contraste. Où l'objectif est d'extraire une seule région de l'image représentant le défaut de soudure. Le troisième chapitre est consacré à l'analyse quantitative de l'image défaut par le calcul des attributs géométriques invariants. Dans le quatrième chapitre, nous étudierons l'analyse en composantes principales (ACP) pour éliminer la redondante entre les différentes variables d'attributs, dans le Cinquième chapitre nous étudions les réseaux de neurones artificiels pour la classification des défauts de soudure. La description détaillée de la méthode d'apprentissage par la rétropropagation du gradient est aussi illustrée. Dans le Sixième chapitre nous avons présenté la conception et la réalisation. Enfin, dans le Septième chapitre nous avons présenté implémentation et test c'est la présentation des résultats de notre logiciel.

I. Objectifs :

L'étude abordée dans ce projet se rapporte à la modélisation et la réalisation d'un système aide à la décision pour l'identification des défauts de soudures.

Un système aide à la décision permet d'assister un opérateur humain, dans la prise de décision dans un domaine bien précis. L'opération d'expertise est parfois très délicate, telle que l'expertise des images de films de radiographie.

Nous avons étudié le problème de segmentation dans un cadre global, en essayant d'aborder les différents aspects.

Une image n'est qu'une représentation d'un univers composé d'entités, issues d'un mélange de phénomènes physiques et géométriques divers. Le but de notre étude est d'extraire les attributs qui caractérisent ces entités; les attributs étudiés correspondent à des mesures de paramètres géométriques d'où à des points d'intérêt on à des zones caractéristiques de l'image.

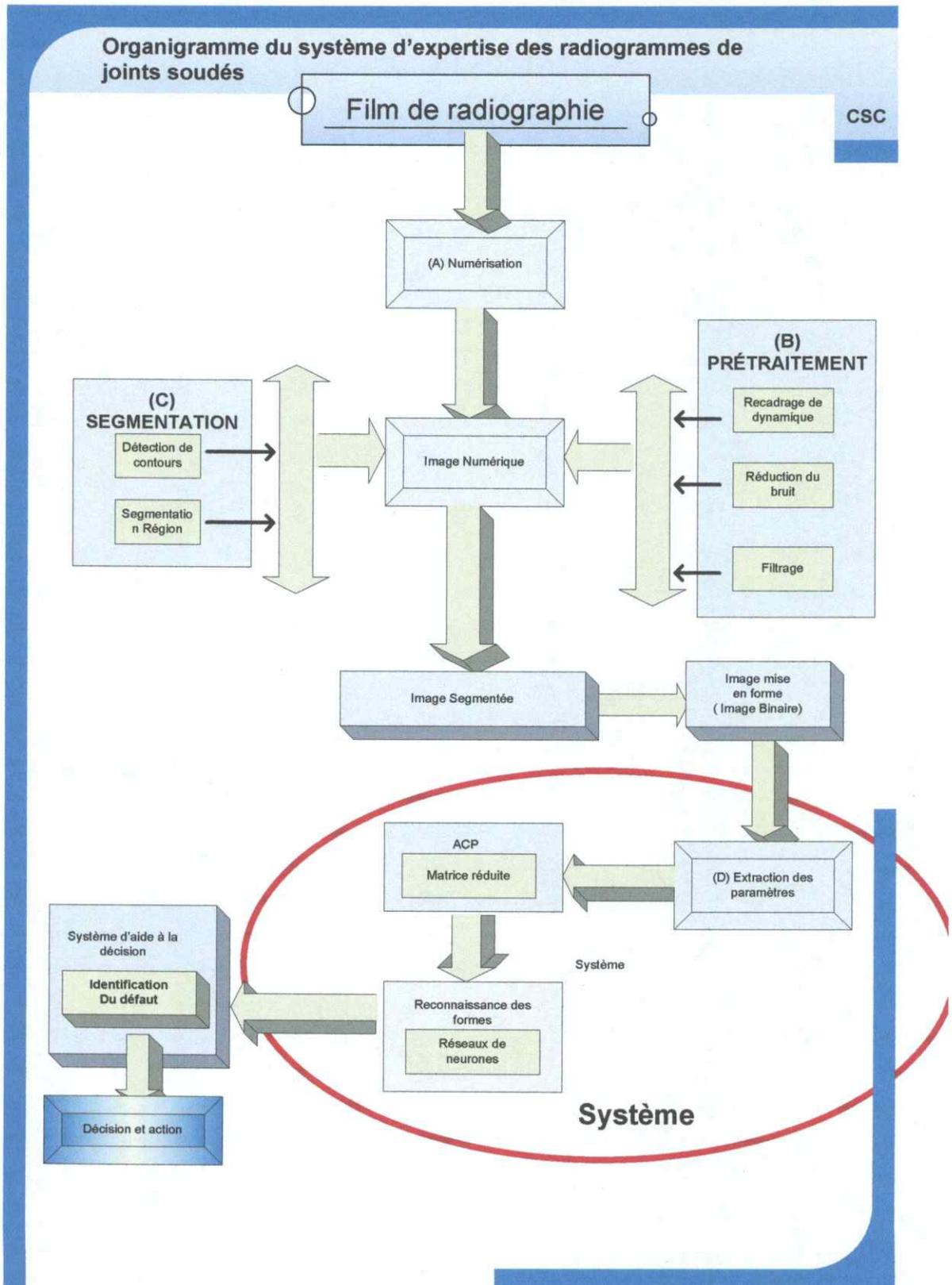
Un défaut de soudure va se présenter dans l'image binaire comme une région, possédant des caractéristiques géométriques, similaires à des modèles de référence, ces modèles de référence sont stockés dans une base de connaissance, regroupant le plus grand nombre de formes possibles.

Soit une région (connexe) de l'image, dont les pixels ont pour coordonnées (X_i, Y_i) ; on peut disposer des attributs suivants: (nous citons que quelques uns) la surface, périmètre, symétrie, centre gravité, ...etc.

Nous avons appliqué l'ACP sur la matrice initiale n individus représentent les images des défauts et p caractères qui représentent les attributs calculés dans la partie analyse quantitative.

Pour l'identification des défauts, nous avons choisi la méthode de classification "réseaux de neurones".

Le schéma suivant montre l'organigramme de notre système:



II. Cahier de charge :

Le système à réaliser sera conçu selon une approche objet. Le langage de spécification des divers aspects du système sera UML. Pour arriver à ce but, les travaux suivants seraient nécessaires :

- 1- Etude sur les généralités du traitement d'image (la prétraitement, la segmentation), analyse quantitative d'image par le calcul d'attribut géométriques, la méthode de réduction de variable en analyse des données ACP (Analyse en composante principale), les notions géométriques.
- 2- Etude sur les méthodes de reconnaissance des formes, choix d'une méthode pour la classification.
- 3- Etude poussée D'UML et apprentissage des outils (un IDE) basé sur UML.
- 4- Etude d'une ou deux méthodologies de développement de système en UML.
- 5- Choix des cibles d'implantation pour chaque alternatives : Plate-forme, environnement d'exécution ; langage de programmation.
- 6- Modélisation UML de chaque alternative: Quelles sont les vues à utiliser pour définir les aspects nécessaires du système ?
- 7- Implémentation effective du système.



CHAPITRE I

SOUDAGE ET RADIOGRAPHIE

INDUSTRIELLE

I Introduction :

La bonne exécution d'un produit manufacturé peut rarement être garantie sans inspection ou contrôle de qualité. A cette fin, on peut soit prélever des échantillons pour les soumettre à des essais, soit mettre au point une méthode globale d'examen non destructif. La nécessité de disposer d'un système complet de contrôle non destructif s'impose d'autant plus lorsqu'il s'agit d'objets fabriqués tels que des assemblages soudés, qui nécessitent un travail minutieux et un grand savoir-faire.

La plupart des examens non destructifs visent à détecter les défauts éventuels pour déterminer, le cas échéant, s'ils sont susceptibles d'influer sur les conditions de fonctionnement ou sur la vie utile du produit. Toutefois, les examens non destructifs permettent également de vérifier des assemblages et de mettre au point des processus de fabrication. De plus, ils se révèlent utiles dans la recherche.

Si les méthodes d'examen non destructif sont nombreuses, seules quelques-unes permettent d'examiner des objets de part en part. D'autres ne révèlent que les défauts de surface.

L'une des méthodes les plus communément utilisées, est la radiographie

Les clichés d'objets obtenus par rayons X ou gamma servent à mettre en évidence d'éventuelles modifications d'épaisseur (défauts internes ou de surface), des modifications de structure, des détails d'assemblage, etc. En matière de radiographie industrielle, l'éventail des techniques et équipements est très large, de sorte qu'il est nécessaire de posséder une certaine expérience et un certain savoir-faire pour choisir la technique la plus appropriée à une application déterminée sans perdre de vue les facteurs de performance, rapidité, coût, sensibilité, etc. Cette introduction à la radiographie industrielle signale les facteurs à prendre en considération lors de la mise au point et de l'utilisation de cette méthode d'examen non destructif

La radiographie industrielle n'est pas nécessairement coûteuse, et certaines nouvelles techniques rendent cette méthode d'inspection très rapide.

II. Définition du soudage :

Le soudage est un procédé qui assure la continuité métallique entre deux éléments à assembler, contrairement aux autres techniques d'assemblage (boulonnage, rivetage, etc.).

Le soudage permet d'assembler la quasi-totalité des métaux, dans pratiquement toutes les positions et dans tous les milieux, tout en assurant une bonne continuité de la matière, une bonne transition des efforts, une bonne résistance et une parfaite étanchéité.

Parmi de nombreuses définitions données dans la bibliographie, la plus indiquée pour la métallurgie de soudage (donnée par l'Institut International de Soudage) est celle qui fait appel à la notion de continuité métallique entre les parties à assembler. Un examen attentif de ce qui se passe à l'échelle de la structure cristalline nous éclairera mieux sur la nature de la continuité métallique qu'établit le soudage. Les métaux et alliages s'y prêtent particulièrement bien, en raison de leur pouvoir de diffusion élevée. L'essor industriel du soudage a été rendu possible grâce à la mise au point de puissantes sources d'énergie thermique. On distingue le soudage exécuté sans effort mécanique extérieur sur les éléments à assembler, qui exige toujours une élévation de la température jusqu'à fusion des bords de ces éléments et nécessite souvent l'introduction dans le joint d'un apport extérieur de métal ; le soudage exécuté avec effort, où la liaison peut être obtenue soit par fusion des éléments à assembler, soit sur les bords demeurant à l'état solide mais portés à la température de plasticité du matériau.

II.1 Description d'un joint soudé:

Le joint soudé est la zone critique de la soudure, il détermine si une soudure peut être acceptable. Il est formé de trois régions (Fig. I.1):

- a. Le métal de base : c'est une région qui ne subit aucune action thermique.
- b. La zone affectée thermiquement (ZAT) : c'est une région qui subit des transformations lors du soudage.
- c. La zone fondue : cette région est formée par la fusion du métal d'apport.

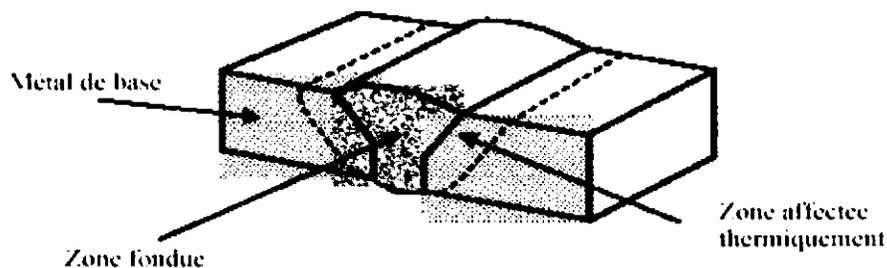


Figure I.1 Schéma représentatif d'un joint soudé.

II.2 Source et types des défauts de soudage

Les défauts de soudage sont en général inhérents aux procédés de soudage. En fait, ils dépendent de plusieurs facteurs : mauvaises exécutions (surcharge, pression, durée d'application, etc.), qualité du métal de base et/ou du produit d'apport, propreté des régions à souder et choix du procédé lui-même [1].

On donne ici une brève description et les causes probables des principaux défauts de soudage:

Fissure : Rupture dans le métal qui résulte de la présence des contraintes localisées qui dépassent la résistance limite du matériau.

Soufflure : Cavité gazeuse emprisonnée durant le processus de solidification, due essentiellement à la présence de l'humidité dans les pièces et les électrodes.

Retassure : Cavité due au retrait du métal au moment de la solidification.

Inclusion : Corps solide étranger emprisonné dans la masse du métal fondu dû à un nettoyage insuffisant des cordons ou à un piquage mal fait du laitier dans le cas du soudage multipasse.

Manque de fusion : Manque de liaison dans le cordon de soudure ou bien entre celui-ci et le métal de base dû à une intensité du courant de soudage trop faible, une avance trop rapide ou des cordons trop bombés formant entre eux un angle aigu dans le cas du soudage multipasse.

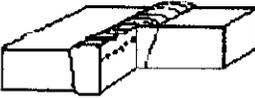
Manque de pénétration : Remplissage incomplet du fond de la rainure de soudage avec le métal de la soudure, dû à l'utilisation d'électrodes de diamètres trop grand, à l'emploi d'un courant d'intensité trop faible ou d'une mauvaise préparation des chanfreins.

Défauts de forme : Cordon trop bombé, dû à une intensité du courant de soudage trop faible.

Cordon trop aplati, dû à une intensité du courant de soudage trop forte. Cordon irrégulier, dû à une avance irrégulière, un mauvais choix de l'intensité de soudage, une mauvaise longueur d'arc ou une utilisation d'électrodes non appropriées au métal de base.

Les défauts de soudage peuvent prendre différentes formes [2], plusieurs types de défauts sont répertoriés dans la figure ci-après :



GROUPE N° 1 - FISSURES			
			
FISSURE LONGITUDINALE	FISSURE TRANSVERSALE	FISSURES RAYONNANTES	FISSURES DE CRATÈRE
GROUPE N° 2 - CAVITÉS OU POROSITÉS			
			
SOUFFLURE SPHEROÏDALE	SOUFFLURES ALLIGNÉES	NID DE SOUFFLURES	SOUFFLURES ALLONGÉES
			
SOUFFLURE VERMICULAIRE	PIQÛRE	RETASSURE INTERDENTRITIQUE	RETASSURE DE CRATÈRE
GROUPE N° 3 - INCLUSIONS SOLIDES			
			
INCLUSION DE LAITIER	INCLUSION DE FLUX	INCLUSION D'OXYDE	INCLUSION MÉTALLIQUE
GROUPE N° 4 - MANQUE DE FUSION			
			
MANQUE DE FUSION DES BORDS OU COLLAGE		MANQUE DE FUSION ENTRE PASSES	
GROUPE N° 5 - MANQUE DE PÉNÉTRATION			
			
MANQUE DE PÉNÉTRATION		MANQUE D'INTERPÉNÉTRATION	

GROUPE N° 6 – DÉFAUTS DE FORME			
 CANIVEAU	 MORSURE	 CANIVEAU A LA RACINE	 SUREPAISSEUR EXCESSIVE
 EXCÈS DE PÉNÉTRATION	 GOUTTE OU EXCÈS LOCAL DE PÉNÉTRATION	 DEBORDEMENT	 DÉFAUT D'ALIGNEMENT
 DÉFORMATION ANGULAIRE	 EFFONDREMENT	 MANQUE D'ÉPAISSEUR	 RETASSURE A LA RACINE

Figure I.2 Les différents types des défauts de soudage

III Aperçu sur les méthodes de contrôle non destructif : [3]

Le contrôle non destructif est devenu une pièce maîtresse dans chaque boîte à outils industrielle. Construire des usines ou des constructions soudées sans le contrôle non destructif aujourd'hui serait comme construire sans dimensionner, nettoyer ou souder. Faire la maintenance des avions, des raffineries ou des machines tournantes sans CND serait comme faire la maintenance sans la lubrification ou sans la vérification de l'étanchéité et de la corrosion.

Le CND a pour but de détecter le ou les défauts présents dans la pièce. Il est très important ensuite de les identifier et de les dimensionner. Par exemple, la taille d'une fissure détectée dans un matériau peut être incluse dans le calcul de la mécanique de la rupture pour estimer la durée de vie de ce matériau fissuré et par conséquent, la durée de vie de l'installation industrielle.

Il y a trois principaux paramètres à considérer pour le choix d'une méthode de contrôle :

- a. les avantages et les limitations de la méthode d'inspection,
- b. les normes d'acceptation,
- c. le coût.

Les méthodes de contrôle non destructif utilisées couramment sont :

- ❖ **Le contrôle visuel**, qui ne doit pas être oublié, est généralement pratiqué par le soudeur lui-même : linéarité et régularité du cordon de soudure, absence de surépaisseur ou de criques, de caniveaux (manque de fusion du métal sur les bords de la soudure), bonne pénétration du métal d'apport, etc.
- ❖ **Le contrôle par ressuage** utilisable pour la recherche des défauts débouchant en surface de tous les métaux non poreux. Après nettoyage et dégraissage de la surface à contrôler, un produit est appliqué au pinceau, par pulvérisation ou immersion et va s'infiltrer dans les criques. Ce produit contient soit un traceur rouge visible en lumière naturelle, soit un traceur fluorescent pour examen en lumière ultraviolette.
- ❖ **Le contrôle par magnétoscopie** qui permet de déceler des défauts affleurant la surface mais il est uniquement utilisable pour des matériaux ferromagnétiques. Sous l'action d'un champ magnétique, les particules magnétiques en suspension dans un liquide s'accumulent à l'endroit des défauts.
- ❖ **Le contrôle par ultrasons** qui utilise des ondes élastiques à fréquence élevée, de l'ordre du Mégahertz, qui se propagent dans le milieu, grâce à la mise en vibration élastique, autour de leur position d'équilibre, des points matériels constituant le matériau. Dans l'examen des soudures, on fait généralement appel aux ondes transversales qui vibrent perpendiculairement au sens de leur propagation. Le contrôle des soudures peut se faire, soit par transmission (deux palpeurs séparés : un émetteur et un récepteur), soit par réflexion ou par échos (un seul palpeur émetteur et récepteur).
- ❖ **Le contrôle d'étanchéité** qui permet de déceler les fuites dues à des défauts de dimensions microscopiques par mise sous pression ou en dépression du conteneur à contrôler au moyen d'un gaz (hexafluorure de soufre, hélium, ammoniac, etc.) ou d'un liquide.
- ❖ **Le contrôle par courant de Foucault** est largement utilisé pour détecter les défauts de surface. Cette méthode s'applique aux matériaux conducteurs d'électricité uniquement. Lorsqu'une bobine parcourue par un courant alternatif est approchée à la pièce, des courants de Foucault sont induits dans celle-ci par induction électromagnétique. L'amplitude des courants de Foucault induits dépend, entre autres, de la présence des discontinuités dans la pièce.
- ❖ **Le contrôle par radiographie** s'effectue principalement avec les rayons X ou les rayons gamma. Ces rayonnements ont la propriété de pénétrer la matière, d'être plus ou moins atténués selon la densité et l'épaisseur des matériaux traversés et de

transporter une "image radiante" de ces derniers, qui peut être convertie en image visible sur un dispositif approprié : un film radiographique (radiographie industrielle) ou un écran fluorescent (radioscopie industrielle).

La section suivante sera consacrée à cette dernière méthode, car l'objet de notre travail se consacrera à l'analyse des images des films de la radiographie et ce, dans le but de la détection et la classification des défauts de soudures.

IV Principe de la radiographie : [4]

En radiographie industrielle, la procédure généralement utilisée pour obtenir un cliché radiographique consiste à placer une source de radiations pénétrantes (rayons X ou gamma) d'un côté de l'objet à examiner, et à poser un détecteur de rayonnement (le film) de l'autre Figure I.3. L'énergie des radiations doit être déterminée de manière à ce que la quantité de rayonnement soit suffisante pour faire réagir le détecteur de manière adéquate.

Le détecteur est généralement un film photographique placé dans une enveloppe opaque ou un chargeur dont la surface frontale est suffisamment mince pour laisser passer les rayons X ou gamma.

Les dimensions de la source de rayonnement sont aussi réduites que possible (quelques millimètres de diamètre) comme les rayonnements ionisants qui se propagent en ligne droite de la source jusqu'au film, ils déterminent une image nette des défauts éventuels, des arêtes ou des discontinuités présentes dans l'objet. Cette image géométrique se forme selon le même principe que l'ombre créée par une source de lumière visible ; de même, la netteté de l'image dépend du diamètre de la source de rayonnement et de sa distance par rapport à la surface sur laquelle se forme l'image.

Normalement, le film contenu dans la cassette est placé tous près de l'objet à examiner. Le générateur de rayons X ou gamma est activé pendant une période de temps adéquate (la durée d'exposition), puis le film est soumis à un traitement photographique: développement, fixage, lavage et séchage. Le cliché obtenu est placé ensuite sur un écran éclairé par le négatoscope.

Les différences d'intensité du faisceau de rayons X ou gamma traversant l'objet entraînent des différences de noircissement (densité photographique du film), qui se traduisent sur le film éclairé par des différences de luminosité.

La radiographie industrielle est principalement concernée par les images d'enregistrement sur le film. Voici donc, les trois bases essentielles dans la production d'une radiographie :

- a. une source de rayonnement des rayons X ou gamma
- b. l'objet à examiner
- c. une cassette contenant le film

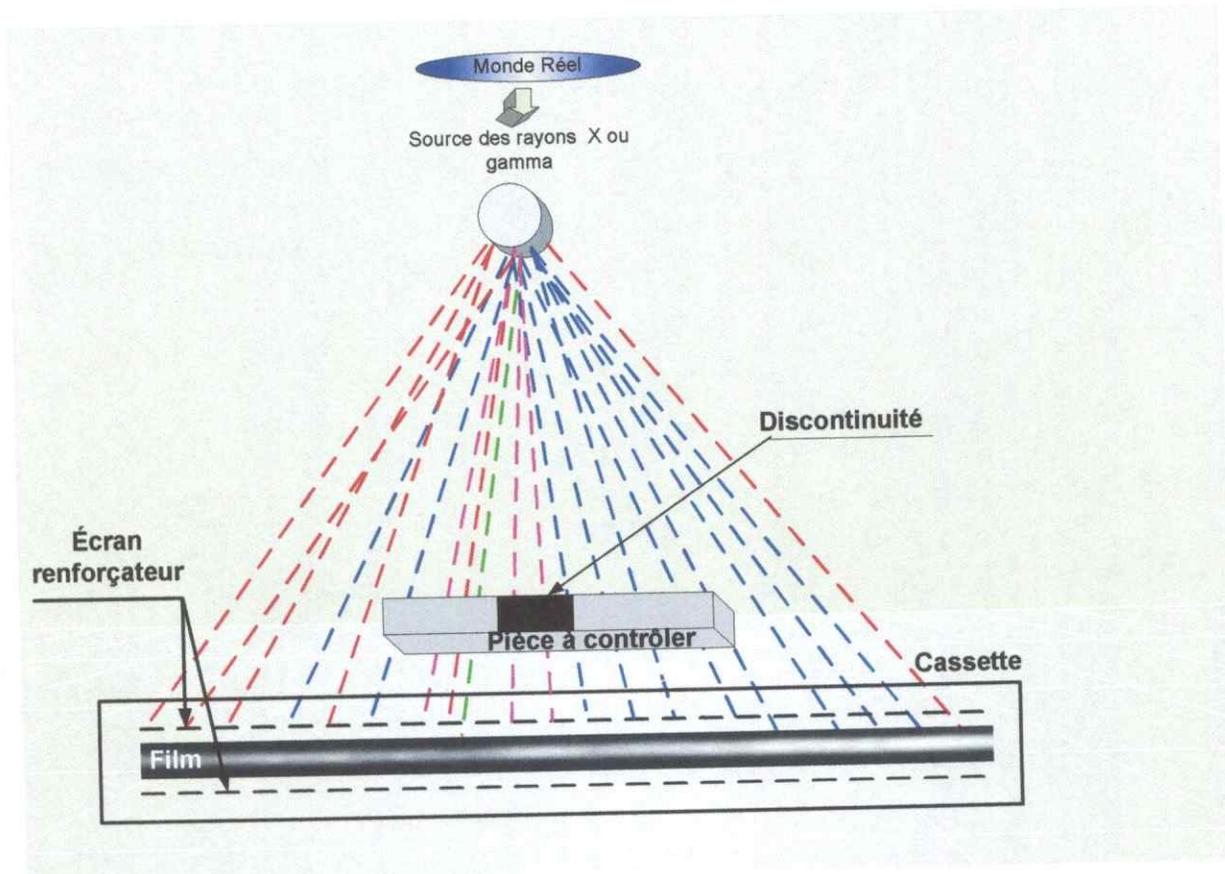


Figure I.3 Montage de base pour la radiographie sur film

La figure I.3 illustre les principaux éléments pour l'obtention d'un film radiographique. La source de rayonnement peut être un générateur de rayons X ou gamma.

IV.1 Qualité du film radiographique: [5]

D'une manière générale, la détectabilité d'un défaut sur une radiographie dépend de :

- La qualité de la radiation.
- La qualité du film.
- Les conditions d'examen du film.

La qualité de l'image du film est déterminée par trois facteurs : le contraste de l'image, la netteté de l'image et la granulation du film.

IV.2 Le contraste d'image :

A titre d'exemple, considérons un objet présentant une série de rainures de différentes profondeurs usinées en surface. La différence entre la densité de l'image d'une rainure et celle du fond de la radiographie est appelée «Contraste d'image», il est déterminé par :

- Les différences de la perméabilité de l'objet et défauts.
- Les différences d'épaisseur de l'objet.
- La forme et la position des défauts.
- La qualité de la radiation.
- L'importance de la radiation diffusée.
- Le filtrage.

IV.3 La netteté de l'image:

Les images peuvent être nettes ou floues, il s'agit là du second facteur : «le degré de netteté» de l'image que la plupart des radiologistes ont l'habitude d'appeler «le flou». Il est déterminé par :

- Les dimensions réelles du foyer ou de la source de radiation.
- La distance source - objet.
- La distance objet - film.
- Le contact entre le film et les écrans.
- Le type d'écran renforçateur utilisé.
- L'énergie de la radiation.

Remarque : Dans les cas limites de détectabilité de l'image, il est possible de montrer qu'il existe un rapport entre le contraste et le degré de netteté et que la visibilité des défauts dépend de l'association de ces deux facteurs,

IV.4 La granulation :

C'est le troisième facteur qui détermine la qualité de l'image du film. L'image reproduite sur film photographique est formée de grains- Elle offre un aspect granuleux dont le caractère plus ou moins prononcé dépend de la taille et de la distribution des particules de grains. Le grain du film peut être considéré comme un «bruit» photographique. Ce paramètre étant lié à :

- L'épaisseur de la couche d'émulsion.
- La concentration des graines d'argent (rapport argent / gélatine).
- La répartition de la taille des graines d'argent.
- L'énergie de rayonnement,

Ces trois paramètres : le contraste, le degré de netteté (flou) et la granulation (le bruit) se retrouvent dans tous les systèmes d'imagerie, se sont des paramètres fondamentaux qui déterminent la qualité du film radiographique. L'essentiel de la technique visant à obtenir une radiographie de qualité satisfaisante se résume dans l'optimisation de ces trois facteurs et à contrôler leur incidence sur la détectabilité des défauts présents dans un matériau.

CHAPITRE II
INTRODUCTION AU TRAITEMENT
D'IMAGE DES FILMS DE
RADIOGRAPHIE

I. Introduction :

Avec le progrès qu'a connu la technologie de l'ordinateur, on assiste actuellement à l'émergence des techniques de traitement numérique des images qui sont appliqués dans divers domaines tels que l'imagerie spatiale et la recherche biomédicale.

Néanmoins, ces dernières années, ces techniques intéressent de plus en plus le domaine du contrôle non destructif par radiographie, où les images des films radiographiques peuvent être numérisés et leur être appliquées les techniques du traitement numérique des images et ce, dans le but d'améliorer leur visibilité et d'en extraire les paramètres pertinents pour effectuer une interprétation [6].

Le traitement numérique des images recouvre l'ensemble des procédés d'amélioration et d'extraction d'informations qualitative des images numériques, selon les utilisateurs et les besoins requis, pour en donner, soit des nouvelles images ou bien des évaluations particulières.

L'expert radiographie a pour rôle d'inspecter chaque film afin d'y détecter la présence d'éventuels défauts qu'il doit alors identifier et mesurer. Ce travail est rendu particulièrement délicat du fait de la faible dimension de certains défauts (une fissure peut avoir une épaisseur inférieure à $200 \mu m$), du manque de contraste et du caractère bruité du film.

L'expert travaille souvent aux limites du système visuel, la part de subjectivité dans les mécanismes de détection et de mesure des défauts n'est pas négligeable.

L'emploi des techniques numériques de traitement d'image a pour but non seulement de détecter et d'identifier automatiquement les défauts mais aussi, d'une part, d'offrir une meilleure visualisation de l'information et d'autre part de formaliser les méthodes d'expertise radiographique afin de les rendre robustes et systématiques.

Avant d'aborder notre travail, l'étude fait appel à certaines notions que nous devront définir pour mieux comprendre et suivre.

II. Généralités :

II.1 Image analogique :

L'image analogique peut être décrite par une fonction analogique bidimensionnelle $F(x,y)$ représentant l'intensité lumineuse au point (x,y) .

II.2 Image numérique :

Elle est décrite par une matrice M , un point $M(i,j)$ représente l'intensité lumineuse d'un point de coordonnée spatiale (i,j) , on dit qu'une image est numérisée quand chaque point

(pixel) de l'image est codée sur un nombre de bits selon la résolution. Elle est le fruit de la quantification de l'image analogique.

II.3 Pixel [1]

C'est l'élément atomique (élémentaire) dans une image. Chaque point $M(i,j)$ de la matrice bidimensionnelle représente un pixel, il est défini par (i,j) ces coordonnées spatiales et sa couleur ou son niveau de gris.

II.4 Niveau de gris [1]

La modélisation de l'image consiste à la transformer en une fonction $f(x, y)$ dont l'amplitude représente l'intensité lumineuse, ou brillance de tout point $M(i,j)$ du plan. Dans le cas de l'image monochrome (dégradation en noir et blanc), l'intensité en chaque point est appelée niveau de gris (gray level), on distingue dans les images de qualité jusqu'à 256 niveaux de gris.

II.5 Bruit: [1]

Le bruit est un phénomène de brusque variation de la couleur ou de niveau de gris d'un pixel par rapport à ses voisins. Il existe plusieurs causes du bruit.

- Bruit lié au contexte de l'acquisition.
- Bruit lié au capteur.
- Bruit lié à l'échantillonnage.
- Bruit lié à la nature de la scène.

II.6 Région d'une image : [10]

Elle est décrite par un ensemble de pixels satisfaisant simultanément la même connexité et la même l'homogénéité.

II.6.1 Connexité

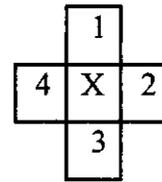
On distingue deux connexités relatives au voisinage utilisé dans une image numérique

- La quatre connexité d'un pixel $M(i,j)$ implique que le voisinage du pixel est composé de quatre pixels de coordonnées spatiales $(x+1, y)$; $(x-1,y)$; $(x,y+1)$; $(x,y-1)$.
- La huit connexité d'un pixel $M(i, j)$ implique que le voisinage du pixel est

compose de huit pixels de coordonnées $(x+1,y+1)$; $(x+1,y-1)$; $(x-1,y+1)$; $(x-1,y-1)$; $(x+1, y)$; $(x-1,y)$; $(x,y+1)$; $(x,y-1)$.

4	3	2
5	X	1
6	7	8

a- La huit connexité



b- La quatre connexité

Figure II.1 les types de connexité

II.6.2 Homogénéité

On dit qu'une région est homogène si l'ensemble de pixels qui la compose ont une spécificité similaire ou uniforme. L'homogénéité est établie par plusieurs facteurs (moyenne, variance, contraste, texture,... etc.)

II.7 Histogramme [13]

L'histogramme des niveaux gris d'une image est une fonction qui donne la fréquence d'apparition de chaque niveau de gris dans l'image.

En abscisses sont représentés les niveaux de gris de 0 à N, en coordonnées est affiché le nombre de pixels affectés par chaque niveau de gris.

L'histogramme permet de donner un grand nombre d'information sur la distribution des niveaux de gris de l'image, ce qui est plus communément appelé la dynamique de l'image. Il permet notamment d'analyser entre quelles bornes est répartie la majorité des niveaux de gris dans le cas d'une image trop claire ou trop foncée.

III. Prétraitement :

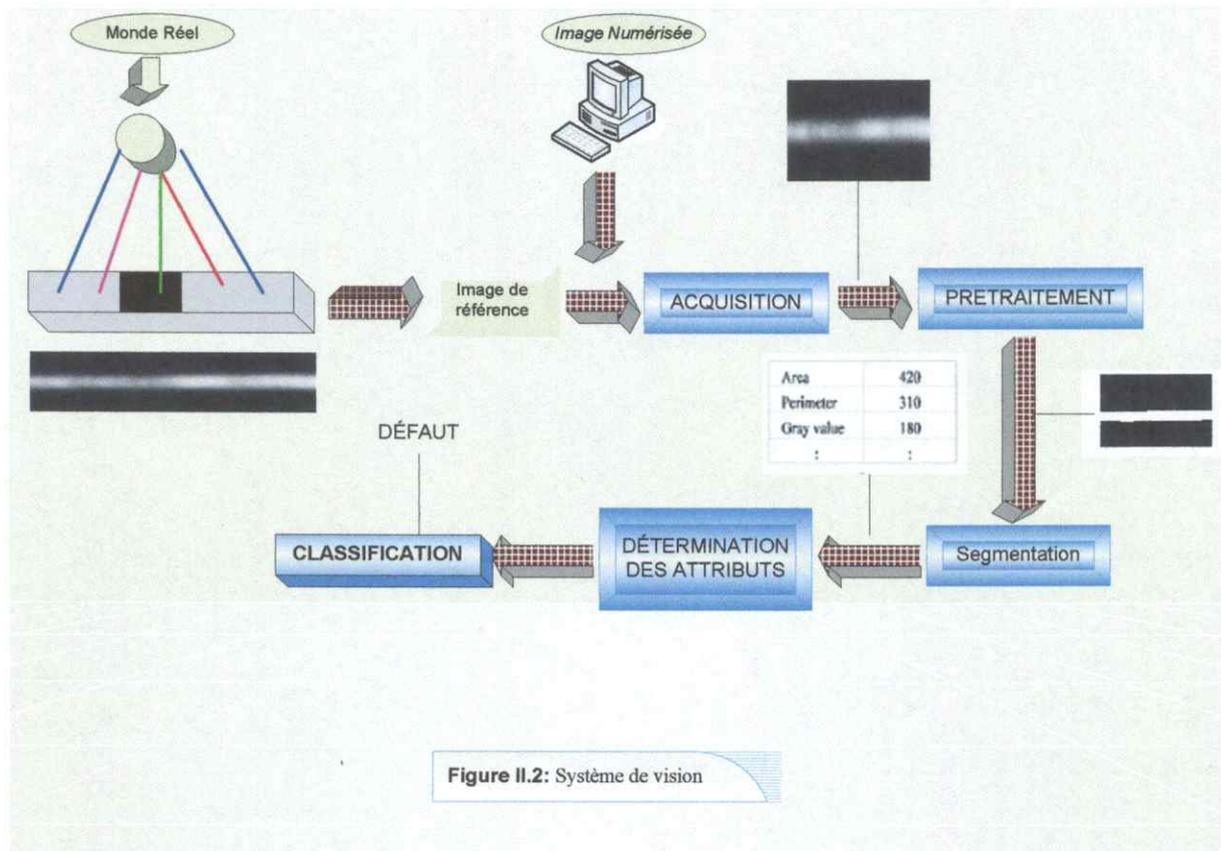


Figure II.2 Description générale d'une chaîne de vision par ordinateur

III.1 Acquisition d'une image :

L'étude de cette étape passe inmanquablement par le système d'acquisition qui fait référence : l'œil. On utilise plus couramment des caméras vidéo, ou récemment pourquoi pas des appareils photos numériques. En médecine, on utilise des imageurs IRM, TEP, scanner X, écho doppler, échographie, scintigraphie etc.

Tous ces systèmes peuvent être comparés à des capteurs. Il ne faut pas oublier qu'il y a une étape de conversion analogique/numérique. C'est souvent cette étape qui limite la résolution de l'image.

L'une des caractéristiques intéressantes de ces capteurs est la taille du plus petit élément (pixel), mais aussi l'intercorrélation de deux éléments voisins : plus cette intercorrélation est faible, meilleure est l'image.

III.2 Prétraitement de l'image : [7]

On parle aussi d'opérateur de prétraitement quand il s'agit d'améliorer l'image. L'amélioration d'image repose principalement sur des procédés mathématiques conçus afin de produire une image mieux adaptée à une application particulière. Cette amélioration peut servir dans un premier à faciliter la visualisation de l'image sur un écran d'ordinateur. Les capacités de vision de l'être humain étant limitées, il est indispensable d'adapter la dynamique de l'image à notre vision.

Le prétraitement regroupe l'ensemble de toutes les techniques par les pixels et manipulent l'image pour l'améliorer : il existe différentes approches visant à améliorer une image :

- Les opérations basées sur l'examen de l'histogramme, elles visent à pallier certains défauts de prise de vue en jouant sur la dynamique de l'image,
- Le filtrage, dont le but est de minimiser l'influence du bruit dans l'image

Ces techniques sont:

III.2.1 Amélioration du Contraste de l'image :

III.2.1.1 Table de conversion ou LUT (Look Up Table) :

L'une des applications les plus courantes des *LUT* est le **recadrage de dynamique**. Certaines images sont initialement trop claires, certaines trop sombres, ou bien peu contrastées. Cela est dû respectivement au fait que les niveaux de gris de l'image sont tassés vers le haut de l'échelle, vers le bas, ou bien sont regroupés dans un intervalle étroit. Ce défaut est très visible sur l'histogramme,

Le but de la *LUT* est alors de redistribuer les niveaux de gris de l'image afin de leur faire occuper toute la bande de nuances possibles. C'est ce que réalise le recadrage de dynamique dont le principe est donné par la figure (II).

Le choix des valeurs *Min* et *Max* du recadrage peut être effectué soit manuellement, soit automatiquement en décidant d'un certain pourcentage de pixels recadrés.

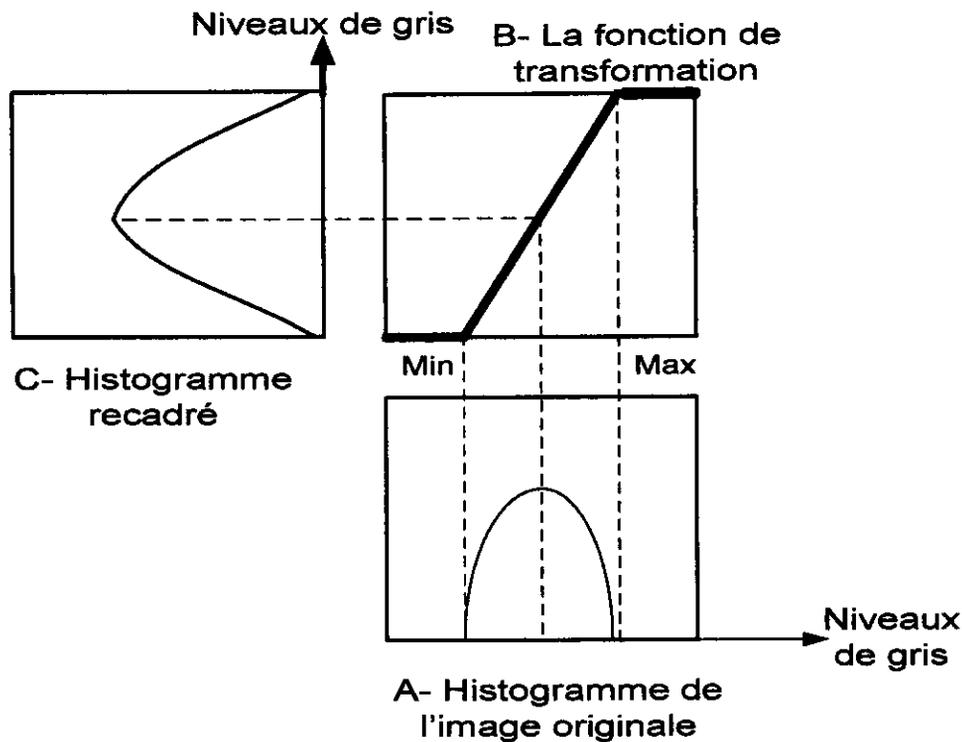


Figure II.3 Principe de recadrage dynamique

III.2.2 Filtrage du bruit de l'image : [1]

Dans de nombreux domaines scientifiques, le bruit joue un rôle fondamental. Il est à l'origine d'un grand nombre de difficultés. Ce bruit est le résultat du bruit électronique du capteur et de la qualité du digitaliseur. Mais malheureusement le bruit est non seulement lié au matériel utilisé, mais aussi à l'image elle-même: Zone de fort contraste, parties plus ou moins sombres. On peut caractériser le bruit impulsionnel, or le bruit affecte de brusques variations à des pixels isolés. Il faut alors utiliser des opérateurs de traitement plus complexes scindés bien souvent en deux sous catégories :

- les opérateurs linéaires,
- les filtres non linéaires.

La première catégorie comprend tous les opérateurs pouvant exprimer leur résultat comme une combinaison linéaire des niveaux de gris d'un voisinage de l'image. Ces filtres possèdent des caractéristiques spectrales, on parle ainsi de filtre passe-bas (l'image devient floue) ou de filtre passe-haut (les contours ressortent).

La deuxième catégorie le résultat du filtrage est déduit par une combinaison linéaire des niveaux de gris, un des exemples du filtrage non linéaire est le filtrage médian, ce dernier remplace l'intensité de chaque pixel par les intensités de voisins choisis à l'avance. La valeur

médiane U_m d'une variable aléatoire U_n prise dans un ensemble de n valeurs, satisfait les deux conditions de probabilité suivantes : $P(U_n < U_m) > 1/2$ et $P(U_n > U_m) > 1/2$.

Ainsi la somme de deux dés est 7, car il y a 36 combinaisons et 21 valent 7, on vérifie que $P(U_n < 7) = P(U_n > 7) = 21/36 > 1/2$. En pratique, on choisit une fenêtre de 3×3 pixels, centrée sur le point à filtrer et par une simple méthode de tri sur 8 nombre, on attribut la valeur médiane des gris environnants au pixel central . On peut comprendre intuitivement que le filtre médian force les pixels d'intensité hors de la normale de son voisinage à être lissés ce filtre est un passe-bas qui élimine le bruit du type "neige", mais en contrepartie il entraîne une petite perte de résolution [1].

IV. SEGMENTATION

IV.1 Introduction :

La segmentation est la première étape de l'analyse d'une image, elle consiste à faire la séparation des divers éléments de l'image en régions connexes ayant les mêmes propriétés. Ces régions peuvent être caractérisées par leurs frontières, c'est dans ce cas la segmentation par extraction de contours ou bien être directement caractérisées par les pixels qui les composent, il s'agit alors de la segmentation en régions homogènes. Les propriétés constituant le critère de segmentation peuvent être choisies de manière à effectuer une segmentation en régions ayant même niveau de gris, même texture, même couleur, etc.

Il est évident que ces deux approches de la segmentation sont duales. Cependant l'information qu'elles mettent en évidence est différente. Les contours possèdent l'essentiel des caractéristiques de forme de la région, la segmentation en régions homogènes privilégie les caractéristiques liées au critère de segmentation, donc au contenu de la région plus qu'à sa forme. Les techniques de segmentation permettent d'extraire d'une scène (image) les formes ou silhouettes, qui sont définies par leurs primitives de contour ou de surface ; ainsi à l'issue de la segmentation, les objets sont caractérisés par une séquence de pixels qui appartiennent soit à leur contour fermé, soit à leur surface.

Le domaine de la radiographie est le moteur de la recherche en segmentation car les besoins de perfectionnement y sont encore très importants et motivants, doit permettre de réaliser une partition de l'image en ensembles connexes homogènes. Il existe deux grandes catégories de segmentations :

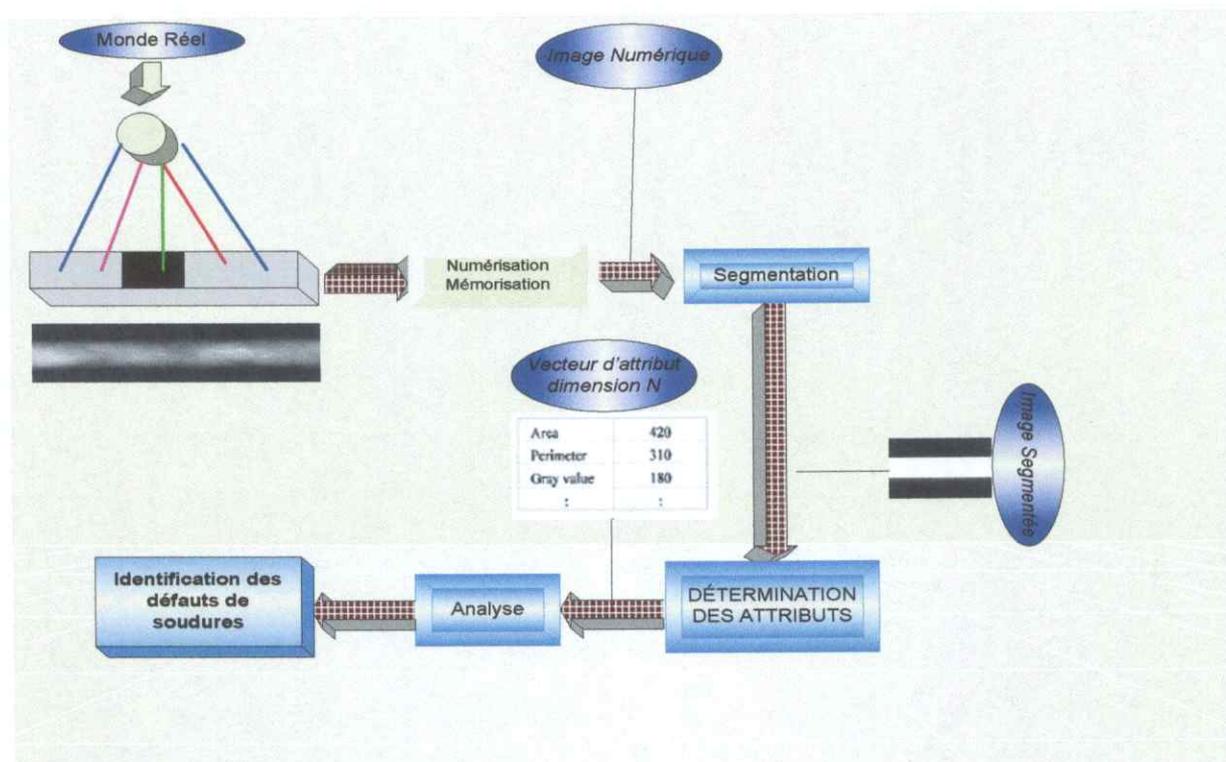


Figure II.4 Schéma de principe du processus de reconnaissance automatique de défaut de soudures.

IV.2 La segmentation en contour :

Le but de l'analyse d'image est d'obtenir une description synthétique des divers éléments qui la constituent à partir de la masse énorme d'information qu'elle contient à l'état brut. L'un des processus fondamentaux dans la chaîne de reconnaissance visuelle consiste à diminuer cette quantité d'information en ne gardant que les points essentiels de l'image. Les points de contour constituent en ce lieu une approche rationnelle.

Mathématiquement, le contour est défini comme une brusque variation de niveau de gris dans l'image. On peut caractériser cette variation par son amplitude (paramètre h) et par sa pente (paramètre p). Cette définition détermine donc toute variation de niveau de gris, aussi faible soit-elle, comme étant un point de contour de plus ou moins grande intensité, avec les opérateurs de différenciation, les opérations: le gradient, laplacien et Atténuation des effets du bruit avec les filtres passe-bas, la mise en évidence des contours, la réduction des contours (ou amincissement des lignes de contours), la binarisation des contours(ou lignes de contour), la description des contours; on se trouve alors confronté à un problème de représentation de résultat par des primitives simples. La segmentation orientée contour connaît de nombreux progrès autour de l'utilisation de *contours actifs* ou des ensembles de niveaux. L'introduction d'aspects probabilistes (chaîne de Markov et champs de Markov) a permis de travailler en

réduisant la connaissance a priori nécessaire pour obtenir un traitement satisfaisant.

Dans cette étape on retrouve souvent une partie de classification des pixels en classes. On essaye de regrouper au sein d'un même ensemble, aussi appelé classe, les pixels présentant une même caractéristique : niveau de gris compris dans un certain intervalle ou dérivée seconde supérieure à un certain seuil.

IV.2.1 Opération de différenciation :

Un premier traitement dans le processus d'extraction de contours consiste à mettre ceux-ci en évidence. Ceci est réalisé en dérivant l'image une fois, on obtient alors le gradient ou bien deux fois, on obtient alors un *laplacien*. La figure suivante montre les effets de la dérivation en présence d'un contour. Notons que ces opérateurs de différenciation peuvent être considérés comme faisant partie des prétraitements. Mais ils peuvent être séparés des autres types de filtrage du fait de leur application spécifique à l'extraction de contours.

- Gauss : rendre l'image floue en calculant une moyenne des niveaux en chaque point, moyenne pondérée par une gaussienne
- Prewitt : un détecteur de contours mélangeant du passe-bas et du passe-haut.
- Laplacien : un détecteur de contour du deuxième ordre, le contour se trouve là où le laplacien change de signe (passage par zéro de la dérivée seconde)
- Canny : un filtre optimal (au sens de trois critères) pour détecter les contours
- Deriche : une variante du filtre de Canny tout aussi efficace
- Sobel : le plus illustre mais aussi le plus simple des détecteurs de contours. Un flou gaussien suivi d'une différence (horizontale pour les contours verticaux et verticale pour les contours horizontaux)

Marr-Hildreth: le laplacien d'une gaussienne, un détecteur de contours permettant de limiter les amplifications des hautes fréquences des dérivées secondes par une gaussienne de variance ajustable.

Les opérateurs "Roberts", "Sobel", "Prewitt" et "Laplacien" ne privilégient aucune direction. Les opérateurs de type "Compass" permettent la détection de contour selon certaines directions.

IV.2.2 La description des contours :

Elle permet d'organiser les points de contour en structures simples telles que segments de droite, arcs de cercle et de détecter les contours fermés afin d'isoler ceux-ci. Ceci dans le but de réduire la quantité de données en isolant chaque contour et le décrivant suivant un code. Il existe plusieurs modes de codage parmi eux :

- Codage cartésien.
- Codage Squelettisation
- Codage de HOUGH
- Approximation par segments de droite.
- Approximation polygonal

IV.3 Segmentation en régions homogènes :

La segmentation en région homogène définit par les méthodes : globales ou par séparation (split) Quat Trees, locales ou par fusion (merge), par séparation fusion (Split and merge). On peut passer à l'étape de segmentation qui doit permettre de réaliser une partition de l'image en ensembles connexes homogènes, Elle vise à segmenter l'image en se basant sur les propriétés intrinsèques de la région selon des critères de segmentation tels que :

- Niveau de gris ou la couleur.
- La texture.
- Le mouvement.

Nous citons trois approches de la segmentation en régions, qui sont indépendantes du critère choisi :

IV.3.1 Méthodes globales ou par séparation (Split) - Quat Trees :

Le principe de ces méthodes repose sur le test de la validité du critère sur une large zone de l'image, s'il est validé de manière globale le traitement s'arrête. Sinon on décompose l'image jusqu'à ce que le critère soit validé,

IV.3.2 Méthodes locales ou par fusion (Merge) :

Ces méthodes ont le principe inverse de la séparation. L'idée consiste à explorer l'image à partir de petites régions et de faire croître celles-ci,

IV.3.3 Méthodes par séparation-fusion (Split and Merge) :

Ces méthodes combinent entre les deux précédentes méthodes. On réalise tout d'abord une division en construisant le Quad- Tree dans le sens racine vers feuilles, puis on parcourt l'arbre dans le sens feuille vers racine en fusionnant les régions ayant même critère.

Conclusion :

La phase prétraitement et segmentation ont présenté les techniques de segmentation de scène (image), c'est à dire le seuillage d'histogramme, extraction du contours et les opérateurs de différenciation et les régions homogènes. Ces techniques permettent d'extraire d'une image les formes, qui à ce stade sont définies par leurs primitives de contours ou de surface, ainsi à l'issue de la segmentation, l'image est caractérisés par une séquence de pixels qui appartiennent soit à leur contours fermé, soit à leur surface, à ce niveau de traitement d'image, si l'on veut aller jusqu'à la reconnaissance des défauts de soudures et la compréhension de l'image dans les uses cases suivants nous exposons les moyens de décrire ou de modéliser un défaut dans un plan de l'image afin de lui donner une identité l'image de défaut ne sera plus une forme géométrique mais deviendra un vecteur d'attribut quand le défaut est identifié, il quitte le plan de l'image et accède au monde abstrait de la classification. Cette opération consiste à entrer chaque défaut dans une classe.



CHAPITRE III
L'ANALYSE QUANTITATIVE DES
IMAGES DES DEFAUTS DE
SOUDURES

I. Introduction:

Une fois la segmentation des images effectuée, la dernière étape en vue de la reconnaissance de forme consiste en l'analyse quantitative, est d'extraire les objets du fond pour calculer les valeurs des attributs les plus appropriés à différencier les objets les uns des autres. C'est ce que nous appelons l'analyse quantitative des images. Souvent, le but de ces mesures est de constituer une base de descripteurs qui sera exploitée pendant la classification. Les images des radiogrammes des joints soudés contiennent souvent des défauts qu'on doit identifier et quantifier afin de décider de leur nocivité en se référant aux normes qui régissent le contrôle non destructif. Après la segmentation d'une image radiographique fournissant une description en terme de régions (le défaut de soudure et le fond), le problème est ensuite d'interpréter leur contenu. Il s'agit donc de déterminer des attributs efficaces permettant de caractériser ces régions de défauts et même de les reconnaître comme éléments de classe aisément identifiables. En radiographie industrielle, on peut obtenir des radiogrammes où les images des défauts de soudure, s'il y a lieu, peuvent avoir diverses résolutions et orientations. A titre d'exemple, une fissure reste une fissure quelles que soient sa grandeur et sa direction et une inclusion est identifiée comme étant une inclusion en dépit de son emplacement ou sa taille [1].

Un problème majeur dans la reconnaissance de tels défauts est que la même forme de défaut peut être vue à partir de plusieurs angles et ce, suivant l'orientation et la distance de l'ouvrage soudé par rapport à la source d'irradiation. Donc, dans le but de comparer une image donnée d'un défaut avec une autre qui est stockée dans une base de données, les méthodes qui ne se basent pas sur les invariants doivent chercher dans un espace de paramètres multidimensionnel, i.e. réaliser plusieurs versions obtenues par transformations appliquées au modèle image pour voir si une version coïncide avec l'image originale. Les attributs géométriques invariants ont donc pour but de faciliter la recherche, en faisant une correspondance entre les images des défauts observées et les images stockées ou de référence à partir des invariants déjà calculés [3].

Une panoplie d'attributs satisfaisant les conditions suscitées sera proposée dans ce chapitre. Ces invariants géométriques découleront du calcul de paramètres géométriques (surface, périmètre, etc.)

II. Extraction de contours : [3]

La segmentation nous a permis d'identifier sur l'image un ensemble R de points image connexes constituant une ou plusieurs régions en niveau de gris qui représentent le défaut de soudure, ainsi que le cordon de soudure.

Le problème qui se pose est comment isoler le défaut des autres défauts voisins d'une part, et du cordon de soudure d'une autre part.

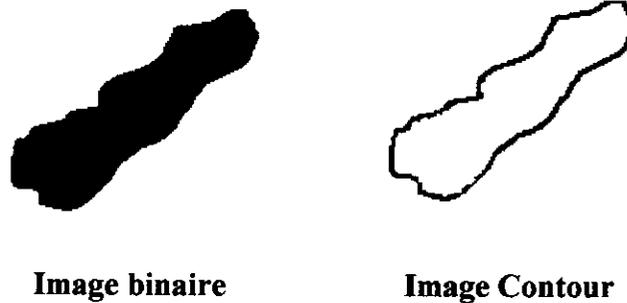


Image binaire

Image Contour

Figure III.1 Extraction de contours

III Détermination des paramètres:

Après le suivi de contour donc l'isolation des défauts de soudure, il est possible de mesurer un grand nombre de paramètres pour caractériser ces défauts pour une éventuelle classification, Mais il est absolument indispensable, que ces paramètres aient une signification physique, morphologique et mathématique.

III.1 Mesure des paramètres géométriques (attributs morphologiques) :

Après que chaque ensemble connexe représentant l'objet d'intérêt (défaut de soudure) ait été isolé, ses attributs morphologiques sont calculés. La définition et la formulation mathématique de chaque attribut, soit (X) étant une région (connexe) de l'image dont les pixels ont pour coordonnées (X_i, X_j) , on peut disposer les attributs suivants:

III.1.1 La surface (Area):[3]

$$A = \sum_{x=1}^M \sum_{y=1}^N b(x, y)$$

Equation III.1

avec $N \times M$ présente la largeur et la hauteur de l'image et

$$b(x,y) = \begin{cases} 1 & \text{Si } (x,y) \in R \\ 0 & \text{Si non} \end{cases}$$

III.1.2 Le périmètre : [3]

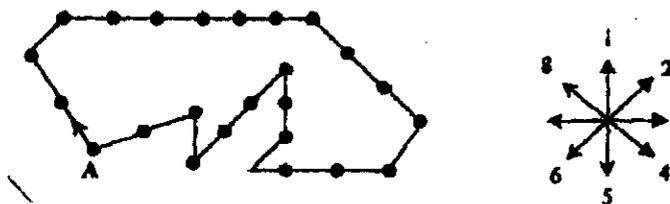
Le périmètre P est calculé comme la somme des distances entre pixels consécutifs du contour (somme de 1 et de $\sqrt{2}$). Le résultat est sensiblement différent s'il est calculé à partir du contour intérieur ou du contour extérieur de la région [29]. Elle est définie par la relation

$$P = \sum_{i=1}^n \left[1 + \left(\frac{\sqrt{2} - 1}{2} \right) (1 - (-1)^{C_i}) \right]$$

Equation III.2

avec C_i est la chaîne numérique (C_1, C_2, \dots, C_i) cette chaîne représente le voisinage d'un pixel. Elle est définie par le code de FREEMAN La chaîne convient à la description de formes constituées de segments droits ayant des longueurs et directions données. Cette représentation s'appuie sur les connectivités 4 et 8, ainsi la chaîne numérique de *FREEMAN*, basée sur la connectivité 8, code tout contour fermé par un chaînage qui emprunte la notation des gradients circulaires à 8 directions. Un exemple est donné par la figure suivante :

Chaîne: 88233333444466777211



Plutôt que de mémoriser une liste de coordonnées, on peut exploiter le fait que deux points successifs de la liste soient des points voisins et se contenter d'indiquer comment passer de l'un à l'autre. *FREEMAN* a proposé de stocker pour le point P la direction suivant laquelle il est obtenu à partir du point P_{i-1} . Cette direction dépend de la notion d'adjacence utilisée :

Soit quatre directions possibles en 4-connectivités et huit directions en 8-connectivités. A chaque direction est associé un code alpha variant entre 1 à 8 Le code alpha est impair si le déplacement est horizontal ou vertical, et pair si le déplacement est diagonal.

III.1.3 Le Centre de gravité (Center of Mass) : [1]

Le centre de gravité d'un objet est la position du centre de l'objet en ce qui concerne sa « masse ». Il représente la position idéale de l'objet dans l'image où il réside. L'abscisse \bar{x} et l'ordonnée \bar{y} du centre de gravité de l'objet peuvent être définies comme suit :

$$\bar{X} = \frac{\sum_{x=1}^N \sum_{y=1}^N x b(x,y)}{A} \quad \bar{Y} = \frac{\sum_{x=1}^N \sum_{y=1}^N y b(x,y)}{A}$$

Equation III.3

III.1.4 L'angle d'orientation:

L'angle d'un objet est défini comme étant l'angle entre l'axe des abscisses et l'axe autour duquel l'objet peut être tourné avec un minimum d'inertie. Soit l'ensemble des coordonnées des pixels appartenant à l'objet et (x, y) désigne la position du centre de gravité de l'objet. La matrice de covariance de l'objet est définie par:

$$\sum_{x,y} \begin{vmatrix} \mu_{2,0} - \sum_i^n (x_i - \bar{x})^2 & \mu_{1,1} - \sum_i^n (x_i - \bar{x})(y_i - \bar{y}) \\ \mu_{1,1} - \sum_i^n (x_i - \bar{x})(y_i - \bar{y}) & \mu_{0,2} - \sum_i^n (y_i - \bar{y})^2 \end{vmatrix}$$

Equation III.4

Avec cette définition de la matrice de covariance qui est une matrice semi définie positive et à coefficients réels, l'angle α de l'objet peut être calculé par :

$$\alpha = \frac{1}{2} \arctan \left(\frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}} \right)$$

Equation III.5

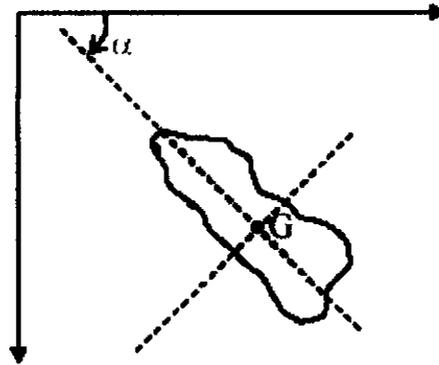


Figure III.3 Orientation de l'objet

III.1.5 La longueur et la largeur du rectangle exinscrit : [3]

Les vecteurs propres de la matrice de covariance (matrice d'inertie) correspondent aux directions des axes principaux d'inertie suivant lesquels l'objet présente le plus fort et le plus faible moment d'inertie [30].

Le premier axe principal est défini par l'équation de la droite faisant un angle α avec l'axe des abscisses et passant par le centre de gravité G de l'objet :

$$y_1 = \tan(\alpha) \times x + b_1, \quad \text{avec } b_1 = \bar{y} - \tan(\alpha) \times \bar{x}$$

Equation III.6

Le deuxième axe principal peut être défini comme précédemment par l'équation de la droite faisant un angle $\alpha + \pi/2$ avec l'axe des abscisses et passant par le centre de gravité de l'objet

$$y_2 = \tan(\alpha - \pi/2) \times x + b_2, \quad \text{avec } b_2 = \bar{y} - \tan(\alpha + \pi/2) \times \bar{x}$$

Equation III.7

Le rectangle exinscrit est le rectangle minimal qui renferme l'objet et dont les cotés sont parallèles aux axes principaux de l'objet. La longueur LNG représente la longueur du rectangle et la largeur LRG est celle de ce dernier.

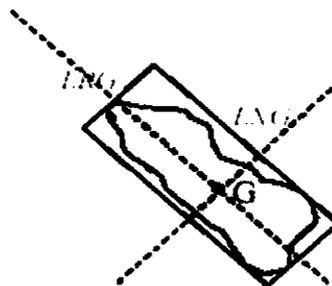


Figure III.4 Longueur et largeur du rectangle exinscrit

III.1.6 Le diamètre maximal:

Le diamètre maximal D_{max} d'un objet est défini comme étant la plus grande distance entre deux points du contour de l'objet

$$D_{max} = \max_{i=1}^n \max_{j=1}^n \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

Equation III.8

III.1.7 Le rayon du cercle maximal inscrit :

On calcule ici le rayon du plus grand cercle inscrit dans la région de l'objet. On l'appelle R_{max} . (Voir Fig. III.5).

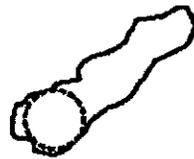


Figure III.5 Cercle maximal inscrit

III.2 Attributs géométriques invariants :

Les attributs que nous allons définir sont invariants par rapport aux transformations géométriques de translation, rotation et changement d'échelle. Quant à l'angle qui définit l'orientation de l'objet, il ne dépend pas de la translation et du changement d'échelle.

III.2.1 La compacité :

Elle est définie à partir de la surface et du périmètre de l'objet.

$$Comp = 4\pi A / P^2$$

Equation III.9

avec A: Surface , P: Périmètre

III.2.2 L'élongation:

Elle est définie par le rapport entre la longueur et la largeur du rectangle exinscrit [31].

$$Elong = LNG / LRG$$

Equation III.10

III.2.3 La rectangularité:

Elle est définie par le rapport entre la surface de l'objet et celle du rectangle exinscrit.

$$Rct = A / LNG. LRG$$

Equation III.11

III.2.4 La symétrie:[3]

La symétrie repose sur la loi donnée ci-dessous

$$Sym = SymH * SymV$$

Equation III.12

SymV et *SymH* représentent respectivement la symétrie par rapport à l'axe principal et la symétrie par rapport l'axe perpendiculaire à l'axe principal (Fig. III.6). Elles sont données par les algorithmes suivants :

Si $(S4+S3) < (S1+ S2)$

$SymV = (S4+S3) / (S1+ S2)$

Si non

$SymV = (S1+ S2) / (S4+S3)$

Si $(S2+S3) < (S1+ S4)$

$SymH = (S2+S3) / (S1+ S4)$

Si non

$SymH = (S1+ S4) / (S2+S3)$

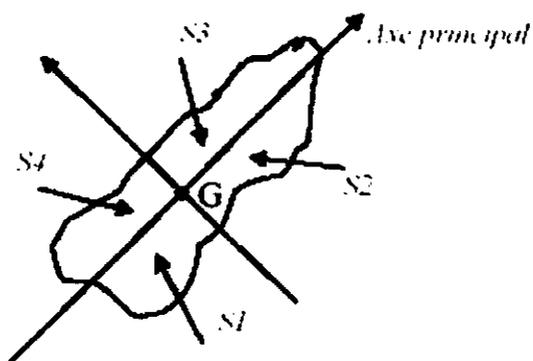


Figure III.7: Illustration de la notion des surfaces partielles

III.2.5 L'indice d'allongement

Il est défini à partir de la surface et du diamètre maximal de l'objet.

$$I_a = \pi \times D_{\max}^2 / 4A$$

Equation III.13

III.2.6 L'indice d'écart au cercle inscrit:

Il est lié à la surface et le rayon du cercle maximal inscrit dans la région de l'objet.

$$I_r = 1 - \pi R_{\max}^2 / A$$

Equation III.14

III.2.7 La planéité:

La planéité est un attribut géométrique invariant définie par l'équation suivante :

$$PLA = M_p / A^2$$

Equation III.15

où A est la surface de la région R et M_p le moment d'inertie polaire donné par

$$M_p = \sum_x \sum_y r^2$$

Equation III.16

avec r la distance entre le centre de gravité et un pixel appartenant à l'objet. Alors, l'expression du M_p devient, en utilisant l'expression des moments centraux normalisés

$$M_p = \sum_x \sum_y [(x - \bar{x})^2 + (y - \bar{y})^2] = \sum_x \sum_y (x - \bar{x})^2 + \sum_x \sum_y (y - \bar{y})^2 = \mu_{20} + \mu_{02}$$

Equation III.17

III.2.8 L'eccentricité:

L'ellipse équivalente d'une image (Fig. III.8) est définie comme étant l'ellipse dont le plus petit et le plus grand moment d'inertie sont égaux à ceux de l'image. Le centre de gravité de l'ellipse est le centre de gravité (\bar{x}, \bar{y}) de l'objet. L'orientation de l'ellipse est la même que celle de l'objet. En utilisant des moments centraux, les longueurs des axes semi majeur et semi mineur de l'ellipse équivalente sont donnés par :

$$(a, b) = \left(\frac{2 \left[\mu_{20} + \mu_{02} \pm \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2} \right]}{\mu_{00}} \right)^{1/2}$$

Equation III.19

L'anisométrie est le rapport des axes semi-majeur et semi-mineur de l'ellipse équivalente :

$$Ecc = b / a$$

Equation III.20

IV. Relation entre les attributs invariants proposés et les types des défauts de soudure :

Compacité (Comp) : La valeur de cet attribut est comprise entre 0 et 1. Elle est proche de 0 dans le cas d'une droite et elle est égale à 1 dans le cas d'un cercle parfait. Cet attribut prend des petites valeurs pour les défauts dont la forme est très fine (fissure, manque de fusion) et des valeurs proches de 1 pour les défauts sphéroïdaux (porosité, inclusion de tungstène, inclusion de solide sphéroïdale).

Elongation (Elong) : Cet attribut nous décrit l'espace occupé par le défaut en termes de zone d'encadrement. Des grandes valeurs du rapport entre la longueur et la largeur de cette zone sont caractéristiques des défauts longitudinaux (fissure, manque de fusion, manque de pénétration, soufflure allongée, caniveau, etc.).

Rectangularité (Rct) : Les valeurs de cet attribut sont comprises entre 0 et 1. Elle est égale à 1 dans le cas d'un rectangle parfait. Cet attribut met en évidence les défauts dont la forme est très proche d'un rectangle. C'est le cas du manque de pénétration.

Eccentricité (Ecc) : Elle est liée à la direction des axes principaux de la matrice d'inertie du défaut. Sa valeur est proportionnelle à l'allongement du défaut. Elle est égale à 1 dans le cas d'un carré ou d'un cercle.

Symétrie (Sym) : Les valeurs de cet attribut sont comprises entre 0 et 1. La valeur 1 correspond à une forme parfaitement symétrique. L'aspect asymétrique des défauts peut être traduit par des petites valeurs de cet attribut. Ce dernier aspect peut caractériser les inclusions de laitiers, les soufflures vermiculaires et les fissures radiales.

Indice d'allongement (Ia) : Les grandes valeurs de cet indice mettent en évidence les fissures fines et rectilignes.

Indice d'écart au cercle inscrit (Ir) : La valeur de cet indice est maximale (proche de 1) pour les défauts allongés et minimale (proche de 0) pour les défauts ronds.

V Conclusion

Nous avons proposé dans ce chapitre, un ensemble d'attributs qui se basent tous sur les caractéristiques géométriques des régions des défauts de soudure extraits. Les valeurs du vecteur d'attributs ainsi calculées seront réduire par l'Analyse en composantes principales abordées dans le prochain chapitre.



CHAPITRE IV

LA REDUCTION DES DONNEES PAR L'ANALYSE EN COMPOSANTE PRINCIPALE (ACP)

I. Généralités:

L'analyse en composantes principales (ACP) s'utilise quand il s'agit de décrire un tableau X de valeurs numériques continues du type " *individus - caractères* ".

Par exemple, évolution des dépenses des ménages Européens dans différents secteurs, comme l'alimentation, les loisirs, l'automobile etc... En fonction des années. Dans notre cas, les attributs géométriques invariants calculés (Compacité, Elongation, Réctangularité,..etc) sont considérés comme *caractères* et les images des défauts de soudure *individus*.

On est ainsi en présence d'un tableau *individus caractères*, X (n,p) de n individus (les attributs) et p caractères les défauts.

Comme on l'a vu précédemment les individus forment un nuage de n points dans l'E.V R^p (espace des individus) et les caractères, un nuage de p points dans l'E.V R^n (espace des caractères).

Si l'on veut représenter les individus (ou caractères) sur un graphique plan on aura une figure déformée de la représentation exacte.

La projection *raccourcit* les distances, on cherche donc un plan (π) dit *plan principal* tel que la somme des carrés des projections des distances soit maximal.

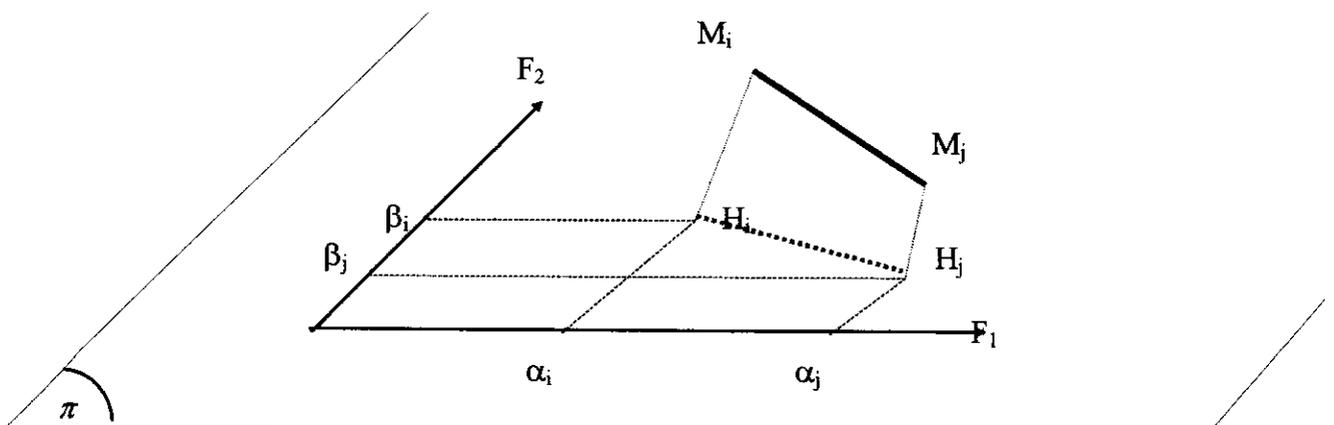
Si on appelle $d(M_i ; M_j)$ la distance entre les points M_i et M_j alors, on cherche (π) le plan principal tel que :

$$\sum d^2(H_i; H_j) \text{ maximum } (H_i, H_j \text{ projections orthogonales, respectivement de } M_i, M_j \text{ sur } (\pi))$$

On aura : $d(H_i, H_j) \leq d(M_i, M_j)$

Soient deux droites F_1 et F_2 ($F_1 \perp F_2$) telles que:

$$d^2(H_i; H_j) = d^2(\alpha_i; \alpha_j) + d^2(\beta_i; \beta_j)$$



La méthode consiste à chercher:

F_1 telle que $\sum d^2(\alpha_i; \alpha_j)$ soit maximal.

puis F_2 telle que $\sum d^2(\beta_i; \beta_j)$ soit maximal.

Remarque: Comme on l'a vu précédemment, on peut étendre le procédé en dehors du plan. Sera alors trouvés les droites $F_3, \dots, F_\alpha, \dots, F_p$ orthogonales, qui seront les *axes principaux* du nuage (support des vecteurs $\bar{u}_1, \dots, \bar{u}_\alpha, \dots, \bar{u}_p$).

Les nouvelles composantes des vecteurs OM_i sur ces axes sont alors les **composantes principales**.

$$C(\alpha, i) = X_i u_\alpha = \sum_{k=1}^p x_{ik} u_{k\alpha}$$

$C(\alpha, i)$ étant la *composante principale* du point M_i (ou vecteur OM_i) sur le $\alpha^{\text{ième}}$ axe principal.

II. Choix d'une métrique en ACP : [7]

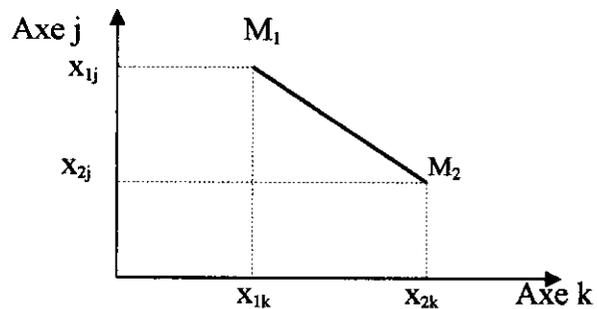
Plaçons-nous par exemple dans l'espace des individus R^p .

Chaque individu est défini par p coordonnées et est un vecteur de R^p

$\bar{e}_i = OM_i$ ayant pour composantes:
(x_{1i}, \dots, x_{ip}).

Comment calculer la distance entre 2 individus

- En physique: Pythagore



$$d^2(M_1; M_2) = (x_{1k} - x_{2k})^2 + (x_{1j} - x_{2j})^2$$

- Différent en statistique où les dimensions sur chaque axe ne sont pas les mêmes.

On part du principe que:

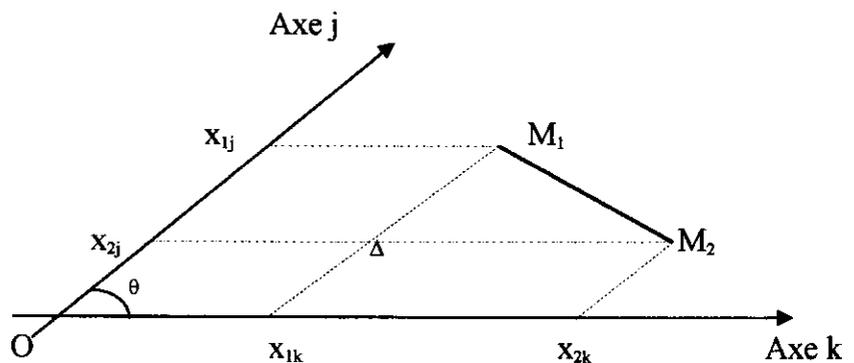
$$d^2(M_1; M_2) = a_1(x_{11} - x_{21})^2 + a_2(x_{12} - x_{22})^2 + \dots + a_p(x_{1p} - x_{2p})^2$$

Ce qui revient à multiplier chaque caractère par $\sqrt{a_j}$, $a_j > 0$

Pythagore n'est valable que si les axes sont orthogonaux.

Si on prend des axes obliques faisant entre eux un angle θ alors:

$d^2(M_1; M_2)$ se calcule avec la différence des 2 vecteurs:



$$\Delta \vec{M}_2 - \Delta \vec{M}_1 = M_1 \vec{M}_2 \Rightarrow d^2(M_1; M_2) = (\Delta \vec{M}_2)^2 + (\Delta \vec{M}_1)^2 - 2 \Delta \vec{M}_2 \Delta \vec{M}_1$$

soit: $d^2(M_1; M_2) = (x_{1k} - x_{2k})^2 + (x_{1j} - x_{2j})^2 - 2(x_{1k} - x_{2k})(x_{1j} - x_{2j}) \cos \theta$

$$d^2(M_1; M_2) = \sum_{k=1}^p \sum_{j=1}^p m_{ij} (x_{1k} - x_{2k})(x_{1j} - x_{2j})$$

matriciellement:

$$d^2(M_1; M_2) = (\mathbf{M}_1 - \mathbf{M}_2)' \mathbf{M} (\mathbf{M}_1 - \mathbf{M}_2)$$

Remarque: Pythagore revient à choisir $\mathbf{M} = \mathbf{I}$.

L'espace des individus étant muni d'une structure euclidienne, $\mathbf{M} = [m_{kj}]$: matrice de l'espace euclidien peut être n'importe quelle matrice symétrique définie positive.

On définit ensuite le produit scalaire de 2 vecteurs de l'espace des individus:

$$\vec{e}_1 = \vec{OM}_1,$$

$$\vec{e}_2 = \vec{OM}_2$$

Le produit scalaire des 2 vecteurs dans l'espace muni de la norme M s'exprime par :

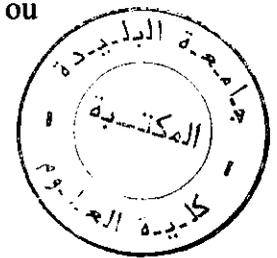
$$\langle \vec{e}_1; \vec{e}_2 \rangle_M = \mathbf{e}_1' \mathbf{M} \mathbf{e}_2$$

Remarque: Le produit scalaire de \mathbf{e}_1 par lui même est noté $\|\mathbf{e}_1\|_M^2$ et $\|\mathbf{e}_1\|_M$ analogue au module de \vec{e}_1 s'appelle la *M norme de \mathbf{e}_1* .

La métrique généralement utilisée en ACP normée est la métrique diagonale qui revient à pondérer les caractères en $1/s^2$ s représentant l'écart-type du caractère, ainsi :

$$\mathbf{M} = \mathbf{D}_{1/s^2} = \begin{pmatrix} 1/s_1^2 & & 0 \\ & 1/s_2^2 & \\ 0 & & 1/s_p^2 \end{pmatrix}$$

Ce qui revient à diviser chaque variable par son *écart-type*.



Ainsi cette métrique donne à chaque caractère la même importance (un caractère dispersé "n'absorbe" pas ainsi toute la variance) et d'autre part les distances entre individus sont adimensionnelles ce qui permet de traiter des caractères d'unités différentes.

En posant $M=T'T$, alors le produit scalaire $\langle \bar{e}_1; \bar{e}_2 \rangle_M = e_1' M e_2$ peut s'écrire:

$$e_1' T' T e_2 = (T' e_1)' (T' e_2) = \langle T e_1; T e_2 \rangle_1$$

Tout se passe comme si on avait transformé les données initiales du tableau X par la matrice T et utilisé ensuite le produit scalaire ordinaire.

En conclusion, il revient à remplacer le tableau X par $Y=XT'$ et à prendre comme *métrique M* la matrice unité I avec $T= D_{1/n}$.

Y est appelée la matrice des **coordonnées centrées réduites**.

La métrique ayant été définie, on analyse les comportements des données dans chaque espace R^p et R^n .

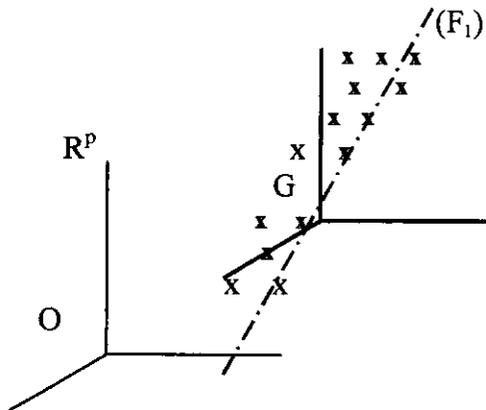
III. Analyse dans l'espace des individus R^p :

III-1. Coordonnées centrées:

Les n points de l'espace R^p sont des individus. Il est nécessaire de prendre comme nouvelle origine le centre de gravité du nuage car il importe de connaître sa forme et non sa position.

Il est pris ainsi comme nouvelle origine le centre de gravité G du nuage.

Si X_{ij} sont les coordonnées brutes des n points M_i dans R^p alors:



$$x_{ij} = (X_{ij} - \bar{x}_j) \quad x_{ij} : \text{Coordonnées centrées}$$

$$\bar{x}_j = \sum_{i=1}^n X_{ij} \quad \bar{x}_j : \text{Moyennes arithmétiques ou coordonnées de G}$$

$$G(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_j, \dots, \bar{x}_p)$$

III-2. Coordonnées centrées réduites:

Comme on l'a vu précédemment, la matrice des coordonnées centrées réduites est:

$$Y = XT' \text{ avec } T'T = M = D_{1/s^2}$$

Il est nécessaire de calculer D_{1/s^2} la variance d'un caractère s'exprimant par :

$$s_j^2 = \frac{1}{n} \sum_{i=1}^n x_{ij}^2 \quad \text{ou} \quad s_j^2 = \frac{1}{n} \sum_{i=1}^n (X_{ij} - \bar{X}_j)^2 \quad \text{dans le cas où il y a équi-pondération}$$

Si on appelle D la matrice diagonale des poids statistique, d'ordre n :

$$D = \begin{pmatrix} p_1 & & 0 \\ & \ddots & \\ 0 & & p_n \end{pmatrix} \quad \text{ici} \quad D = \begin{pmatrix} 1/n & & 0 \\ & \ddots & \\ 0 & & 1/n \end{pmatrix}$$

On peut déterminer la matrice V de variance covariance:

$$V = X'DX$$

Par exemple si $X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$ on obtient:

$$X'DX = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} x_{11}^2 + x_{21}^2 & x_{11}x_{12} + x_{21}x_{22} \\ x_{12}x_{11} + x_{22}x_{21} & x_{12}^2 + x_{22}^2 \end{pmatrix}$$

où les éléments diagonaux (termes carrés), représentent la *variance* et les autres éléments (termes rectangles), la *covariance*

D'une manière générale si X est la matrice (n,p) des coordonnées centrées, alors V est une matrice (p,p) symétrique de terme général :

$$s_{ij} = \frac{1}{n} \sum_{k=1}^n x_{ki} x_{kj}$$

Pour $i = j$ $s_{ii} = s_{jj} = s_j^2 = \frac{1}{n} \sum_{k=1}^n x_{kj}^2$ on définit ainsi la variance d'un caractère ce qui permet de

déterminer la matrice de la métrique utilisée D_{1/s^2}

Il faut donc prendre la diagonale de la matrice V :

$$D_{1/s^2} = \text{diag}(V)^{-1} = M$$

On détermine ainsi la matrice T telle que:

$$\mathbf{T}'\mathbf{T} = \mathbf{M} \quad \text{avec} \quad \mathbf{T} = \mathbf{D}1/s \quad \text{et}$$

$$\mathbf{T} = \begin{pmatrix} 1/s_1 & 0 \\ 0 & 1/s_p \end{pmatrix}$$

Ainsi, la matrice \mathbf{Y} des coordonnées centrées réduites s'exprime par :

$$\mathbf{Y} = \mathbf{X}\mathbf{T}'$$

Remarque: Pour calculer la matrice \mathbf{V} de variance-covariance de terme général s_{ij} on aurait pu diviser chaque terme x_{ij} de la matrice des coordonnées centrées par : $\frac{1}{\sqrt{n}}$, ce qui donne la

nouvelle matrice $\tilde{\mathbf{X}} = [\tilde{x}_{ij}]$ de terme général $\tilde{x}_{ij} = \frac{s_{ij}}{\sqrt{n}}$

La matrice de variance-covariance s'exprime alors :

$$\mathbf{V} = \tilde{\mathbf{X}}'\tilde{\mathbf{X}}$$

Egalement pour la matrice \mathbf{Y} on définit la matrice $\tilde{\mathbf{Y}}$ de terme général $\tilde{y}_{ij} = \frac{x_{ij}}{s_j\sqrt{n}}$ qui servira par la suite à calculer la matrice des coefficients de corrélation entre caractères.

III.3. Axes principaux d'inertie:

On a vu précédemment que les directions principales sont données par les vecteurs propres de la matrice $\mathbf{X}'\mathbf{X}$.

En ACP elles sont données maintenant par les vecteurs propres de la matrice :

$$\mathbf{R} = \tilde{\mathbf{Y}}'\tilde{\mathbf{Y}} \quad \text{ou bien} \quad \mathbf{R} = \mathbf{Y}'\mathbf{D}\mathbf{Y}$$

dont le terme général est $r_{ij} = \frac{\frac{1}{n} \sum_{k=1}^n x_{ki}x_{kj}}{s_i s_j}$ qui n'est autre que le coefficient de corrélation

$$r_{ij} = \frac{s_{ij}}{s_i s_j} \text{ entre les variables } i \text{ et } j.$$

Remarque: Il est possible de déterminer \mathbf{R} d'une autre manière :

\mathbf{V} étant la matrice de variance-covariance

$$\mathbf{V} = \mathbf{X}'\mathbf{D}\mathbf{X} \quad \text{avec:} \quad \begin{cases} \mathbf{X} = [x_{ij}] \text{ (coordonnées centrées)} \\ \mathbf{D} = \begin{pmatrix} 1/n & 0 \\ 0 & 1/n \end{pmatrix} \text{ (matrice diagonale des poids)} \end{cases}$$

On remplace \mathbf{Y} par sa valeur $\mathbf{Y} = \mathbf{X}\mathbf{T}'$ dans \mathbf{R} :

$$\mathbf{R} = \mathbf{Y}'\mathbf{D}\mathbf{Y} = (\mathbf{T}\mathbf{X}')\mathbf{D}(\mathbf{X}\mathbf{T}') \Rightarrow \mathbf{R} = \mathbf{D}_{1/s}\mathbf{V}\mathbf{D}_{1/s}$$

Exemple_1: Tableau individus-caractères. On distingue 3 caractères, indice géographique, revenu, âge

	géographie	revenu	âge
Individu_1	8	30	55
Individu_2	2	6	40
Individu_3	5	15	30
Individu_4	7	22	40

On peut déterminer la matrice des coefficients de corrélation par : $\mathbf{R} = \tilde{\mathbf{Y}}'\tilde{\mathbf{Y}}$ où $\tilde{\mathbf{Y}} = \frac{1}{\sqrt{n}}\mathbf{Y}$ soit

$$\tilde{\mathbf{Y}} = \frac{1}{2}\mathbf{Y} = \begin{pmatrix} 0.5455 & 0.6644 & 0.7702 \\ -0.7638 & -0.6927 & -0.0700 \\ -0.1091 & -0.1838 & -0.6301 \\ 0.3273 & 0.2120 & -0.0700 \end{pmatrix} \quad \text{et}$$

$$\mathbf{R} = \begin{pmatrix} 1.0000 & 0.9810 & 0.5195 \\ 0.9810 & 1.0000 & 0.6612 \\ 0.5195 & 0.6612 & 1.0000 \end{pmatrix}$$

On détermine ensuite les matrices des vecteurs propres et valeur propres, respectivement \mathbf{u} et \mathbf{vap} de \mathbf{R} .

$$\text{On obtient : } \mathbf{u} = \begin{pmatrix} -0.5988 & -0.6515 & -0.4658 \\ -0.6277 & 0.7430 & -0.2322 \\ -0.4974 & -0.1534 & 0.8539 \end{pmatrix} \quad \text{et} \quad \mathbf{vap} = \begin{pmatrix} 2.4598 & 0 & 0 \\ 0 & 0.0034 & 0 \\ 0 & 0 & 0.5368 \end{pmatrix}$$

Les deux plus grandes valeurs propres $\lambda_1=2.4598$ et $\lambda_2=0.5368$, donnent la valeur de la variance expliquée par les 2 axes factoriels choisis pour le plan principal.

Ces axes factoriels sont générés par les 2 vecteurs propres correspondant à ces valeurs propres.

Le taux de variance expliquée par les q premiers axes factoriels (ou inertie sur ces axes), relatif aux q plus grandes valeurs propres est égal à :

$$\tau_q = \frac{\sum_{\alpha=1}^q \lambda_\alpha}{\sum_{\alpha=1}^p \lambda_\alpha} \quad \text{dans notre cas } q=2 \quad \text{et} \quad \tau_q = \frac{2.4598 + 0.5368}{3} = 0.9989 \quad \text{ce qui veut dire que}$$

99.89% de la variance est expliquée par les deux axes factoriels choisis. Le nuage de points est très voisin d'un plan, l'échantillon, dans ce cas, se prête avec une grande précision à une ACP.

Remarque : Lorsque le taux de variance expliquée par 2 axes factoriels choisis est faible, l'échantillon ne se prête pas à priori à une analyse factorielle, les résultats seront donnés avec une précision médiocre.

Il est toutefois possible de déterminer les distances entre points du nuage, représentés par des vecteurs, en s'aidant du produit scalaire entre ces vecteurs.

Lorsque les valeurs propres sont voisines ou égales, il est nécessaire dans certains cas de s'intéresser aux axes factoriels correspondants et donc d'effectuer l'analyse pour un nombre plus grand de vecteurs propres.

$$\text{La matrice des composantes principales est : } \mathbf{C} = \mathbf{Y}\mathbf{u} = \begin{pmatrix} -2.2536 & 0.0403 & 0.4984 \\ 1.8540 & -0.0128 & 0.9137 \\ 0.9882 & 0.0624 & -0.8891 \\ -0.5886 & -0.0899 & -0.5230 \end{pmatrix}$$

On ne retient que les deux composantes principales correspondants aux 2 plus grandes valeurs propres ou aux 2 vecteurs propres \mathbf{u}_1 et \mathbf{u}_3

	C_1	C_3
Individu 1	-2.2536	0.4984
Individu 2	1.8540	0.9137
Individu 3	0.9882	-0.8891
Individu 4	-0.5886	-0.5230

Représentation des individus dans le plan principal :

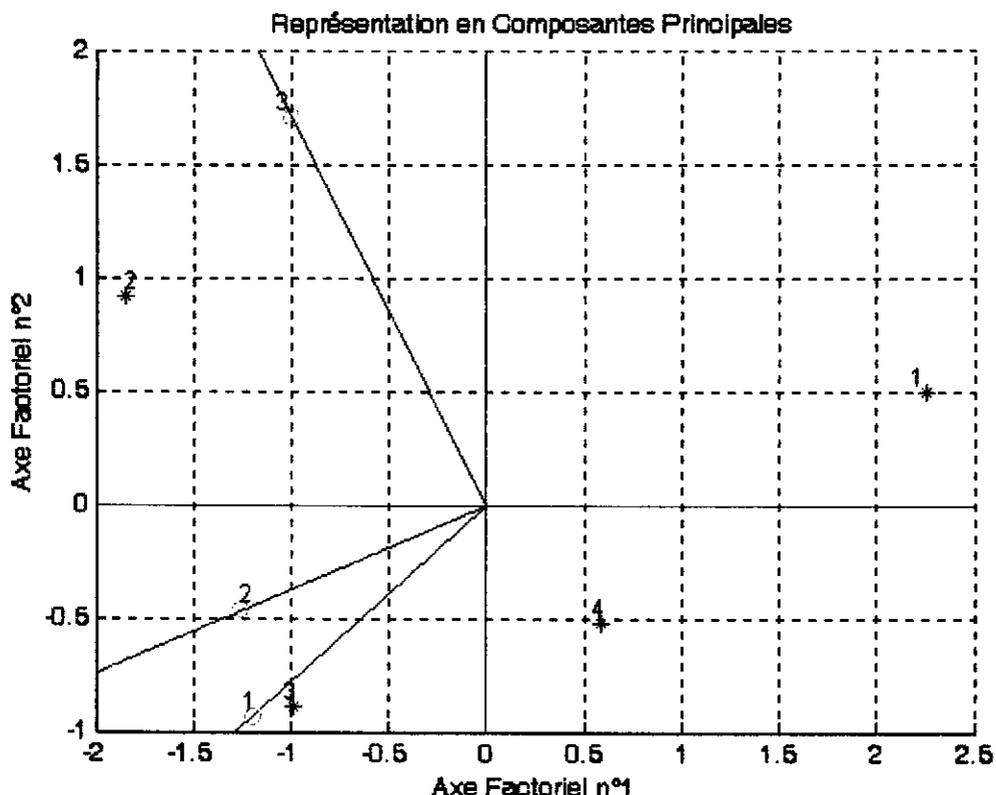


Figure IV.1 Représentation des individus dans le plan principal

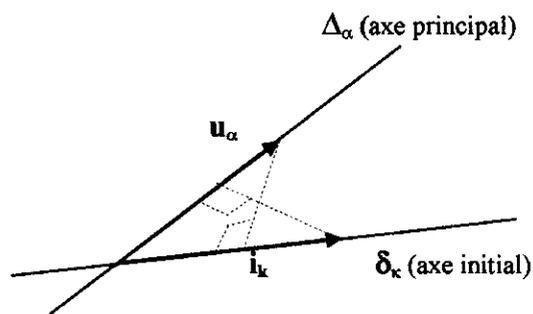
Les points * sont les points individus (au nombre de 4) et les points ° sont les directions des axes initiaux ou tendances (au nombre de 3).

Les directions des axes initiaux ou axes caractères dans le plan principal sont les droites tracées de l'origine aux points °

Pour obtenir dans l'espace des individus, ces axes initiaux, il faut connaître les composantes des vecteurs propres \vec{u}_α

Il est égal au produit scalaire $\vec{u}_\alpha \cdot \vec{i}_k$ qui est commutatif.

Ce dernier représente la composante de l'axe initial (δ_k) sur le $\alpha^{\text{ième}}$ axe principal (Δ_α)



En appelant i la matrice des vecteurs de la base canonique de l'espace des caractères, les directions des tendances ou axes-caractères sont données par le produit $i u$ qui, comme on l'a dit précédemment, est commutatif ou bien par la matrice u^{-1} qui dans ce cas est égale à u' , car les vecteurs sont unitaires. :

Dans notre exemple, la matrice des vecteurs propres est :

$$u = \begin{matrix} & \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{matrix}$$

\vec{i}_1	-0.5988	-0.6515	-0.4658
\vec{i}_2	-0.6277	0.7430	-0.2322
\vec{i}_3	-0.4974	-0.1534	0.8539

En appelant c la matrice des vecteurs directions initiales ou axes caractères :

$$c = u' = u^{-1} =$$

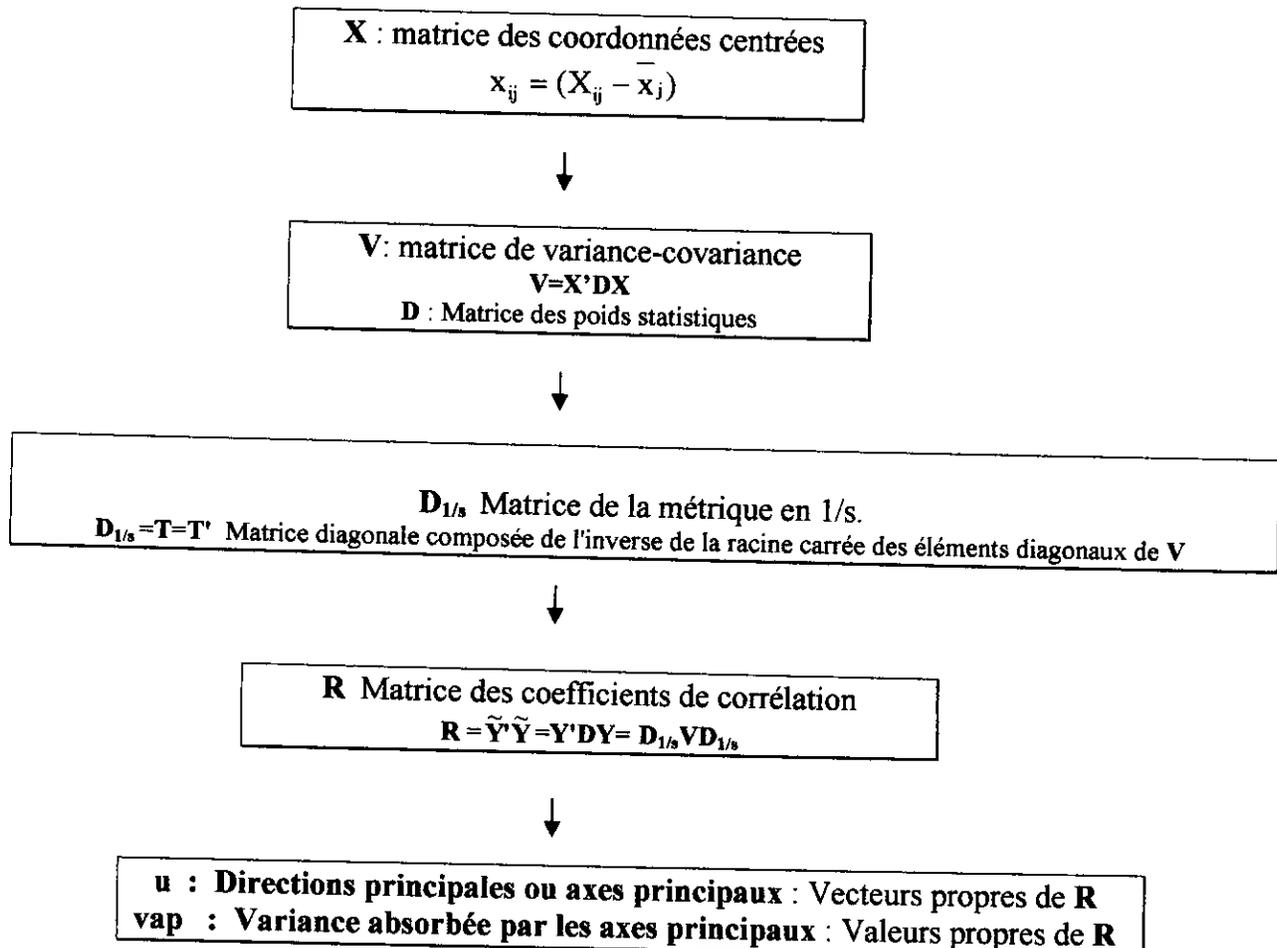
	\vec{i}_1	\vec{i}_2	\vec{i}_3
\vec{u}_1	-0.5988	-0.6277	-0.4974
\vec{u}_2	-0.6515	0.7430	-0.1534
\vec{u}_3	-0.4658	-0.2322	0.8539

On ne retiendra que les composantes sur \vec{u}_1 et \vec{u}_3 , soit :

$$c =$$

	\vec{i}_1	\vec{i}_2	\vec{i}_3
\vec{u}_1	-0.5988	-0.6277	-0.4974
\vec{u}_3	-0.4658	-0.2322	0.8539

Remarque : Dans les cas de notre figure, les coordonnées ont été multipliées par un coefficient d'homothétie de 2 pour une représentation plus facile. *En conclusion:* Dans l'espace R^p des individus on déterminera :





Qualité de la représentation

(variance absorbée par les 2 premiers axes principaux)
C=Yu : Tracé des points individus,
u' : Tracé des directions initiales (tendances ou axes-caractères)

Remarque: La distance entre 2 points M_i et M_j dans R^p est :

$$d^2(i, i') = \frac{1}{n} \sum_{k=1}^p (x_{ik} - x_{i'k})^2 / s_k^2.$$

IV Analyse dans l'espace des caractères R^n :

Pour l'espace des caractères, l'analyse devient alors celle de p points Q_j ($j=1, \dots, p$) dans R^n .
 L'analyse dans R^p et l'analyse dans R^n sont différentes car i et j ne jouent pas des rôles similaires.

Ainsi dans R^p la transformation $x_{ij} = X_{ij} - \bar{x}_j$ qui était une translation de l'origine au centre de gravité du nuage, est maintenant dans R^n une projection parallèle à la 1^{ère} bissectrice.

- le changement d'échelle des axes, c'est-à-dire, la division de chaque coordonnées de R^p par $s_j \sqrt{n}$ devient ici une déformation du nuage qui ramène chacun des points Q_j à la distance 1 du centre de gravité pris comme origine.

En effet :

$$d^2(j, 0) = \frac{1}{n} \sum_{i=1}^n (X_{ij} - \bar{x}_j)^2 / s_j^2 \quad \text{et}$$

$$d^2(j, 0) = \frac{\frac{1}{n} \sum_{i=1}^n (X_{ij} - \bar{x}_j)^2}{s_j^2} = 1$$

Calcul de la distance entre 2 point Q_i et Q_j de R^n :

La distance entre 2 point Q_i et Q_j s'écrit:

$$d^2(j, j') = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_{ij}}{s_j} - \frac{x_{ij'}}{s_{j'}} \right)^2 \quad \text{soit}$$

$$d^2(j, j') = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_{ij}^2}{s_j^2} \right) + \frac{1}{n} \sum_{i=1}^n \left(\frac{x_{ij'}^2}{s_{j'}^2} \right) - \frac{2}{n} \sum_{i=1}^n \left(\frac{x_{ij} x_{ij'}}{s_j s_{j'}} \right)$$

$$= \frac{s_j^2}{s_j^2} + \frac{s_{j'}^2}{s_{j'}^2} - 2r_{jj'} \quad \text{et}$$

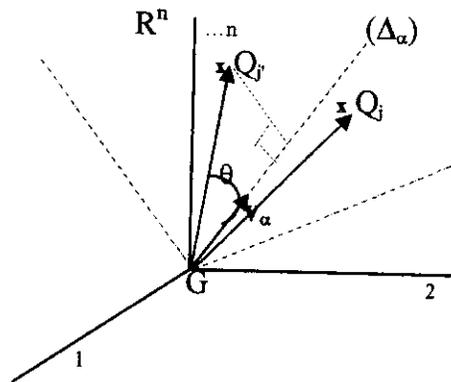
$$d^2(j, j') = 2(1 - r_{jj'})$$

$r_{jj'}$: coefficient de corrélation entre les variables liées à j et j' (caractères)

Les points sont très proches si leur corrélation $r_{jj'}$ est positive, voisine de 1 ou très éloignés si elle est voisine de -1.

Les coordonnées des p point-variables sur l'axe (Δ_α) axe principal sont les composantes du vecteur:

$$Y' v_\alpha = \sqrt{\lambda_\alpha} u$$



Par définition la corrélation entre 2 caractères Q_j et $Q_{j'}$ est

$$r_{jj'} = \frac{\sum_{i=1}^n x_{ij} x_{ij'}}{s_j s_{j'}} \quad (x_{ij} \text{ variables centrées})$$

Si on considère les vecteurs :

$$G\bar{Q}_j = \begin{pmatrix} x_{1j} \\ \vdots \\ x_{ij} \\ \vdots \\ x_{nj} \end{pmatrix} \quad G\bar{Q}_{j'} = \begin{pmatrix} x_{1j'} \\ \vdots \\ x_{ij'} \\ \vdots \\ x_{nj'} \end{pmatrix}$$

$$\|G\bar{Q}_j\|^2 = \sum_{i=1}^n x_{ij}^2 \quad \text{et} \quad \|G\bar{Q}_{j'}\|^2 = \sum_{i=1}^n x_{ij'}^2 \quad \text{soit :}$$

$$\begin{cases} \frac{\|G\bar{Q}_j\|^2}{n} = \frac{1}{n} \sum_{i=1}^n x_{ij}^2 = s_j^2 \quad \text{et} \\ \frac{\|G\bar{Q}_{j'}\|^2}{n} = \frac{1}{n} \sum_{i=1}^n x_{ij'}^2 = s_{j'}^2 \end{cases}$$

Les modules (au carré) sont égaux à la variance du caractère j à $1/n$ près
Si l'on fait le produit scalaire :

$$G\bar{Q}_j \cdot G\bar{Q}_{j'} = \sum_{i=1}^n x_{ij} x_{ij'} = \|G\bar{Q}_j\| \|G\bar{Q}_{j'}\| \cos \theta$$

$$\theta = (G\bar{Q}_j, G\bar{Q}_{j'}) \Rightarrow \cos \theta = \frac{G\bar{Q}_j \cdot G\bar{Q}_{j'}}{\|G\bar{Q}_j\| \|G\bar{Q}_{j'}\|} = \frac{\frac{1}{n} \sum_1^n x_{ij} x_{ij'}}{\frac{1}{n} \left(\sum_i x_{ij}^2 \sum_i x_{ij'}^2 \right)^{1/2}}$$

cos θ est égal au coefficient de corrélation entre les deux caractères j et j'

- Représentation dans l'espace des caractères Rⁿ

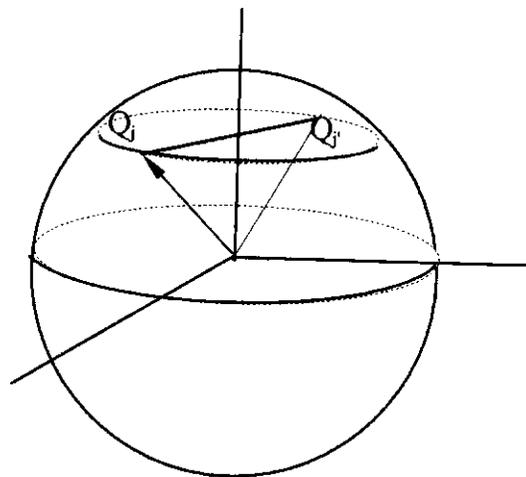
C'est la projection de $G\bar{Q}_j$ sur l'axe (Δ_α) qui donne sa composante sur cet axe ou bien en termes statistiques, représente la corrélation entre le α -ième axe principal et le caractère j

$$Y' v_\alpha = u_\alpha \sqrt{\lambda_\alpha}$$

Interprétation :

La projection de l'hyper-sphère dans R² est appelée le cercle des corrélations.

Pour que 2 caractères soient effectivement corrélés, il faut d'une part qu'ils soient *proches du cercle des corrélation* et d'autres part *proches entre eux* (ou diamétralement opposés lorsque $r_{jj'} \rightarrow -1$)



Exemple 1 : La matrice des coordonnées dans l'espace des caractères es donnée par :

$$RC = u \sqrt{\lambda} \quad (\lambda = \text{vap matrice des valeurs propres} =$$

$$\mathbf{vap} = \begin{pmatrix} 2.4598 & 0 & 0 \\ 0 & 0.0034 & 0 \\ 0 & 0 & 0.5368 \end{pmatrix} \quad \sqrt{\mathbf{vap}} = \begin{pmatrix} 1.5684 & 0 & 0 \\ 0 & 0.0586 & 0 \\ 0 & 0 & 0.7327 \end{pmatrix} \text{ et}$$

$$\mathbf{RC} = \begin{pmatrix} -0.5988 & -0.6515 & -0.4658 \\ -0.6277 & 0.7430 & -0.2322 \\ -0.4974 & -0.1534 & 0.8539 \end{pmatrix} \begin{pmatrix} 1.5684 & 0 & 0 \\ 0 & 0.0586 & 0 \\ 0 & 0 & 0.7327 \end{pmatrix} = \begin{pmatrix} -0.9392 & -0.0382 & -0.3413 \\ -0.9845 & 0.0436 & -0.1701 \\ -0.7801 & -0.0090 & 0.6256 \end{pmatrix}$$

Dont on ne retiendra que la première et 3^{ième} colonne : composantes sur \vec{v}_1 et composantes sur \vec{v}_3
 En divisant par le total des valeurs propres (qui est égal au nombre de variables si les données sont centrées réduites), on obtient cette valeur sous forme de pourcentage. La quantité d'information doit décroître rapidement sinon, cela veut dire que le nuage de points est de forme semblable à une sphère et que donc il n'y a aucun angle de prise de vue qui permette d'avoir une image satisfaisante. Ici on a 80% en cumulé dès le troisième axe.

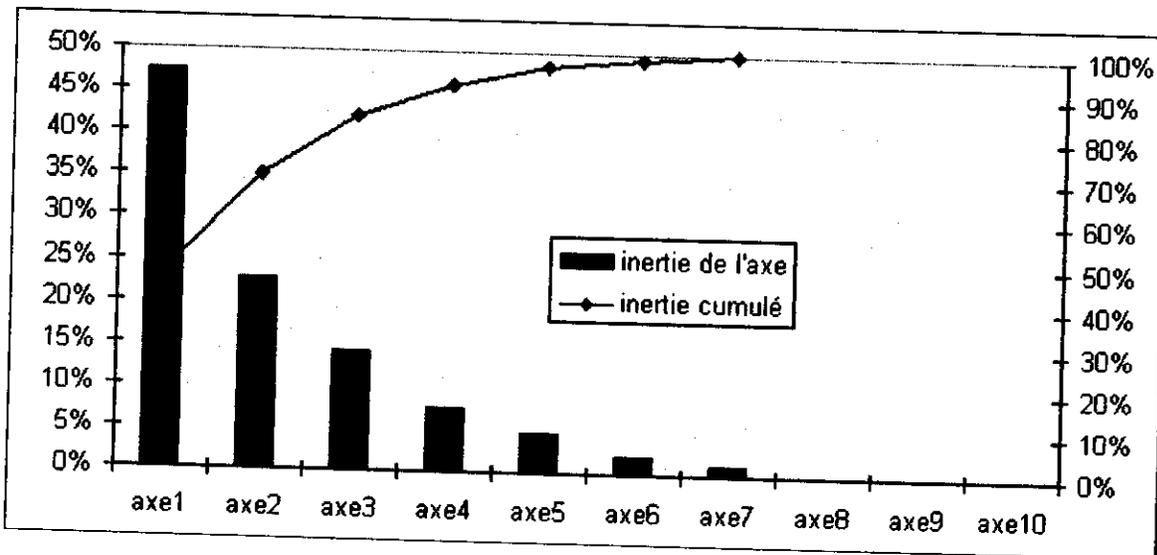


Figure IV.2 Représentation les caractères dans le plan principal

Une fois ces axes calculés il faut en connaître la signification. Pour cela on va regarder les projections des variables sur des plans formés par un couple d'axes factoriels. Les vecteurs propres sont multipliés par les racines des valeurs propres pour être affichés dans un cercle de rayon 1. On va regarder pour chaque axe factoriel quelles sont les variables qui se suivent et s'opposent. Au plus une variable a d'importance au plus elle sera représentée proche du cercle. Dans le cas contraire, une variable n'expliquant pas un axe se trouve vers le centre.

Représentation dans l'espace des caractères :

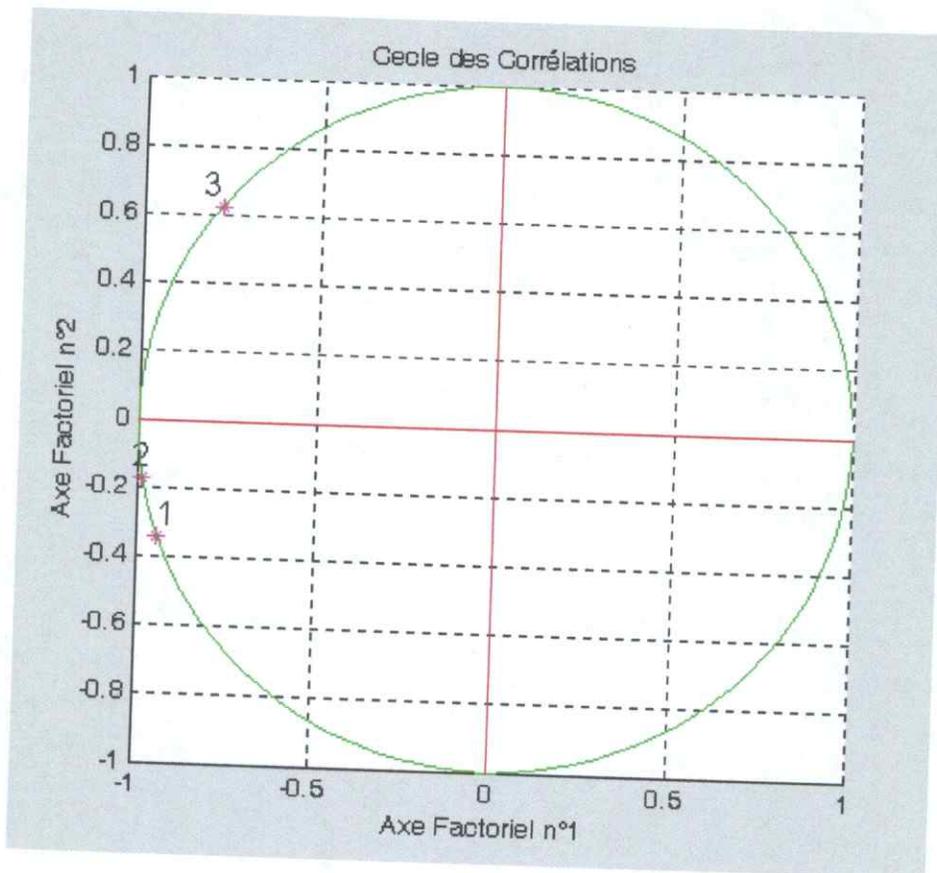


Figure IV.3 Représentation dans l'espace des caractères

En résumé :

Pour l'analyse dans \mathbf{R}^n on calculera :

$$\mathbf{RC} = \mathbf{Y}'\mathbf{v} = \sqrt{\lambda} \mathbf{u}$$

\mathbf{RC} : Matrice des points caractères ($\mathbf{Y}'\mathbf{v}_\alpha = \mathbf{u}_\alpha \sqrt{\lambda_\alpha}$)



Tracé du cercle des corrélations

En conclusion

$\mathbf{R}^p \rightarrow$ classification entre individus, tendances caractérielles.

$\mathbf{R}^n \rightarrow$ corrélation entre caractères

RESUME : ANALYSE EN COMPOSANTES PRINCIPALES

X : matrice des coordonnées centrées

$$X_{ij} = (X_{ij} - \bar{X}_j)$$



V : matrice de variance-covariance $V=X'DX$ **D** : Matrice des poids statistiques



D_{1/s} Matrice de la métrique en 1/s.

$$D_{1/s} = T = T'$$

Matrice diagonale composée de l'inverse de la racine carrée des éléments diagonaux de **V**



R Matrice des coefficients de corrélation

$$R = \tilde{Y}' \tilde{Y} = Y'DY = D_{1/s} V D_{1/s}$$



Directions principales ou axes principaux : Vecteurs propres de **R**
Variance absorbée par les axes principaux : Valeurs propres de **R**



Qualité de la représentation

(Variance absorbée par les 2 premiers axes principaux)

Tracé des points individus, directions principales et initiales (tendances)

Conclusion :

L'objectif dans ce chapitre reste donc, de réduire la base d'attributs pour la recherche des meilleurs attributs qui caractérisent les images des défauts de soudure et qui répondent aux deux critères suivants : discrimination maximale et redondance minimale. Ces meilleurs attributs ont l'objectif de maintenir et/ou améliorer la performance du classifieur utilisé par le système de reconnaissance, pour minimiser le nombre de neurones d'entrées au réseau tout en améliorant la performance du système de reconnaissance. Ceci va permettre un gain de temps de calcul considérable si un grand nombre d'entrées est supprimé.

CHAPITRE V

**LES RESEAUX DE NEURONES ARTIFICIELS ET
CLASSIFICATION DES DEFAUTS DE SOUDURE
DANS LES FILMS DE RADIOGRAPHIE**

I. Introduction :

Les méthodes de reconnaissance à base de réseaux de neurones ont été étudiées depuis plusieurs années dans le but de réaliser des performances proche de celles observées chez l'humain. Ces réseaux de neurones sont composés de plusieurs éléments (ou cellules) de calcul opérant en parallèle et arrangés à la manière des réseaux de neurones biologiques. Lippmann [8] a précisé que la force des réseaux de neurones réside dans leur capacité à générer une région de décision de forme quelconque, requise par un algorithme de classification, au prix de l'intégration de couches de cellules supplémentaires dans le réseau de type perceptron multicouche avec un apprentissage par rétropropagation du gradient de l'erreur. Ainsi, nous utilisons les réseaux de neurones à rétro propagation du gradient à la classification des défauts de soudure dans les images de radiographie industrielle ayant comme traits descriptifs les attributs géométriques invariants.

Les techniques neuronales sont par nature des méthodes de classification paramétriques puisque la phase d'apprentissage consiste à estimer les paramètres qui permettront de choisir la fonction de discrimination dans un ensemble de fonctions possibles. Toutefois, ces fonctions de discrimination sont tellement complexes qu'on peut plutôt rapprocher les techniques neuronales à des méthodes de classification non paramétriques.

II Présentation des réseaux de neurones :

Un réseau se distingue en général par le type de neurone formel qu'il utilise, la règle d'apprentissage qui le décrit et l'architecture définissant les interconnexions entre neurones. Le neurone formel, qui peut être considéré comme une modélisation élémentaire de neurone réel, est un automate possédant n entrées réelles z_1, \dots, z_n et dont le traitement consiste à affecter à sa sortie o le résultat d'une fonction d'activation f de la somme pondérée de ses entrées.

$$o = f\left(\sum_{i=1}^n w_i z_i\right) = f(\text{net}_i)$$

Equation V.1

où les w_i sont les coefficients ou poids synaptiques associés aux entrées z_i . La fonction d'activation f est généralement de type sigmoïde. La règle d'apprentissage tente de déterminer les valeurs (optimales) des coefficients synaptiques à l'aide d'exemples permettant de minimiser une fonction de coût d'erreur définie entre la sortie effective du réseau et la sortie désirée. Enfin l'architecture du réseau définit les interconnexions entre neurones et les

contraintes liant les coefficients synaptiques. L'architecture peut être fixée au départ, généralement de façon heuristique ou construite progressivement pendant la phase d'apprentissage. L'architecture adoptée et la nature des connexions entre neurones définissent la règle de propagation dans le réseau encore appelée la dynamique du réseau.

Les réseaux à couches sont les modèles connexionnistes les plus couramment utilisés. Leur architecture, organisée en couches successives, comprend une couche d'entrée, une couche de sortie et une ou plusieurs couches intermédiaires appelées couches cachées car elles ne sont pas vues de l'extérieur. Chaque couche est composée d'un certain nombre de neurones. Les connexions sont établies entre les neurones appartenant à des couches successives mais les neurones d'une même couche ne peuvent pas communiquer entre eux dans le cas des réseaux à couches.

Parmi les plus simples des réseaux à couches, on cite le réseau mono couche appelé perceptron. Historiquement, le perceptron est parmi les premiers réseaux efficaces qui a été proposé et étudié en détail. Il est composé uniquement d'une couche d'entrée et d'une couche de sortie (voir figure 1). La couche d'entrée comporte autant de neurones que le vecteur de caractéristiques (primitives) a de composantes. Les connexions entre ces deux couches sont modifiables et bidirectionnelles. C'est la couche de sortie qui remplit la tâche de classification et les unités de la couche de sortie réalisent une fonction à seuil. Cette fonction d'activation réalisée par le neurone formel est un échelon qui vaut 1 si la somme pondérée de ses entrées est supérieure au seuil et 0 sinon. Chaque neurone sépare donc linéairement l'espace des entrées en deux classes. Les connexions sont modifiées par un apprentissage supervisé selon le principe de correction d'erreurs : si une unité de la couche de sortie n'est pas dans l'état désiré, toutes les liaisons que cette unité réalise avec les unités de la couche d'entrée sont augmentées ou diminuées selon le type d'action que l'unité de la couche d'entrée réalisait sur l'unité de sortie considérée. Autrement dit, la règle d'apprentissage fixe le poids des connexions selon le principe de coût d'erreurs. A partir d'exemples présentés sur la couche d'entrée du perceptron, les poids associés aux connexions de chaque neurones de la couche de sortie sont modifiés jusqu'à obtenir la sortie désirée de ce neurone. Ainsi, m neurones binaires sur la couche de sortie permettent de séparer linéairement $2m$ classes.

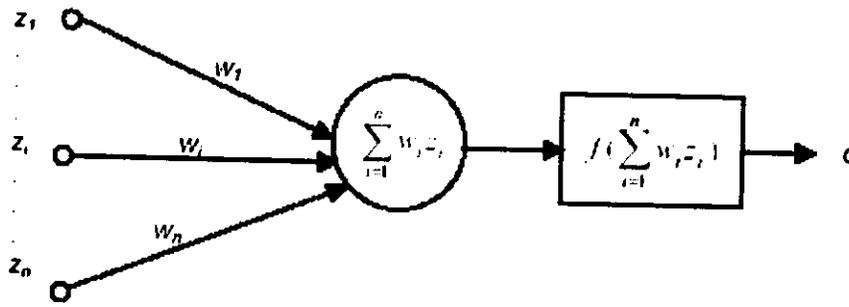


Figure V.1 Modèle d'un réseau de neurone, perceptron [8]

Les méthodes d'apprentissage sont divisées en deux catégories : l'apprentissage *avec supervision* et l'apprentissage *sans supervision* [9]. L'apprentissage supervisé consiste à ajuster par interpolation, approximation ou optimisation les poids du réseau de neurones en comparant la sortie obtenue et la sortie désirée de façon à minimiser l'erreur. Les algorithmes d'apprentissage permettent de déterminer automatiquement certains paramètres libres du réseau (par exemple : poids) qui correspondent en fait à des paramètres permettant de définir les frontières des classes. Le réseau de neurones génère une ou plusieurs valeurs de sorties basées sur la performance de calcul des valeurs d'entrées. Les résultats de ces calculs dépendent des valeurs de poids assignées à chaque entrée obtenues après la phase d'apprentissage.

Pour faire un lien entre les réseaux de neurones et d'autres types de classifieurs, rappelons les rapports étroits qui existent entre les classifieurs bayésiens et certains modèles connexionnistes. Il a été démontré dans la littérature que pour un problème de classification à m classes, un perceptron multicouches avec m neurones de sorties utilisant la méthode de rétropropagation du gradient de l'erreur et ayant comme sortie désirée un neurone à 1 indiquant la classe correcte et les autres à 0, réalisera une estimation des probabilités *a posteriori*. Plus formellement, pour attribuer un vecteur Z à une classe parmi m , un classifieur bayésien estimera les probabilités *a posteriori* $P(C_i|Z)$ pour chaque classe C_i , $i=1, \dots, m$ et assignera au vecteur Z la classe de probabilité maximale. Contrairement à la méthode bayésienne qui estime les probabilités *a posteriori* à partir des probabilités *a priori*, le calcul par un perceptron sera direct sans passer par les probabilités *a priori*. On appelle probabilité *a posteriori* la probabilité conditionnelle $P(C_i|Z)$ de la classe C_i connaissant Z . En d'autres termes, quand on a une entrée Z , $P(C_i|Z)$ représente la probabilité que Z appartienne à la classe C_i . La probabilité conditionnelle $P(Z|C_i)$ représente la probabilité d'avoir Z comme entrée quand on sait que l'on est dans la classe C_i .

II.1 Perceptron multicouche:

L'introduction de couches intermédiaires dans le réseau MLP permet de résoudre des problèmes plus complexes que la simple séparation linéaire. Lorsqu'il existe au moins une couche cachée, les états internes du réseau ne peuvent plus être donnés directement par les exemples et les sorties désirées puisque les sorties des neurones appartenant aux couches intermédiaires sont inconnues. La figure 2 représente un réseau de neurones multicouche avec comme entrée le vecteur de primitives et en sortie les classes où seront classées les formes. Pour déterminer le nombre de couches cachées dans un réseau cela dépendra du problème à résoudre (comment les classes des formes sont-elles séparées?).

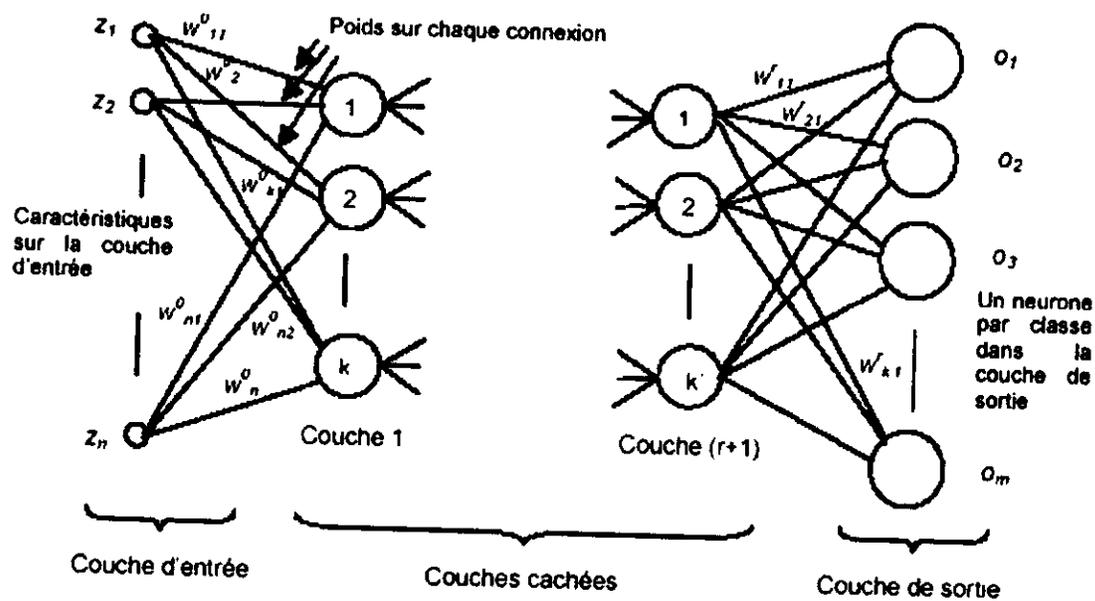


Figure V.2 Réseau de neurones multicouches [10]

Plusieurs types de réseaux de neurones multicouches ont été développés. Nous focalisons notre étude sur le type MLP. Dans le processus de la construction de ce type de classifieur, l'apprentissage a comme objectif d'adapter les poids des connexions entre les neurones de sorte que le réseau donne en sortie la classe d'appartenance des formes qui lui sont proposées en entrée. Ce qui revient à minimiser l'erreur commise par le réseau sur l'ensemble de formes de la base d'apprentissage. Pour une forme de la classe C_i , l'état du neurone O_i de la couche de sortie du MLP doit être proche de la valeur maximale et tous les autres proches de la valeur minimale. Ce problème de minimisation de l'erreur a été résolu par des méthodes de *rétropropagation du gradient d'erreur* [10]. Toutefois, ces méthodes ne garantissent pas l'obtention du minimum global de la fonction à optimiser (blocage dans un minimum local). Le terme «*rétropropagation*» est utilisé pour décrire l'apprentissage du réseau de neurones de

type MLP utilisant la descente du gradient appliquée à la fonction de la somme des erreurs quadratiques [11]. Bishop [12] et Mangalagiu [13] ont démontré qu'un réseau de type MLP à une couche cachée peut estimer n'importe quelle fonction dans R^n avec une précision arbitraire. Ainsi, le MLP est capable d'estimer des hyper-surfaces discriminantes très complexes. Les fonctions estimées sont soit linéaires ou non linéaires. Elles sont appelées fonctions d'activation. La fonction non linéaire que nous utiliserons est la fonction *sigmoïde*. Par contre la difficulté majeure rencontrée lors de l'utilisation de ce type de classifieur consiste à déterminer le nombre de couches cachées, le nombre de neurones dans chacune des couches et les poids des connexions entre les différentes couches (pente de la sigmoïde, taux d'apprentissage...). De ce fait, pour une application donnée, la construction du classifieur de type MLP utilise des règles empiriques et nécessite un certain nombre d'essais afin d'obtenir des performances de généralisation intéressantes.

Pour déterminer les poids de toutes les connexions du réseau, l'utilisation des algorithmes d'apprentissage est obligatoire. L'objectif des algorithmes d'apprentissage est de minimiser l'erreur de décision effectuée par le RNA en ajustant les poids à chaque présentation d'un vecteur d'entraînement. Pour ce qui est de l'ajustement des poids à une étape donnée de la phase d'apprentissage, l'erreur à minimiser est habituellement celle produite lors de l'application d'un vecteur de l'ensemble d'apprentissage à l'entrée du réseau. Pour ce qui est de l'évaluation de la qualité d'apprentissage du réseau, l'erreur cumulée par tous les vecteurs de l'ensemble d'entraînement est évaluée. Cette erreur cumulée est calculée pour tous les cycles de la phase d'entraînement et est définie à partir de l'erreur quadratique. Cette mesure de l'erreur illustre la précision obtenue après P cycles d'apprentissage. De même l'ajout d'un momentum permettra d'accélérer la vitesse de convergence de l'algorithme d'apprentissage par la rétropropagation des erreurs. Cela consiste à ajouter à la valeur courante d'ajustement des poids une fraction de l'ajustement précédent.

En résumé, nous utiliserons pour l'apprentissage du réseau l'algorithme de rétropropagation avec minimisation du gradient d'erreur qui est défini par les étapes suivantes [10, 11] :

1. Initialiser les poids à des petites valeurs et les seuils du réseau.
2. Insérer à l'entrée du réseau une observation (exemple) de la base de données en forme de vecteur de caractéristiques, puis calculer sa valeur d'activation et sa fonction d'activation en utilisant les formules (V.3) et (V.4).
3. Évaluer le signal d'erreur des sorties du réseau en utilisant la formule (V.5).
4. Ajuster les poids en utilisant la formule (V.2).
5. Évaluer le signal d'erreur pour chaque couche cachée en utilisant la formule (V.6).

6. Ajuster les poids de la couche cachée en utilisant la formule (V.2).

7. Répéter les étapes 2 à 6 pour l'ensemble des observations de la base d'apprentissage tant que le critère d'arrêt n'a pas été atteint. Il existe plusieurs critères d'arrêts. Ces critères peuvent être combinés entre eux. Le premier critère est basé sur l'amplitude du gradient de la fonction d'activation, puisque par définition le gradient sera à zéro au minimum. L'apprentissage du réseau du type MLP utilise la technique de recherche du gradient pour déterminer les poids du réseau.

Le second critère d'arrêt est de fixer un seuil que l'erreur quadratique ne doit pas dépasser. Toutefois ceci exige une connaissance préalable de la valeur minimale de l'erreur qui n'est pas toujours disponible. Dans le domaine de la reconnaissance de formes, il suffit de s'arrêter lorsque tous les objets sont correctement classés. Le troisième critère consiste à déterminer un nombre fixe de cycles à atteindre.

Finalement, la méthode du *cross-validation* [10] peut être utilisée pour surveiller l'évolution de l'apprentissage appelée la performance du système en généralisation. Toutefois l'algorithme s'arrête lorsqu'il n'y a plus d'amélioration de la performance du système de reconnaissance. Cette méthode fonctionne sur deux bases de données : la base d'apprentissage pour entraîner le réseau et la base validation pour mesurer la performance en généralisation du réseau de neurones. Durant la phase d'apprentissage, la performance du réseau sur la base d'apprentissage continue à s'améliorer mais sur la base de validation à un moment donné une chute de sa performance peut être observée. Si cette dégradation de la performance sur la base de validation existe alors on dit que le réseau est en sur-apprentissage. Si ce cas de surapprentissage est rencontré alors l'algorithme d'apprentissage sera arrêté à ce moment. Les trois premiers critères sont sensibles aux choix des paramètres (par exemple : le nombre de nœuds dans la couche cachée, le seuil d'erreur, ...) et si le choix n'est pas bon alors les résultats obtenus seront mauvais ou le temps de calcul de la performance du système de reconnaissance sera plus lent (par exemple définir un grand nombre de nœuds dans la couche cachée). Cependant, la méthode du crossvalidation n'a pas ce genre de problème.

En général, les formules utilisées par cet algorithme sont :

pour l'ajustement de poids entre le nœud j (sortie ou cachée) et le nœud i (cachée ou entrée)

$$\Delta w_{ij} = \eta \delta_j o_i$$

Equation V.2

où η est la valeur de la constante d'apprentissage. En général, $0.1 < \eta < 0.9$

et a_i est la valeur d'activation du neurone i tel que $O_i = f(net_i)$

$$f(net_i) = \frac{1}{1 + e^{-net_i}}$$

Equation V.3

avec $f(net_i)$ est la fonction d'activation. La fonction utilisée dans le cas de notre projet est la fonction sigmoïde définie

$$f(net_i) = \frac{1}{1 + e^{-net_i}}$$

Equation V.4

pour le calcul du signal d'erreur du neurone j de la couche de sortie

$$\delta_j = (d_j - o_j) o_j (1 - o_j)$$

Equation V.5

avec d_j est la valeur désirée du neurone j .

. Pour le calcul du signal d'erreur du neurone j de la couche cachée avec k nœuds

$$\delta_j = o_j (1 - o_j) \sum_k w_{jk} \delta_k$$

Equation V.6

Ces formules ont été dérivées de la formule de calcul de l'erreur quadratique de l'ensemble de la base d'apprentissage définie comme suit

$$L_E = \frac{1}{2} \sum_p \left(\sum_k (d_{pk} - o_{pk})^2 \right)$$

Equation V.7

où p est l'indice d'un exemple de la base et k est l'indice du nœud de sortie. L'objectif est de minimiser cette erreur.

- Lors de l'ajout du momentum, la formule (V.2) devient alors :

$$\Delta v_j(n+1) = \eta \delta_j v_j + \alpha \Delta w_j(n)$$

Equation V.8

où α est le momentum tels que $0.1 < \alpha < 0.8$ et $\Delta w_{ij}(n)$ représente l'ajustement à l'étape précédente.

L'inconvénient des réseaux à couches est le manque d'éléments théoriques permettant de relier d'une part le nombre de couches cachées et le nombre de neurones par couche et d'autre part le type et la complexité du problème à traiter. Il existe des heuristiques pour déterminer le nombre de neurones dans une couche cachée. La formule (V.9) nous montre une heuristique pour déterminer ce nombre

$$N_c = \left(\frac{N_e + N_s}{2} \right) + 1$$

Equation V.9

où N_e est le nombre d'entrées du MLP.

N_s est le nombre de neurones en sortie.

N_c est le nombre entier de neurones cachés.

Pour déterminer le nombre maximal de neurones cachés, la formule (V.10) présente ce nombre

$$N_c \leq (2 \times N_e) + 2$$

Equation V.10

Pour définir le nombre de couches cachées et de neurones par couche, le concepteur doit effectuer un grand nombre d'expériences. Par exemple, on fait varier la taille du réseau puis on effectue un apprentissage complet pour chaque taille et enfin on choisit la structure qui conduit aux meilleurs résultats.

III. Travaux Futur :

III.1 Construction des classes des défauts de soudures :

La classification proposée consiste à affecter les principaux types des défauts de soudure à quatre classes ou catégories, en se basant sur les caractéristiques des formes des défauts

couramment rencontrés dans l'annexe C. Les attributs utilisés dans cette classification sont des attributs invariants aux transformations géométriques usuelles. De ce fait, de tels attributs ne caractérisent que la forme des défauts et ne peuvent discriminer les défauts sur la base d'autres paramètres. En examinant un nombre considérable des films de radiographie et selon les caractéristiques morphologiques des défauts, nous pouvons déduire quatre classes principales :

Première Classe :

Elle englobe les défauts dont la forme est allongée, fine et plus ou moins rectiligne. On peut associer à cette catégorie les fissures, les caniveaux, les manques de fusion latéraux, etc. On peut associer à cette catégorie les défauts de soudures : de D_F_1 à D_F_8. (Voir l'annexe C)

Deuxième Classe :

Elle englobe les défauts dont la forme est allongée, unie et rectangulaire. Elle peut comporter le manque de pénétration, les soufflures allongées, etc. On peut associer à cette catégorie les défauts : de D_F_9 à D2_F_13.

Troisième Classe :

Elle englobe les défauts de forme sphéroïdale. Elle concerne les porosités, les inclusions de tungstène, etc. On peut associer à cette catégorie les défauts : de D_F_14 à D2_F_17.

Quatrième Classe :

Elle englobe les défauts dont la forme est biscornue, non allongée et non sphéroïdale. Cette catégorie peut inclure les inclusions de solide, les inclusions de laitier, les soufflures vermiculaires, etc. On peut associer à cette catégorie les défauts : de D_F_18 à D_F_23.

L'architecture du réseau de neurones utilisée est illustrée comme suit :

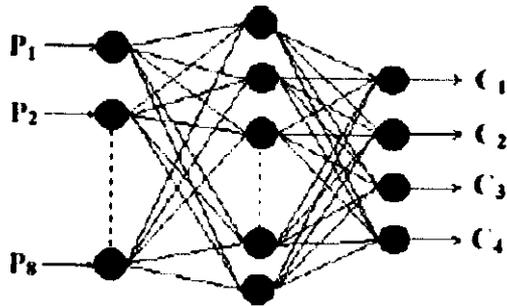


Fig. V.3 : Configuration du RNA utilisé

Conclusion

Ce chapitre nous a permis d'avoir un aperçu général sur les réseaux de neurones. Notre étude s'est intéressée principalement aux réseaux de neurones de type MLP. Nous avons décrit les étapes de l'apprentissage du réseau et les différents critères d'arrêt de cet algorithme. L'objectif de notre projet est de minimiser le nombre de neurones d'entrées au réseau tout en améliorant la performance du système de reconnaissance. Ceci va permettre un gain de temps de calcul considérable si un grand nombre d'entrées est supprimé. Pour se faire, nous utilisons les éléments d'une théorie de l'optimisation.



CHAPITRE VI

**LA CONCEPTION ET LA
REALISATION**

Introduction :

Nous avons utilisé le langage de modélisation et spécification UML (voir l'annexe A) pour modéliser notre système. Comme UML n'impose pas de méthode de travail particulière, il peut donc être intégré à n'importe quel processus de développement logiciel de manière transparente. UML est une sorte de boîte à outil, qui permet d'améliorer progressivement nos méthodes de travail. Intégrer UML par étapes dans un processus, de manière pratique, est tout à fait possible.

Notre choix est basé sur un processus de modélisation qui s'appelle RUP (Rational Unified Process). (voir l'annexe A)

L'architecture que nous avons utilisés permet de couvrir le cycle de vie d'un logiciel depuis l'analyse du besoin jusqu'au test. Cette notation est d'une très grande richesse. Elle piloté par les cas d'utilisation, centré sur l'architecture 4+1, itératif et incrémental [5].

Elle permet de couvrir toutes les phases du développement :

- Les besoins des utilisateurs du futur système, exprimés à l'aide de cas d'utilisation.
- La spécification complète du système, sous forme de diagrammes couvrant les parties statiques et dynamiques du système.
- La conception détaillée, jusqu'à un niveau très proche du code en langage objet de l'implémentation.
- Les suites de tests permettant de s'assurer qu'une implémentation candidate est effectivement conforme à la spécification élaborée.

I La Vue Cas d'utilisation (Diagramme de cas d'utilisation):

Le but de cette vue est de comprendre et structurer les besoins du client. Il faut d'abord clarifier, filtrer et organiser les besoins.

Une fois identifiés et structurés, ces besoins : définissent le contour du système à modéliser (ils précisent le but à atteindre), permettent d'identifier les fonctionnalités principales (critiques) du système.

Les fonctionnalités du système sont décrites comme un ensemble de cas d'utilisation.

- Chaque cas représente un flot spécifique d'événements vers le système.
- La description du cas d'utilisation définit ce qui arrive dans le système lors de sa réalisation.

I.1 Définition des acteurs :

Interpréteur : est l'acteur unique et principal du système qui est chargé d'effectuer plusieurs fonctions :

- établit une Image de défaut par la radiographie (scanner les joins de soudures).
- Prendre la décision sur l'identification des défauts en suivant la chaîne de contrôle

(voir chapitre I).

I.2 Cas d'utilisation de haut niveau : Il décrit les interactions essentielles entre le système et les utilisateurs.

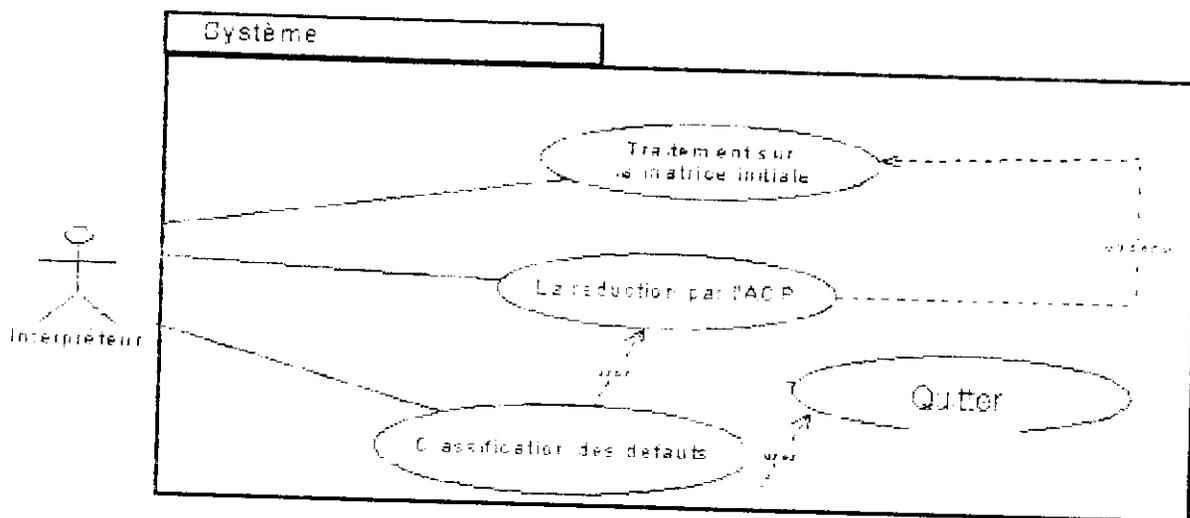


Figure VI.1 : Diagramme de haut niveau du système.



L'interpréteur peut interroger au système via trois cas d'utilisation:

Le cas d'utilisation "la réduction par l'ACP" utilise le cas "traitement sur la matrice initiale", ainsi que "la classification des défauts" utilise "l'ACP".

I.3 Cas d'utilisation de bas niveau:

Les cas d'utilisation de bas niveau sont regroupés selon les paquetages suivants :

- 1- Paquetage traitement sur la matrice initiale.
- 2- Paquetage réduction la matrice par l'ACP¹

¹ Analyse en composantes principales.

3- Paquetage classification des défauts par réseaux de neurones

3.1 Paquetage traitement sur la matrice initiale:

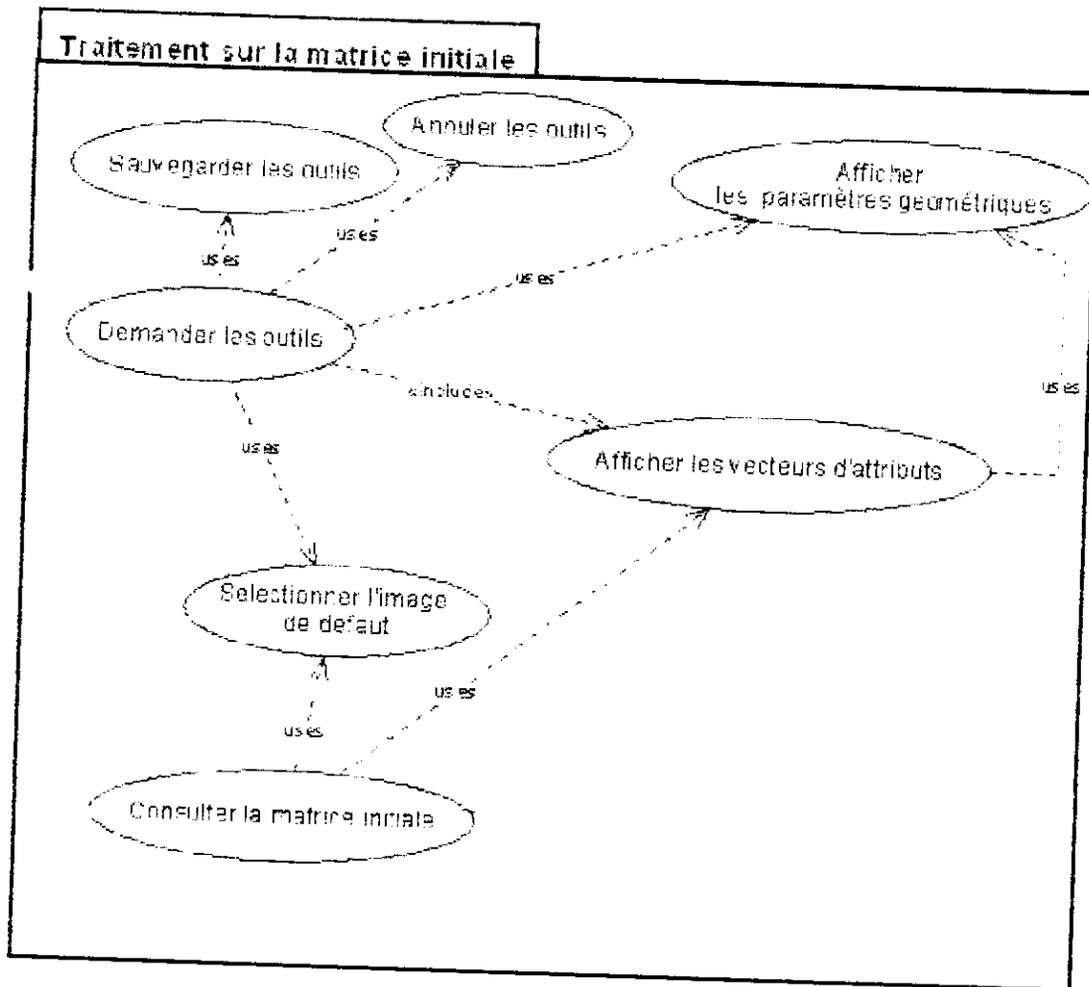


Figure VI.2 : Cas d'utilisation traitement sur la matrice initiale.

Le cas d'utilisation "demander les outils" inclus les mesures et le vecteur d'attribut.

Le cas d'utilisation "Afficher les mesures" extend :

- "Afficher les paramètres géométriques"

- "Afficher les paramètres topologiques"

Les cas d'utilisation "consulter la matrice" et "demander les outils" utilise "sélectionner l'image de défaut"

I.3.2 Paquetage réduction la matrice par l'ACP :

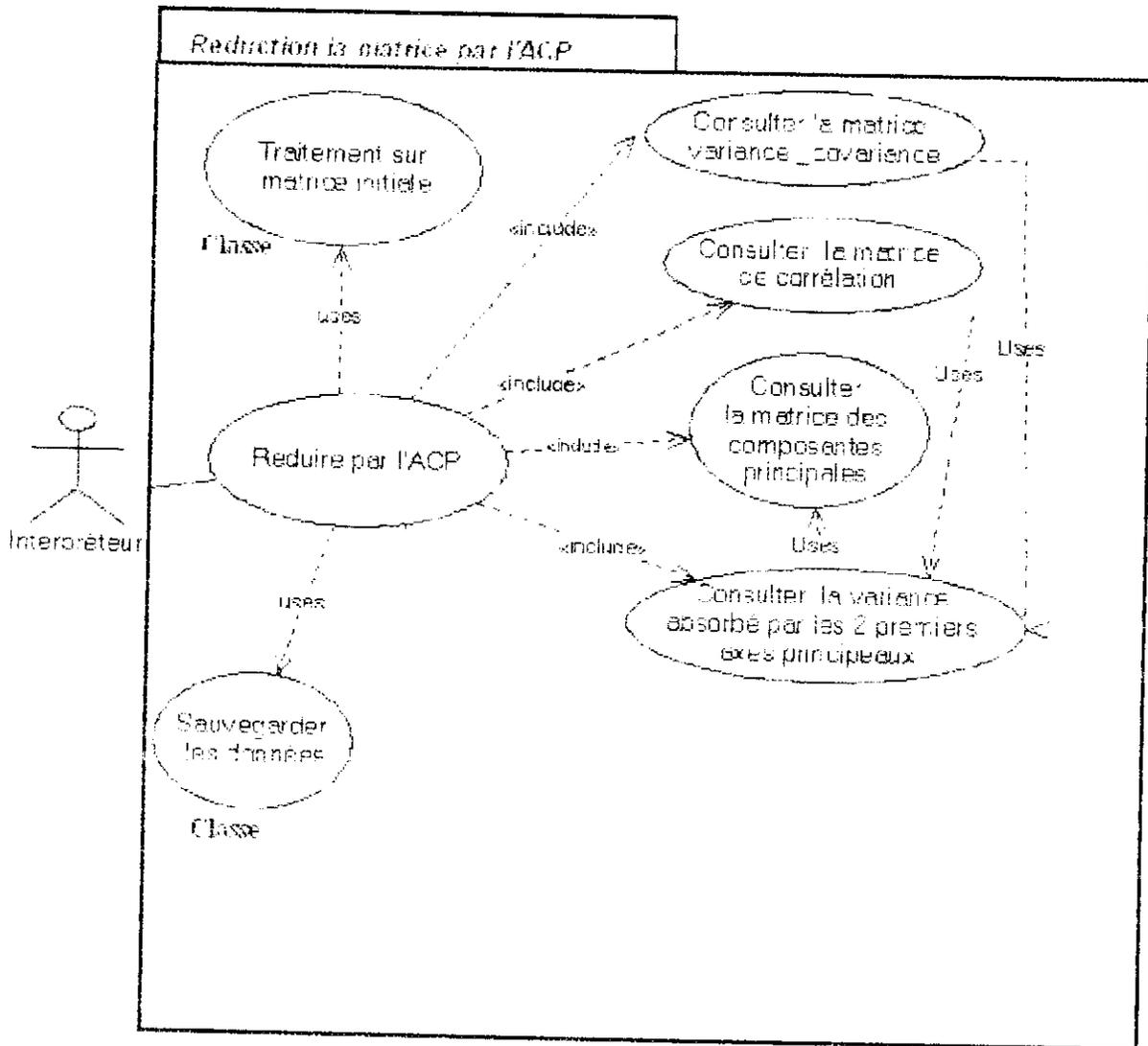


Figure VI.3 : Cas d'utilisation réduction la matrice par l'ACP

- l'utilisateur (interpréteur) peut avoir la matrice initiale et la matrice réduite.
 le cas d'utilisation " Réduction par ACP" utilise le cas d'utilisation "traitement sur la matrice initiale" pour obtenir une matrice réduite par la méthode ACP. Cette dernière est une technique mathématique permettant de réduire un système complexe de corrélations en un plus petit nombre de dimensions géométriquement.

" Réduction par ACP" inclus:

- Consulter la matrice variance_covariance.
- Consulter les axes principaux d'inertie.
- Consulter les composantes principales.

I.3.1 Paquetage classification des défauts par réseaux de neurones artificiels :

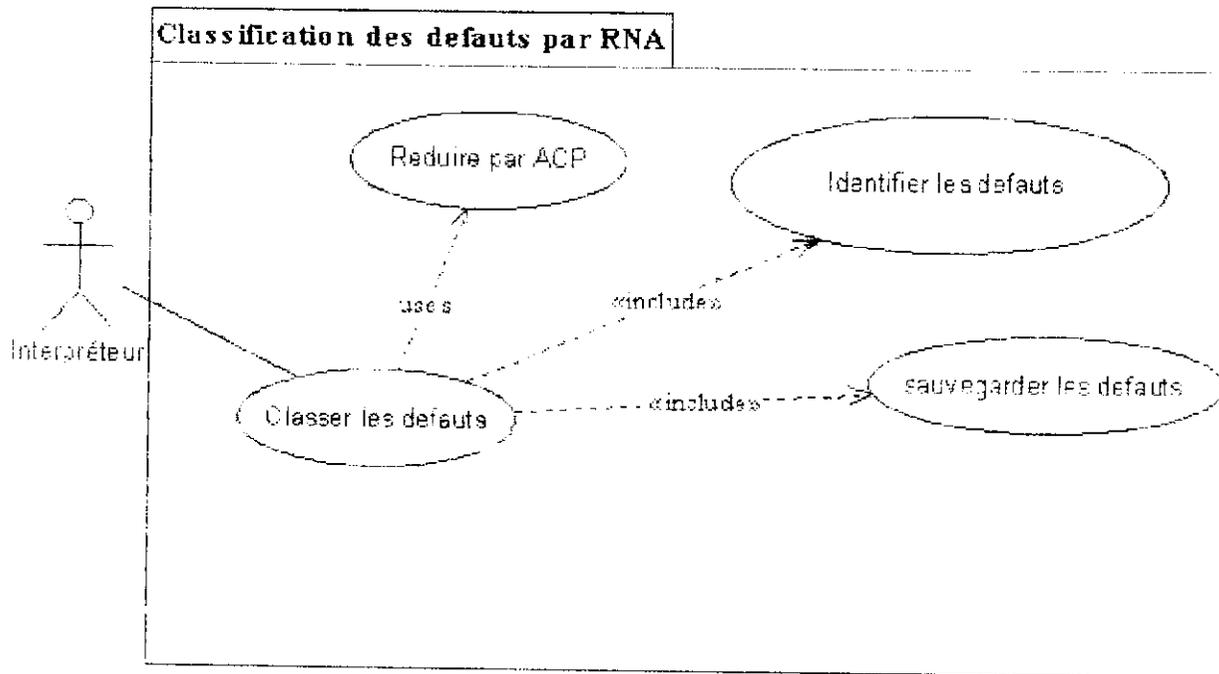


Figure VI.4 : Cas d'utilisation classification des défauts par réseaux de neurones

Le cas d'utilisation "Classifier les défauts" aide l'interpréteur de prendre la décision d'identification des défauts, il inclus "identifier les défauts", "sauvegarder les défauts".

I.4 Diagramme de collaboration:

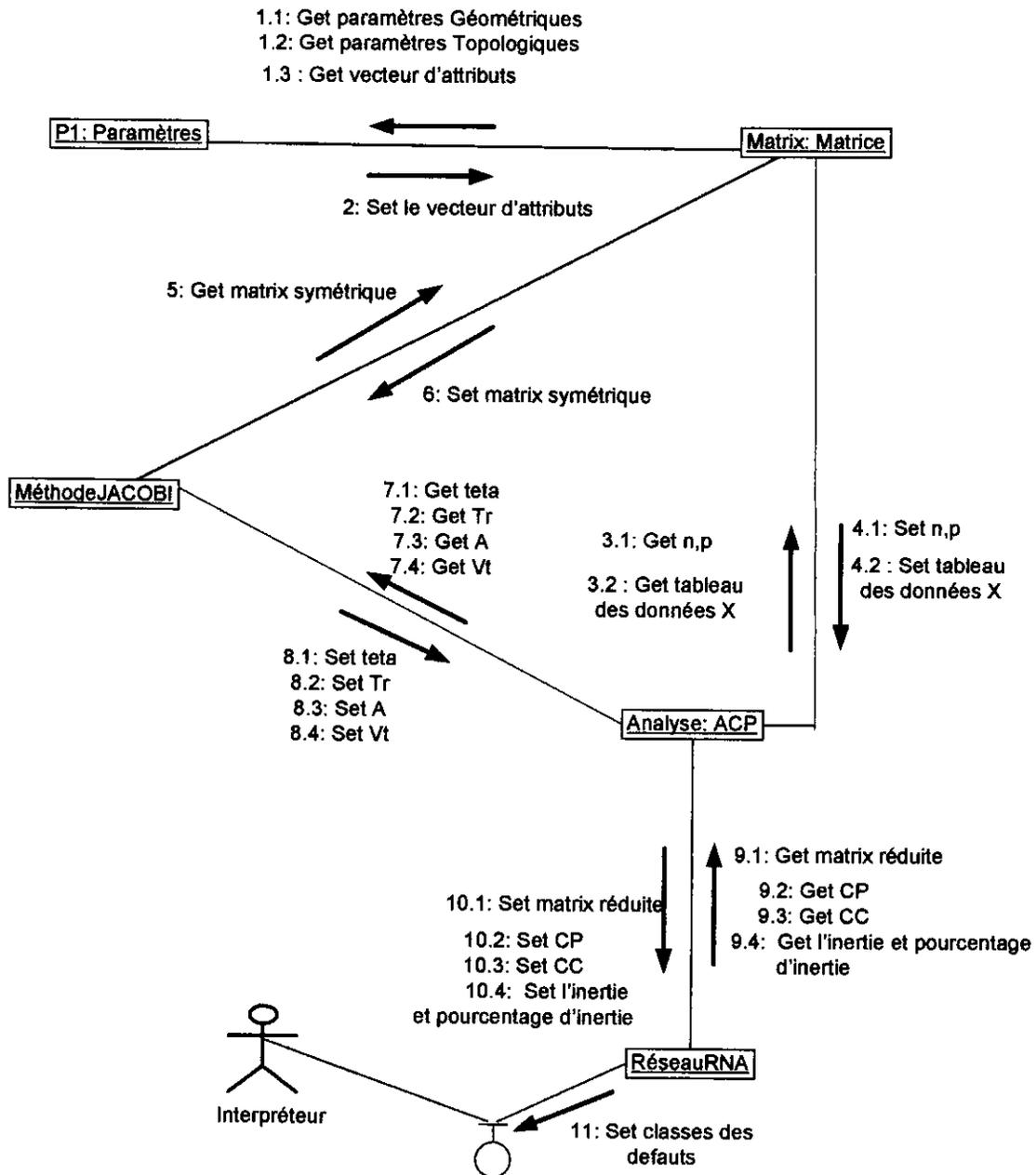


Figure VI.5 : Diagramme de collaboration

II La Vue Logique (Diagramme de classes):

Afin de permettre au lecteur de bien comprendre notre système, dans cette section, nous utilisons la vue logique qui est représentée par le diagramme de classes. Nous présentons notre système par un ensemble de paquetage chacun correspond à un sous système. Ces paquetages sont répond aux besoins des cas d'utilisation. Ensuite nous allons décomposer et détailler chaque paquetage de classes à part.

II.1 Diagramme de classe :

Le diagramme de classe nous a permet de modéliser, d'une manière logique, les relations qui existent entre l'ensemble de classes. Il développe d'une part la structure des entités du système et d'autre part celle d'un code orienté objet.

La figure suivante présente les paquetages des diagrammes de classes et les relations entre eux :

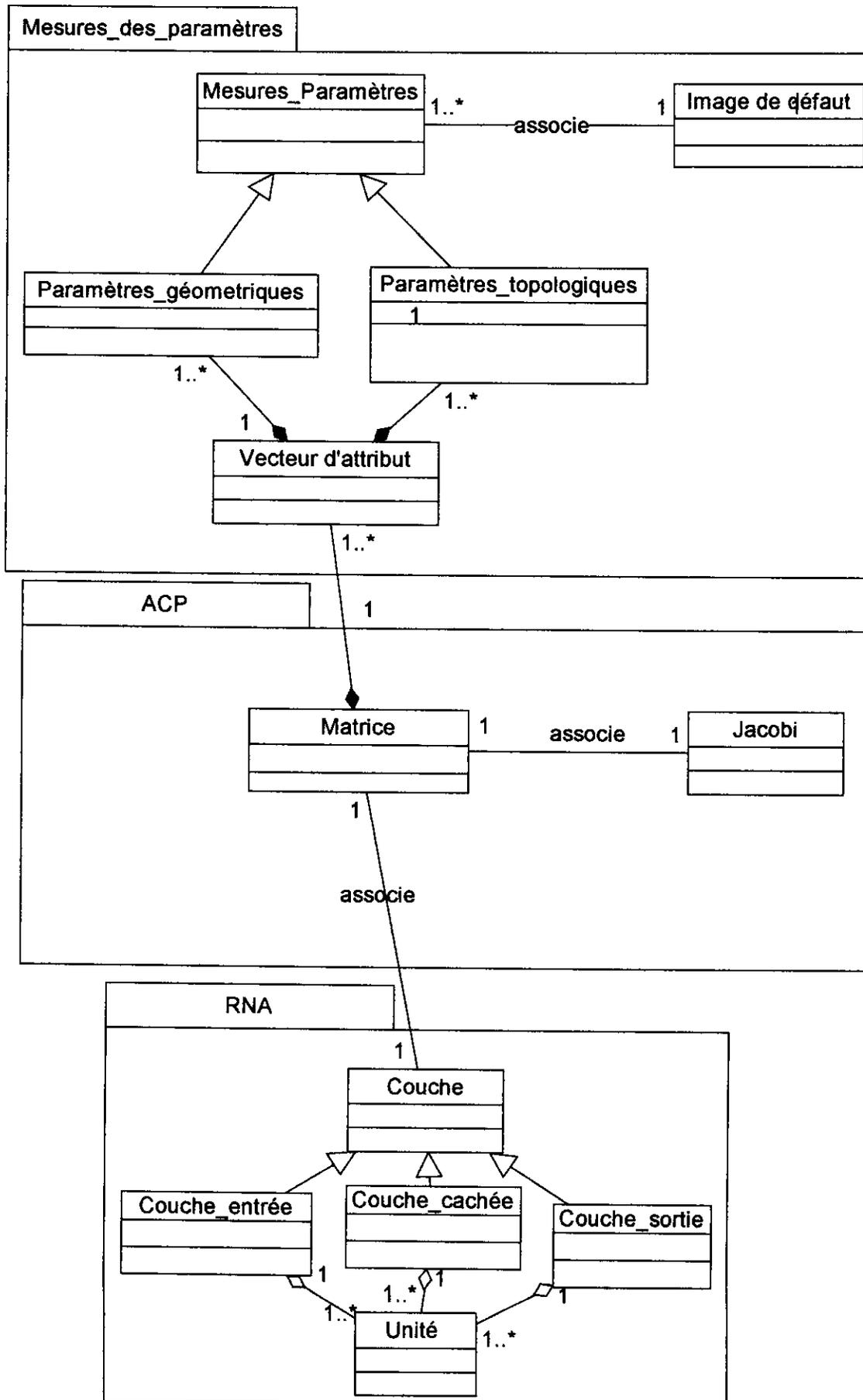


Figure VI.6 : Diagramme de classe globale

Le diagramme globale contient les trois paquetages nous avons traité chaque paquetage dans un diagramme indépendant:

Paquetage de diagramme des classes : Mesures des paramètres

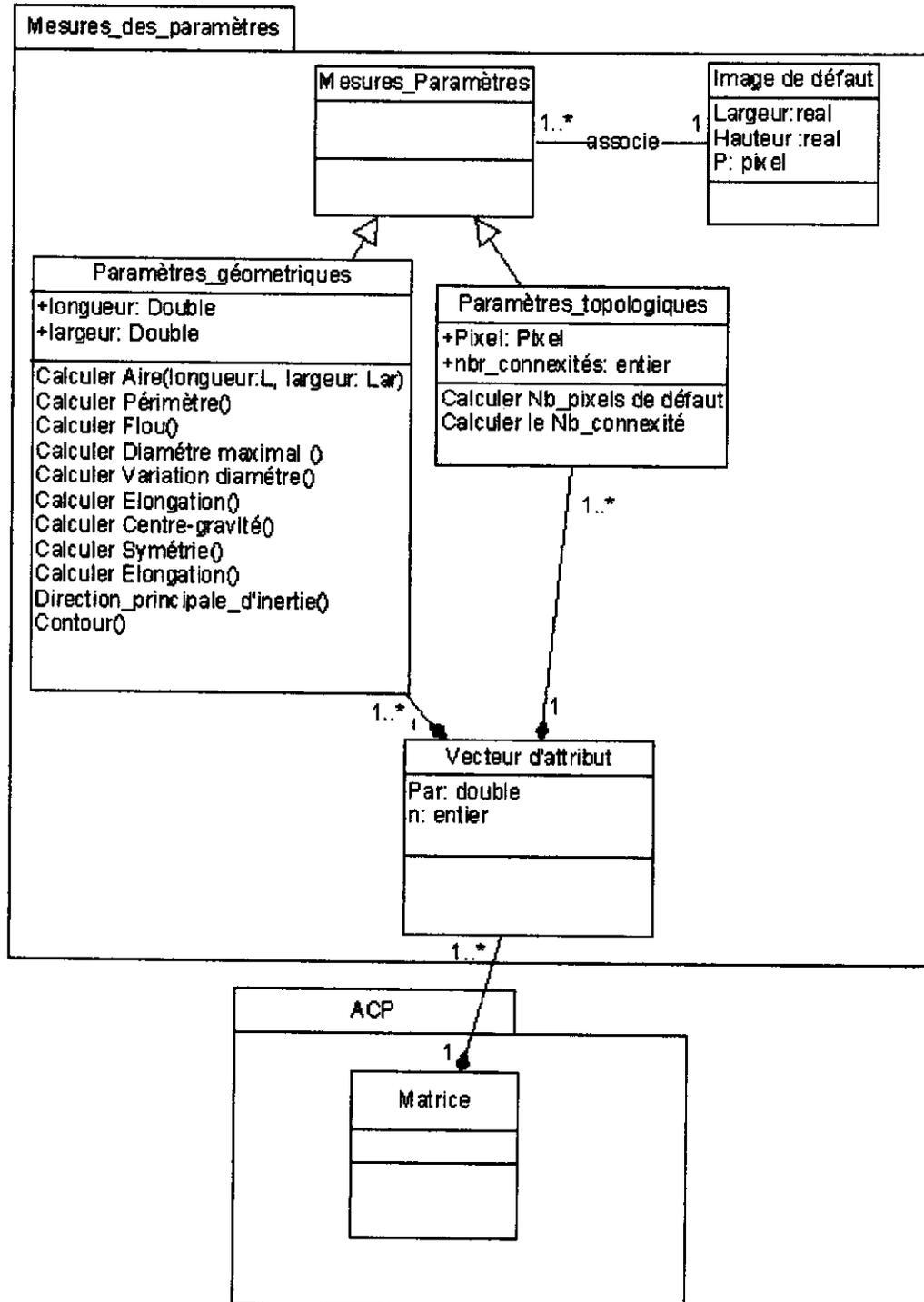


Figure VI.7: Paquetage de diagramme des classes : Mesures des paramètres

Les systèmes de reconnaissance nécessitent généralement deux étapes : une étape d'extraction des attributs et une étape de classification dans laquelle des règles de décision pour séparer les

classes dans les images, plusieurs travaux ont porté sur l'extraction de la base d'attributs quantitatifs.

Afin de caractériser nos images de défauts, nous avons opté pour des primitives géométriques sur le contour et sur les pixels définissant certaines formes topologiques. Ceci revient à la classe "paramètres_géométriques"opérer une reconnaissance grossière via les propriétés structurelles et une reconnaissance plus raffinée par analyse des distributions spatiales des formes composant l'objet à reconnaître. Ces statistiques constitueront un seul vecteur de primitives (vecteur d'attribut) qui présente la classe" Vecteur d'attribut" Nous décrivons ci-dessous les primitives choisies qui est enregistré dans ce vecteur et leur extraction:

Périmètre, variation diamètre, aire (surface), diamètre de la région, planéité, élongation, centre gravité, direction principale d'inertie, longueur, largeur, angle, symétrie, la non linéarité, flou, position, compacité, contours appartient (voisinage directe, chemin, ensemble connecté).

plus les paramètres topologiques qui sont décrits dans la classe "paramètres_topologique"

-La classe "Mesures_paramètres" consiste à déterminer un certain nombre de paramètres caractéristiques de la forme de défaut, des paramètres ayant une nature *topologique*, obtient par la classe"paramètres_topologiques", des paramètres ayant une nature géométriques , obtient par la classe" paramètres_géométriques".

- Chaque classe "Mesures_paramètres" associe à une image de défaut.

-Image est un ensemble de pixels.

Paquetage de diagramme des classes : ACP

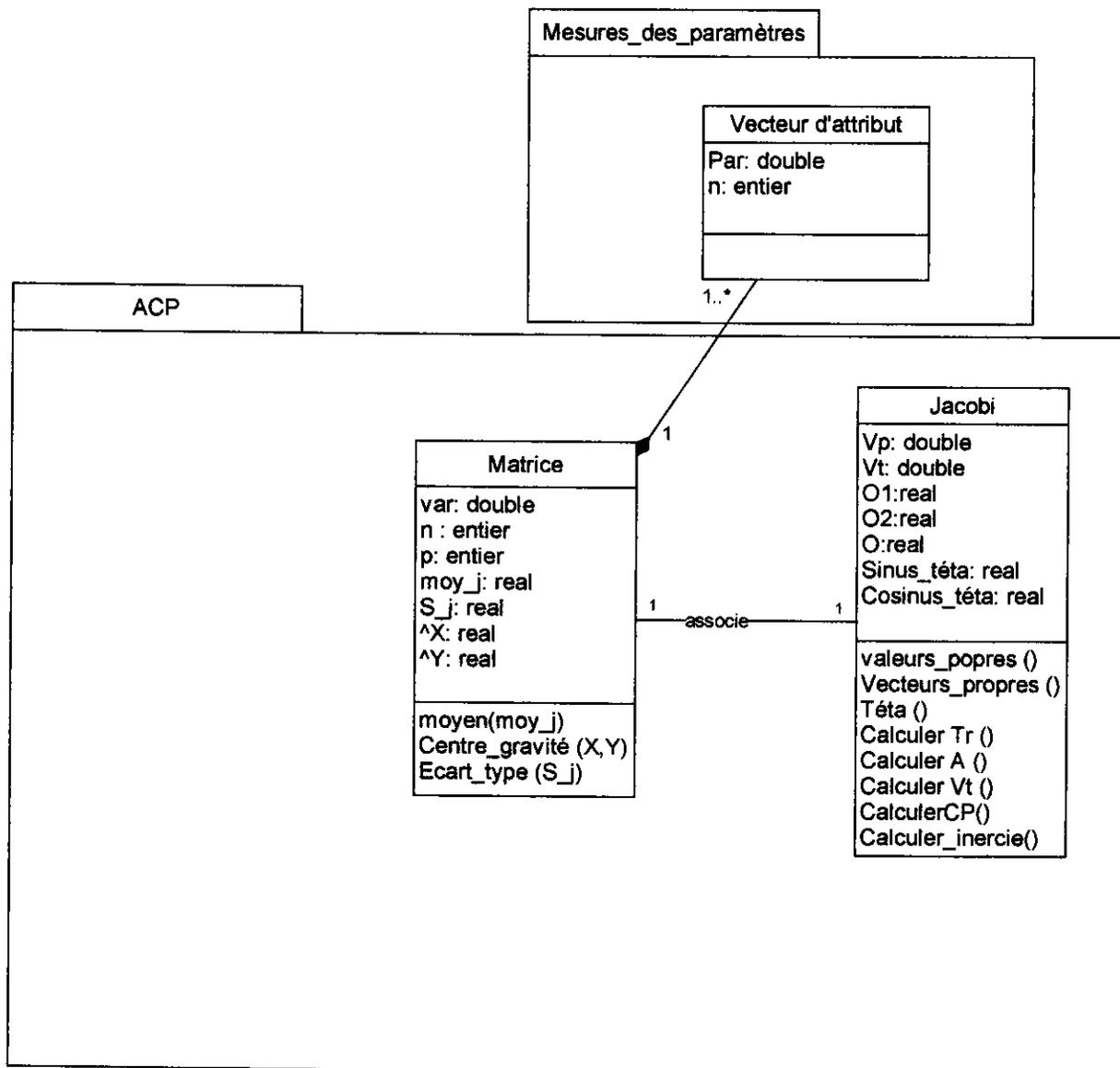


Figure VI.8: Paquetage de diagramme des classes : ACP

-Le système crée une matrice de donnée initiale de n lignes et m colonnes dont les lignes représentent les attributs géométriques invariants calculés dans l'analyse quantitative et m colonnes représentent les images des défauts de soudures.

- le paquetage "Mesures_paramètres" interrogé avec le paquetage "ACP" pour obtient une matrice réduite par la méthode ACP. Pour réduire la base d'attributs pour la recherche des meilleurs attributs qui caractérisent les images des défauts de soudure et qui répondent aux deux critères suivants : discrimination maximale et redondance minimale. Ces meilleurs attributs ont l'objectif de maintenir et/ou améliorer la performance du classifieur utilisé par le

système de reconnaissance, pour minimiser le nombre de neurones d'entrées au réseau tout en améliorant la performance du système de reconnaissance. Ceci va permettre un gain de temps de calcul considérable si un grand nombre d'entrées est supprimé.

III La Vue Dynamique (Diagramme d'activités) :

Cette vue est très importante dans les environnements multitâches; elle montre la décomposition du système en terme de processus et elle permet de définir les fonctions du système.

La vue dynamique est représentée par le diagramme d'activités pour identifier, à partir du cahier des charges, les principales composantes du système à construire.

Diagramme d'activités:

Diagramme d'activité pour le processus extraction des primitives:

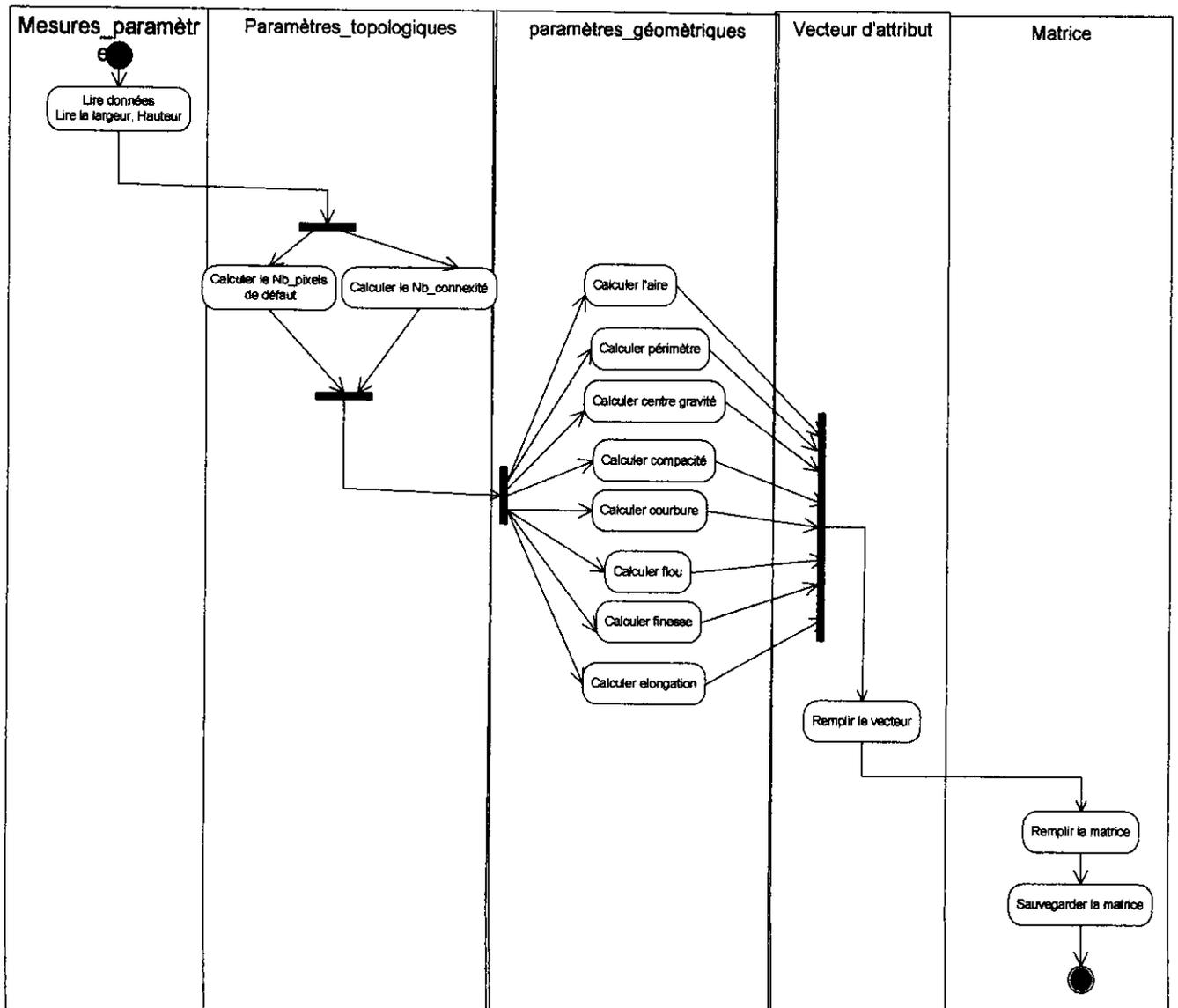


Figure VI.10: Diagramme d'activité pour le processus extraction des primitives

Description:

Grâce à des notions topologiques (Activités calculer Nbr_pixel de défauts, Activité calculer Nbr_connexités), nous pouvons déterminer des paramètres ayant une nature géométrique, à cet aspect nous allons définir les quelques notions géométriques nécessaires pour décrire paramétriquement une forme géométrique.

Activités: - Calculer l'aire,

- Calculer périmètres,
- Calculer centre de gravité,
- Calculer courbure,
- Calculer flou,
- Calculer élongation.
- etc.

Le modèle que nous avons choisi est le modèle paramétrique. Ce type de modèle consiste à caractériser un objet suivant un vecteur d'attribut dont les éléments sont des caractéristiques des paramètres topologiques et géométriques.

Le vecteur d'attribut sera enregistré dans une matrice n lignes et p colonnes dont les caractères de cette matrice sont les paramètres de ce vecteur d'attribut et les individus sont les images des défauts.

Diagramme d'activité de l'ACP:

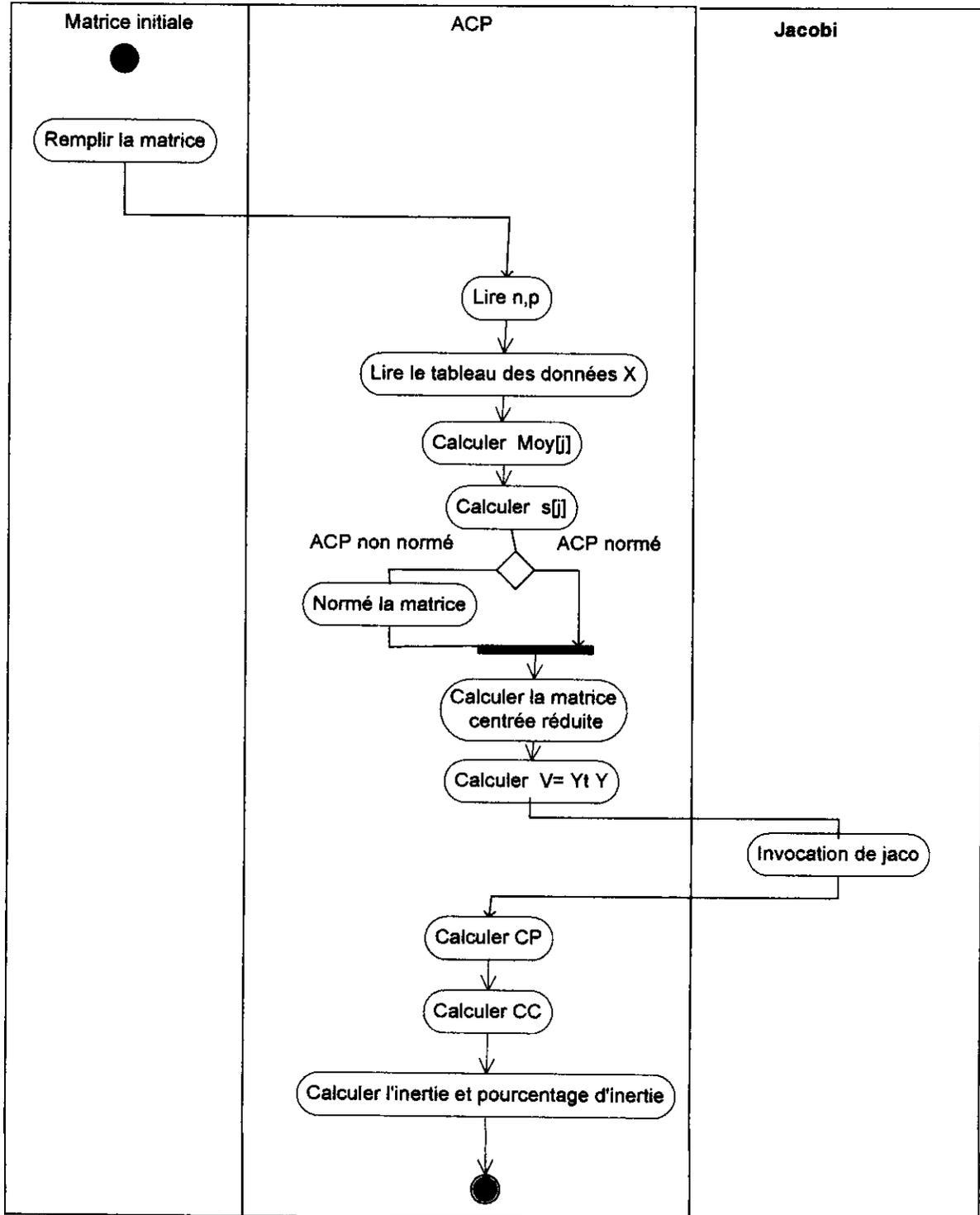


Figure VI.11: Diagramme d'activité pour le processus ACP

Description:

A ce niveau nous avons déjà la matrice initiale pour la quelle nous allons appliquer ACP(l'analyse de composantes principales)

ACP vas exécuter les activités suivantes:

- Lire n lignes et p colonnes de la matrice initiale.
- Lire le tableau de données X (la matrice initiale)
- Calculer la Moyenne Moy[j]
- Calculer écart type s[j].
- Tester si l'ACP est normé.
- Calculer la matrice centrée et réduite.
- Calculer la matrice de covariance $V = YtY$.
- Invoquer la méthode Jaco de la classe jacobi.
- Calculer CP = X.Vt coordonnées des individus.
- Calculer CC= Vt coordonnées des variables.
- Calculer l'inertie et pourcentage d'inertie.

Diagramme d'activité de Jacobi:

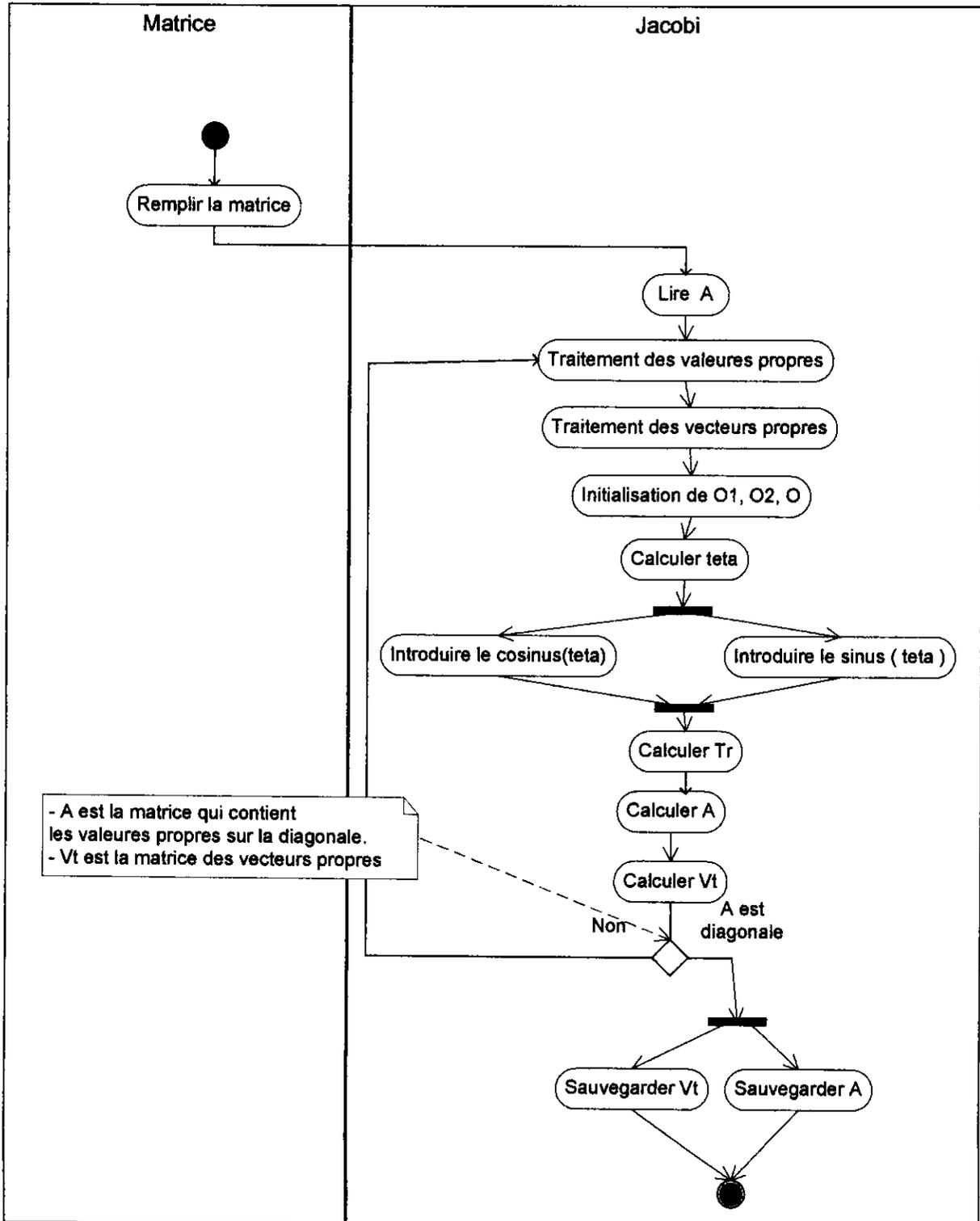


Figure VI.12: Diagramme d'activité de la méthode Jacobi

Description:

La classe ACP demande la classe Jacobi pour calculer les valeurs propres et leurs vecteurs propres.

Le procédé utilisé par la classe Jacobi pour calculer les vecteurs propres d'une matrice A, consiste à substituer à celle-ci sa matrice semblable A, qui a la particularité d'être diagonale.

Le but final recherché est de remplacer la matrice A par une matrice équivalente A' diagonale

Pour calculer des n valeurs propres d'une matrice A, la classe Jacobi, revient à exécuter les activités suivantes:

- Lire la matrice Symétrique A
- Traitement des valeurs propres.
- Traitement des vecteurs propres.
- Initialisation de O1, O2, O.
- Calculer teta.
- Introduire le cosinus (teta), le sinus (teta).
- Calculer Tr
- Calculer A est la matrice qui contient les valeurs propres sur la diagonale.
- Calculer Vt est la matrice des vecteurs propres

-Tester si la matrice A est Diagonale. Pour arriver à construire la matrice **diagonale**, nous procédons l'annulation des éléments extradiagonaux ligne par ligne, mais il se trouve que la transformation visant à annuler un certain élément d'indice i,j modifie les valeurs des éléments des lignes et colonnes i et j . Ainsi un élément déjà annulé appartenant à ces lignes ou colonnes, va prendre une nouvelle valeur non nulle. Nous sommes donc amenés à répéter le cycle d'annulation à partir du début , et cela Jusqu'à ce que les éléments extra diagonaux soient tous théoriquement nuls. Seulement à la fin de chaque cycle d'annulation, les éléments extra diagonaux prennent des valeurs de plus en plus petites.

Pratiquement la matrice est dite diagonale lorsque le plus grand terme extradiagonal est inférieur à la précision de la machine ou à une précision fixée par le calculateur. Ces éléments diagonaux d'autres que les valeurs propres de la matrice A' et donc de A.

- Sauvegarder Vt
- Sauvegarder A

Conclusion :

Pour développer notre Système nous avons commencé par la spécification des besoins, qui décrit les fonctionnalités du système, présentés par les diagrammes des cas d'utilisation, collaboration, et d'activités. Puis nous avons entamé l'étape d'analyse pour déterminer les éléments intervenants dans notre système.

La troisième étape est la phase la plus importante dans le développement de Système, elle commence par une conception globale qui décrit l'architecture du système suivie par une autre détaillée. L'étape qui se suit à pour but d'implémenter notre solution. Quand a la dernière étape, une validation est présenté après un test sur quelques images de défauts.



CHAPITRE VII

IMPLEMENTATION ET TEST

I. Introduction :

A ce niveau, il s'agit d'implémenter la solution retenue au niveau de la phase de conception, c'est la phase au cours de laquelle les structures et les algorithmes définis pendant la conception sont traduits dans un langage de programmation et une BDD.

Les tests dans une méthode 4+1 sont effectués à la fin de chaque itération; lors du développement de notre système nous sommes passés par plusieurs étapes, chacune d'elles nous donnait un résultat qu'on a évalué selon les besoins actuels. Aussi nous avons fait des choix selon des tests de performances.

II. Choix du langage de programmation :

Une des décisions les plus importantes dans la conception et la réalisation d'un logiciel est le choix du langage de programmation. Après le choix de la méthode de modélisation, nous pouvons maintenant choisir le langage de programmation. Dans notre cas le langage de programmation devait être choisi parmi les langages suivants : C, C++, java et visual Basic. Notre choix s'est porté sur le langage **Visual Basic.6** parce que c'est un langage orienté objet, et la programmation dans un langage objet est la manière la plus commode de traduire une conception objet en un langage exécutable par machine.

Pour réaliser ce système nous avons étudié le format d'un fichier BMP.

III. Test :

Dans cette partie nous allons exposer un ensemble de résultats obtenus pendant le développement de notre système.

III.1 Menu Principale :

Cette figure présente la fenêtre principale pour accéder au notre système

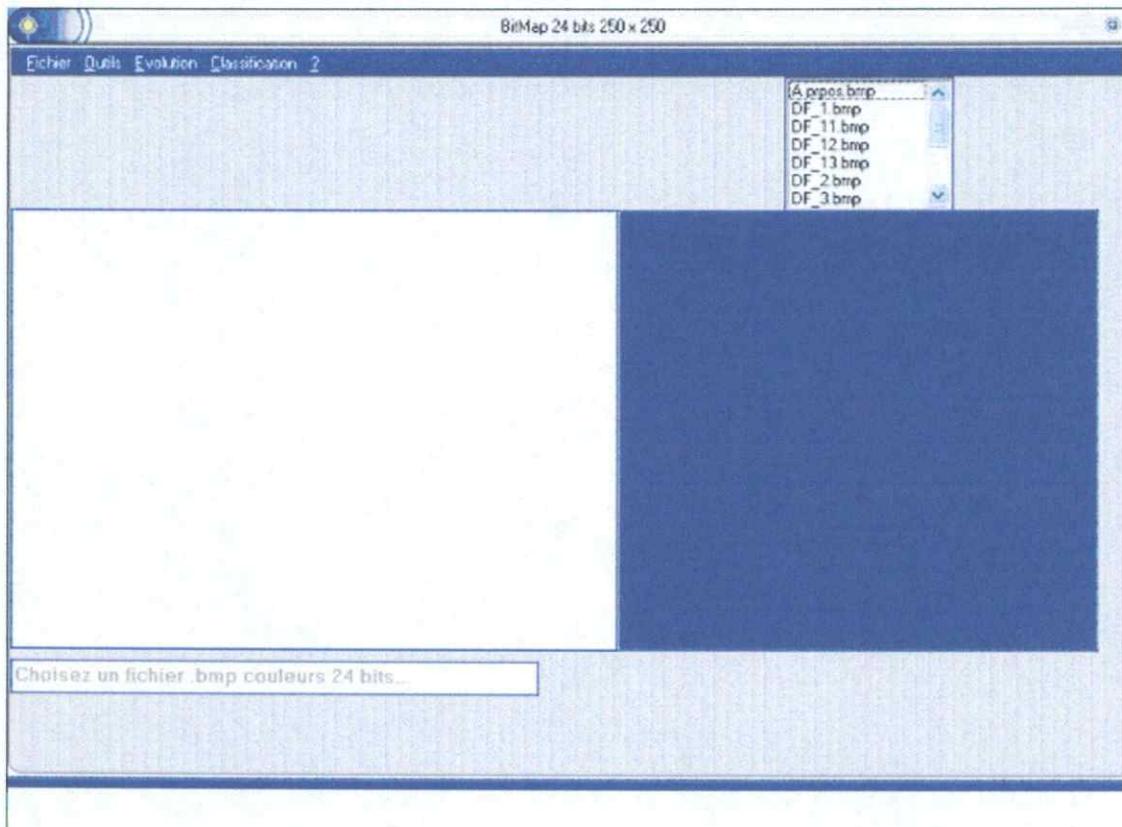
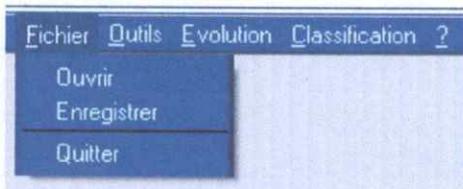


Figure VII.1 présentation le Menu Principal.

Elle appartient le menu suivant qui répond aux besoins spécifier au use cases :

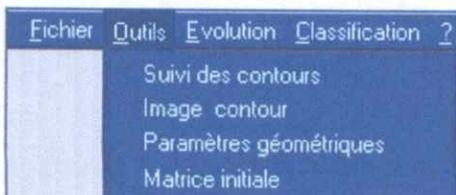


Ouvrir : ouvrir une image de défaut.

Enregistrer : sauvegarder la mise à jour fait aux images de défauts.

Quitter : pour fermer l'application retourne au Windows.

Menu des outils :



Suivi des contours : Extraire le contour de la forme (le défaut).

Image contour : afficher le contour de défaut.

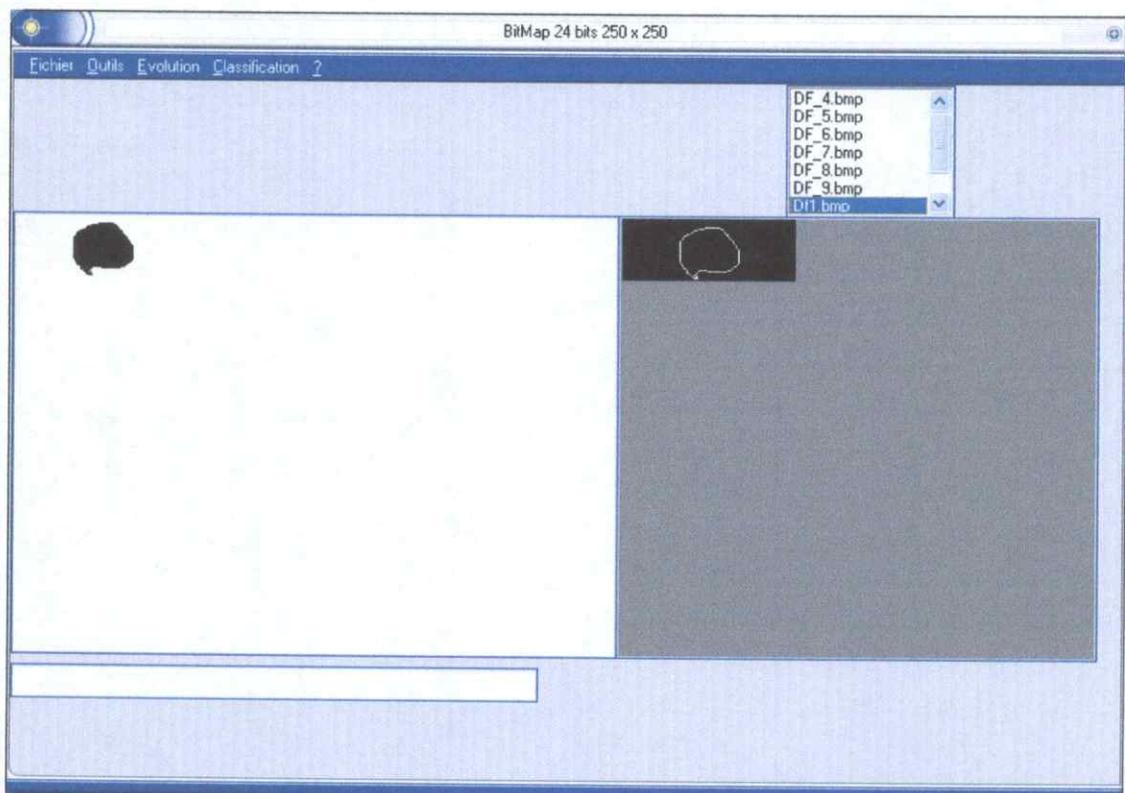
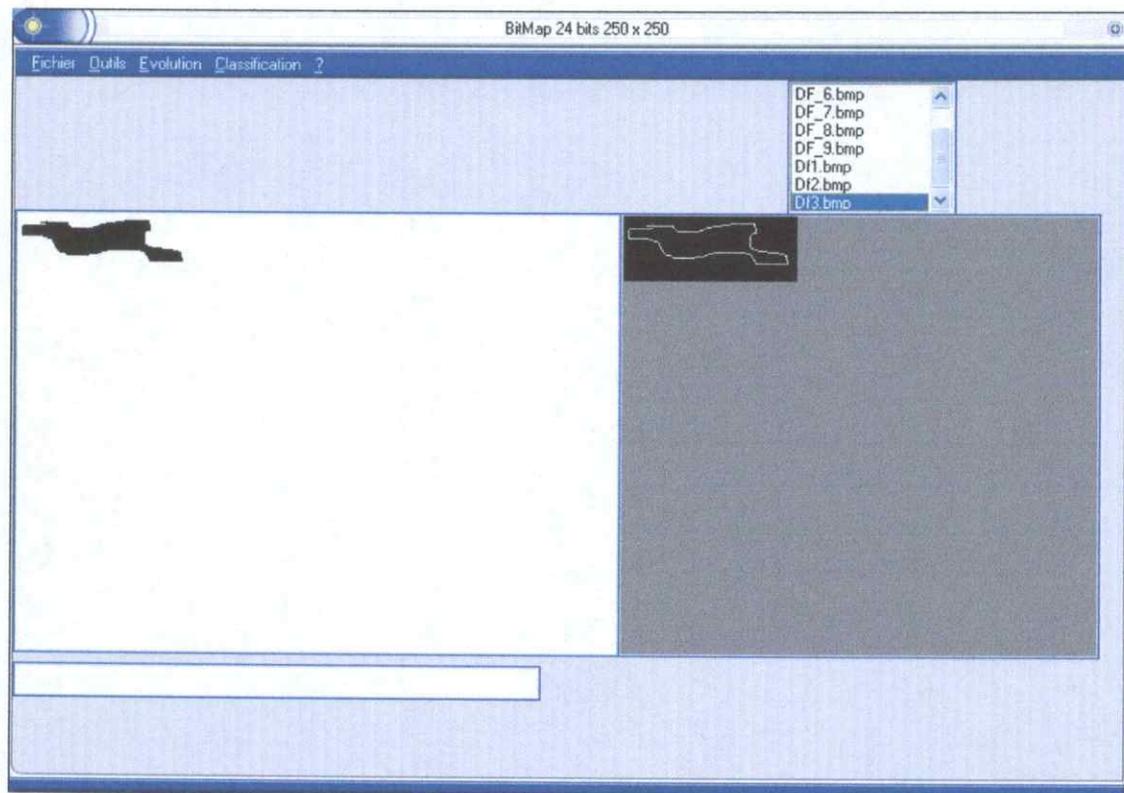


Figure VII.2 Présentation quelque image de contour associés aux défauts de soudures

Paramètres géométriques :

On illustre dans les figures VII.3 et la figure VII.4 quelques exemples de paramètres géométriques définis dans le chapitre III Analyse Quantitative, calculés pour deux types différents de défauts de soudure qui servent dans la construction des bases d'apprentissage et de test.

Les paramètres géométriques			
Surface	1838	Diamètre_maximal	232.56
Périmètre	273	elongation	0.75
Centre_gravité_X	133.284548422198	Symétrie	0.235
Centre_gravité_Y	53.4940152339499	planéité	0.01
Angle_d'orientation	0.06	Compacité	0.041
Longueur_rectangle_exinscrit	78	Anisométrie	0.382
Largeur_rectangle_exinscrit	38		

Quitter

Figure VII.3 Illustration des paramètres géométriques aux images des défauts utilisés
Cette figure contient les attributs invariants suivants :

Compacité (Comp) : Elle est proche de 0 dans le cas d'une droite et elle est égale à 1 dans le cas d'un cercle parfait.

Elongation (Elong) : Cet attribut nous décrit l'espace occupé par le défaut en termes de zone d'encadrement. Des grandes valeurs du rapport entre la longueur et la largeur.

Rectangularité (Rct) : Elle est égale à 1 dans le cas d'un rectangle parfait.

Anisométrie (Ani) : Elle est liée à la direction des axes principaux de la matrice d'inertie du défaut. Sa valeur est proportionnelle à l'allongement du défaut. Elle est égale à 1 dans le cas d'un carré ou d'un cercle.

Symétrie (Sym) : La valeur 1 correspond à une forme symétrique.

Indice d'allongement (Ia) : Les grandes valeurs de cet indice mettent en évidence les fissures fines et rectilignes.

Indice d'écart au cercle inscrit (Ir) : La valeur de cet indice est maximale (proche de 1) pour les défauts allongés et minimale (proche de 0) pour les défauts ronds.

Illustration graphique des paramètres géométriques :

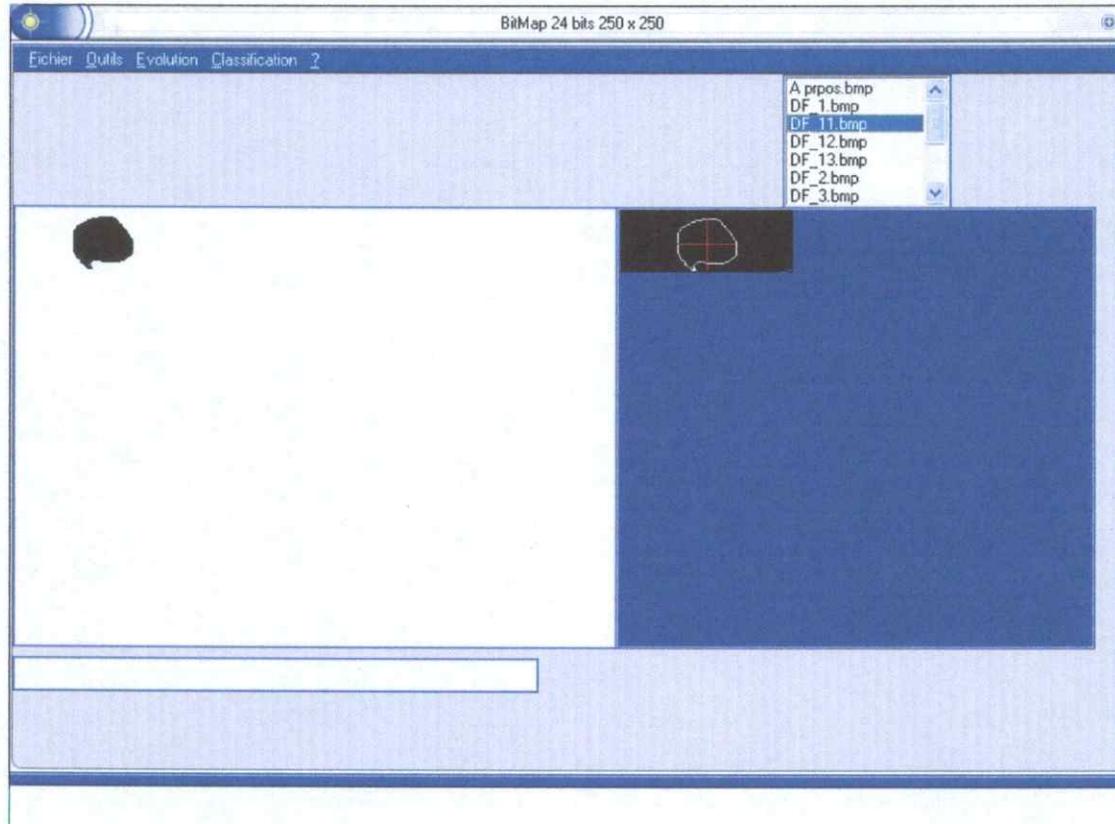


Figure VII.4 Illustration graphique des premiers Axes Principaux aux images des défauts utilisés

Cette fenêtre permet d'afficher :

Les axes principaux : pour dessiner le rectangle exinscrit (réctangularité), ellipse représentative.



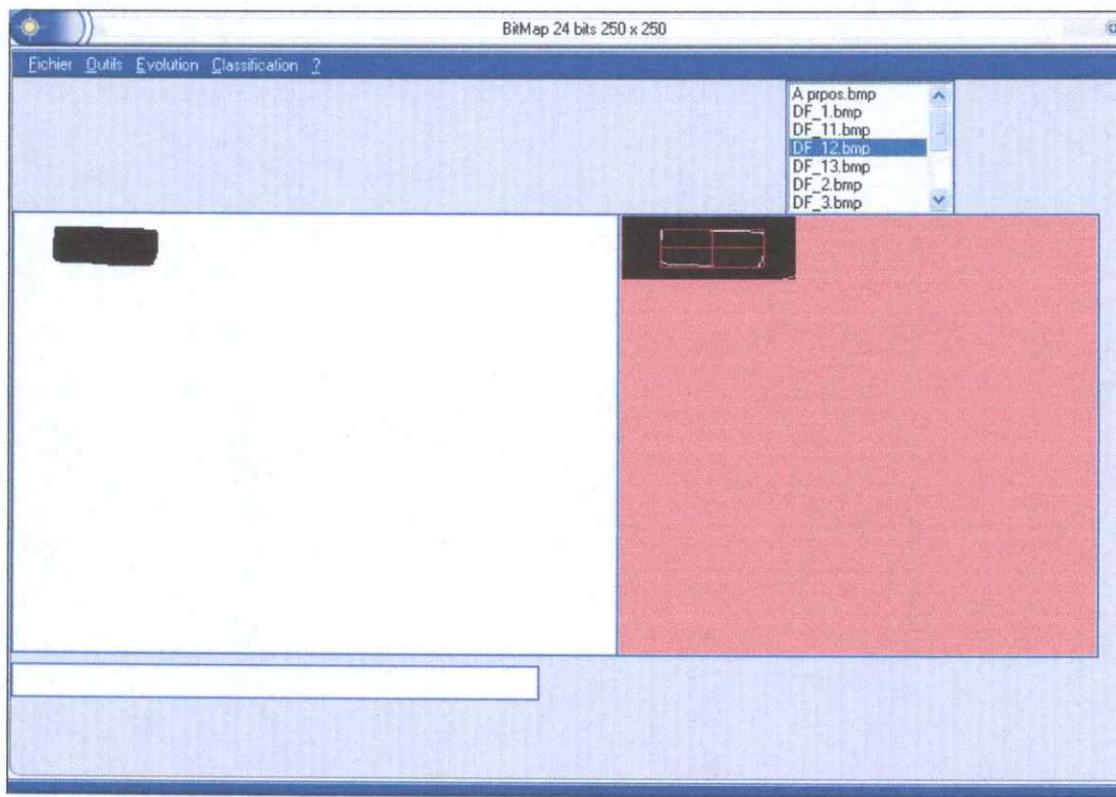


Figure VII.5 Illustration graphique de rectangle exinscrit aux images des défauts utilisés.

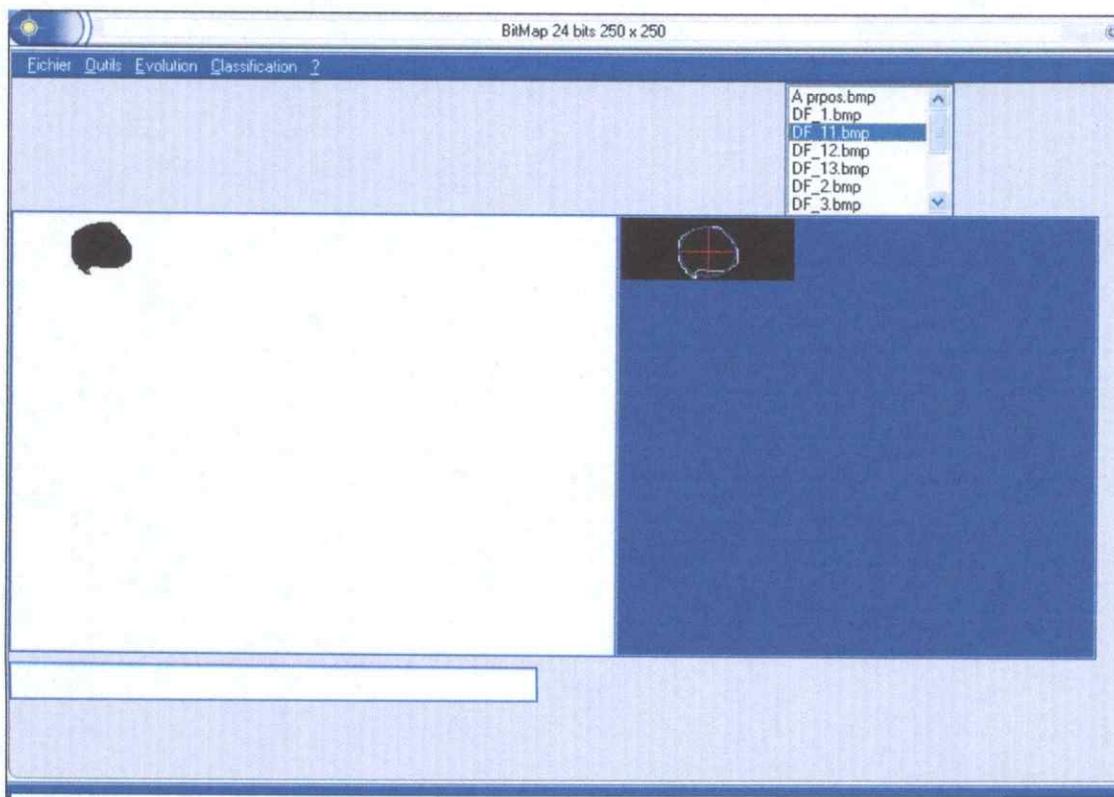


Figure VII.6 Illustration graphique de Cercle exinscrit aux images des défauts utilisés.

Matrice Initiale : Consulter la matrice originale c'est-à-dire la base de connaissance.

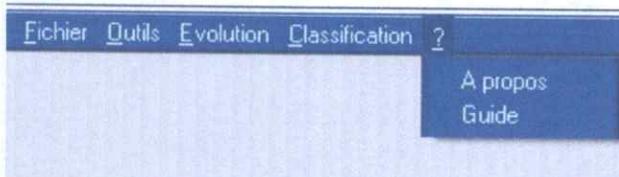
La base de Données des attributs géométriques invariants des images des défauts

Defaut	Compacité	Elongation	Rectangularité	Symétrie	Planeité	allongement	Eccentricité
10.bmp	0,357	0,277	0,522	0,731	1,862	0,826	0,59
3.bmp	0,64	0,128	0,385	0,511	0,668	0,865	0,217
6.bmp	0,641	0,434	0,603	0,494	1,742	0,437	0,58
2.bmp	0,663	0,74	0,973	0,85	3,852	0,62	0,081
4.bmp	0,64	0,128	0,385	0,299	0,668	0,386	0,217
5.bmp	0,237	0,875	0,64	0,717	0,238	0,494	0,441
1.hmn	0,863	1,135	0,781	0,737	0,957	0,133	0,4

Figure VII.7 La matrice initiale contient les attributs associés aux défauts de soudures

La figure VII.7 donne les valeurs des attributs mesurées pour les défauts présentés dans les paramètres géométriques enregistrés dans une matrice de n*m dont n individus représentent les images de défauts dans notre cas nous avons fait les testes sur trois défauts DF_1, DF_2, DF_3, m caractères représentent les attributs calculés dans la fenêtre paramètres géométriques.

Menu d'aide :



Fenêtre A propos :

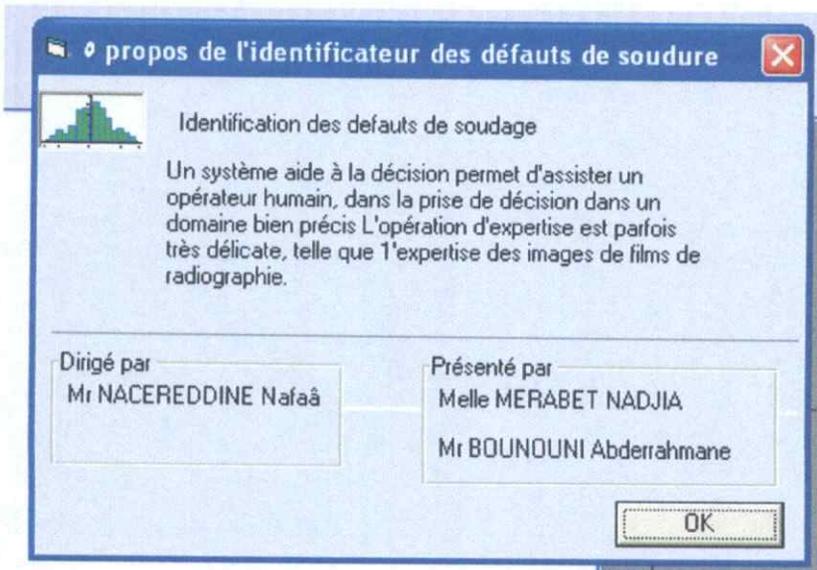


Figure VII.8 A propos de l'application

Matrice réduite :

Matrice initiale	[0,357][0,277][0,522][0,731][1,862][0,826][0,59]	Valeurs propres	Variance cumulée
Matrice centrée	[0,64][0,128][0,385][0,511][0,668][0,865][0,217]	[36]	[0,395]
Matrice réduite	[0,641][0,434][0,603][0,494][1,742][0,437][0,58]	[25]	[0,669]
Vecteur propre	[0,663][0,74][0,973][0,85][3,852][0,62][0,081]	[16,012]	[0,844]
Les composantes principales	[0,64][0,128][0,385][0,299][0,668][0,386][0,217]	[9]	[0,942]
	[0,237][0,875][0,64][0,717][0,238][0,494][0,441]	[4]	[0,985]
	[0,863][1,135][0,781][0,737][0,957][0,133][0,4]	[1,001]	[0,995]
		[0]	[0,995]



CONCLUSION GENERALE

Notre travail est réalisé dans le but est la classification, passant par l'analyse quantitative des images basée sur la recherche d'attributs.

Quant à l'analyse quantitative des formes des défauts ainsi extraites, le souci majeur reste comment construire un ensemble d'attributs qui caractérisent le plus efficacement possible ces formes, tout en tenant compte des spécificités des défauts qu'elles représentent et de la subjectivité et des aléas de leur interprétation. C'est pourquoi, il se peut qu'un seul attribut joue un rôle décisif dans la discrimination de deux classes de défauts réellement distinctes, soit par l'origine de leur obtention ou par la sévérité des normes par rapport à leur interprétation. Ajoutant à cela, que dans la reconnaissance d'un tel défaut, la même forme de ce dernier peut être vue à partir de plusieurs angles selon l'orientation et la distance de l'ouvrage soudé par rapport à la source d'irradiation. C'est pourquoi, la recherche d'attributs invariants aux transformations géométriques de rotation, de translation et de changement d'échelle s'avère nécessaire. Ainsi, une panoplie d'attributs géométriques satisfaisant les conditions suscitées est proposée dans le troisième chapitre qui découlent du calcul des paramètres géométriques (surface, périmètre, etc.). Vu la grande plage qu'occupent les valeurs de ces attributs et vu d'autres considérations liées à la classification neuronale, la normalisation de ces valeurs dans l'intervalle $[0,1]$ s'impose. La réduction du nombre de variables du vecteur d'attributs doit se faire judicieusement, en prenant en considération le nombre de mesures effectuées (nombre d'observations) et l'analyse en composante principale entre les différentes variables.

Dans le cinquième chapitre, nous avons développé les paradigmes des réseaux de neurones artificiels (RNA) qui suscitent un intérêt particulier, surtout dans le domaine de la classification. Cette dernière se base sur l'utilisation des caractéristiques quantitatives mesurées. Comme nous l'avons déjà vu, ces caractéristiques sont regroupées dans un vecteur d'attribut et dont les composantes constituent les entrées du classificateur neuronal. Les types des défauts de soudure constituent ses sorties. La difficulté principale est de trouver un ensemble d'apprentissage représentatif. Dans notre travail, la classification proposée consiste à affecter les principaux types des défauts de soudure à quatre classes ou catégories de défauts qui se basent sur les caractéristiques morphologiques des défauts couramment rencontrés.

En se basant sur les résultats de ladite classification, nous pouvons énoncer que l'avenir de l'identification automatique des défauts de soudure à travers la classification neuronale est très prometteur. La procédure d'un travail futur

Notre travail est réalisé dans le but est la classification, passant par l'analyse quantitative des images basée sur la recherche d'attributs.

Quant à l'analyse quantitative des formes des défauts ainsi extraites, le souci majeur reste comment construire un ensemble d'attributs qui caractérisent le plus efficacement possible ces formes, tout en tenant compte des spécificités des défauts qu'elles représentent et de la subjectivité et des aléas de leur interprétation. C'est pourquoi, il se peut qu'un seul attribut joue un rôle décisif dans la discrimination de deux classes de défauts réellement distinctes, soit par l'origine de leur obtention ou par la sévérité des normes par rapport à leur interprétation. Ajoutant à cela, que dans la reconnaissance d'un tel défaut, la même forme de ce dernier peut être vue à partir de plusieurs angles selon l'orientation et la distance de l'ouvrage soudé par rapport à la source d'irradiation. C'est pourquoi, la recherche d'attributs invariants aux transformations géométriques de rotation, de translation et de changement d'échelle s'avère nécessaire. Ainsi, une panoplie d'attributs géométriques satisfaisant les

I UML (L'Unified Modeling Language):

L'Unified Modeling Language (UML) est un formalisme de modélisation orienté objet. Comme son nom l'indique, Il est issu de la fusion des trois méthodes de modélisation, tirant donc profit des avantages de chacune :

- méthode **Booch** de Grady Booch,
- méthode **OOSE** de Ivar Jacobson,
- méthode **OMT** de James Rumbaugh, [10].

UML est un langage graphique de modélisation objet permettant de spécifier, de construire, de visualiser et de décrire les détails d'un système logiciel.

Il comprend un vocabulaire et un ensemble de règles centré sur la représentation conceptuelle et physique du système.

UML supporte les concepts suivants:

- Des concepts (structurels, comportementaux, annotationnels, groupements).
- Des relations (association, généralisation, les agrégations, les compositions).
- Des diagrammes (statiques et dynamiques).

La section suivante a pour but de présenter le langage de modélisation de données UML:

I.1 Les concepts:

UML supportent quatre (04) types de concepts : [11]

Les concepts structurels : représentés par les classes, les interfaces, les collaborations...

Les concepts comportementaux : représentés par les interactions et les états des objets.

Les concepts annotationnels : représentés par les notes.

Les concepts de groupement: Représentés par les sous-systèmes, les paquetages.

Paquetages (Packages) : Ils offrent un mécanisme générale pour la partition des modèles et le regroupement des éléments de modélisation. Il permet de définir des sous-systèmes formés d'éléments. [12]

Chaque package est représenté par un dossier, et doit avoir un nom différent de celui des autres packages, et peut être composé d'autres éléments, y a compris d'autres packages.

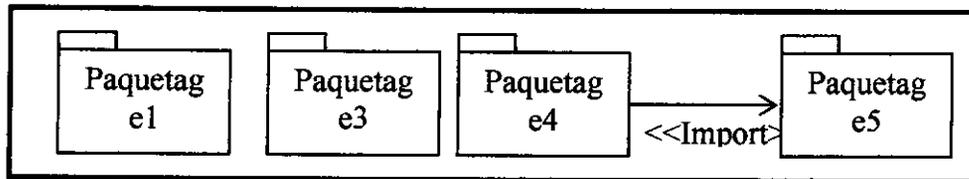


Figure A.1 Représentation graphique d'un package

I.2. Les relations : Elles permettent de relier les concepts entre eux.

Exemple : les associations, les généralisations, les agrégations, ... etc. [13].

I.2.1. L'association: Elle représente une relation précisant que les objets d'un élément sont reliés aux objets d'un autre élément.

I.2.2. La généralisation: Elle représente une relation entre un élément général et un élément dérivé de celui-ci, mais plus spécifique (désigné par sous-élément ou élément fils).

Le plus souvent, la relation de généralisation est utilisée pour représenter une relation d'héritage.

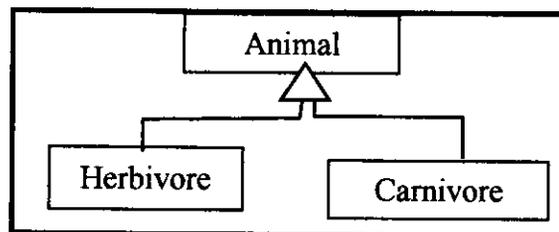


Figure A.2 Représentation graphique de la généralisation

I.2.3. L'agrégation : Elle représente une association non symétrique dans laquelle une extrémité joue un rôle prédominant par rapport à l'autre extrémité [6].

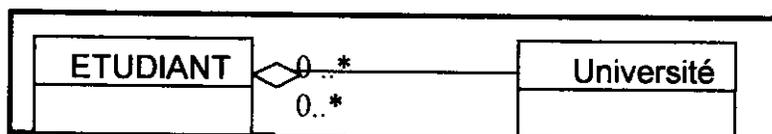


Figure A.3: Représentation graphique d'une agrégation

I.3. Les diagrammes d'UML :

UML définit neuf types de diagrammes qui sont présentés dans l'extrait du méta modèle suivant :

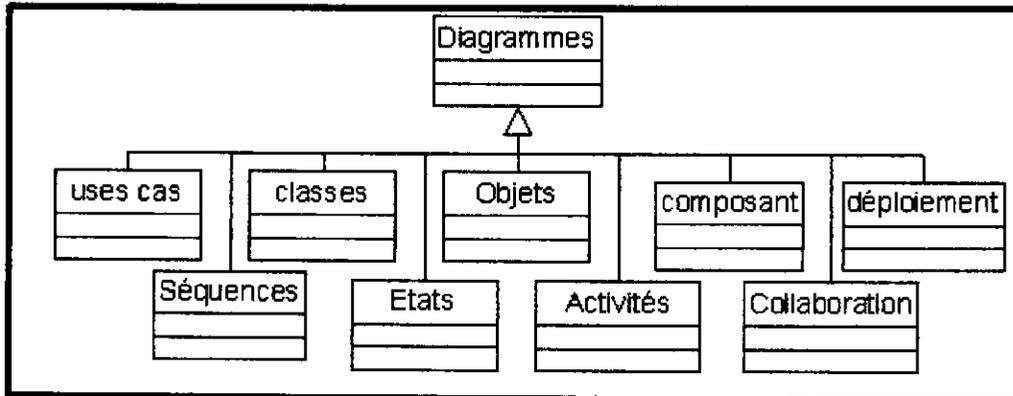


Figure A.4 : Les diagrammes d'UML [10]

I. 3.1 Diagrammes statiques:

I.3.1.1 diagramme de cas d'utilisation :

Un cas d'utilisation est une interaction effectuées par un acteur sur un système.

Un acteur peut représenter un rôle joué par une personne ou une logiciel qui interagit avec le système [12].

Un diagramme de cas d'utilisation permet de décrire les interactions entre les acteurs de l'organisation et le système dans chacun des cas d'utilisation envisagés. Il décrit le comportement d'un système au point de vue de l'utilisateur et fixe les limites du système et les relations entre le système et l'environnement [12].

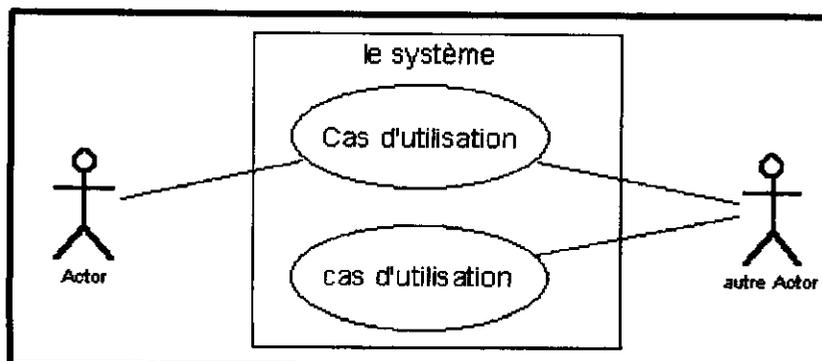


Figure A.5 : Représentation graphique de cas d'utilisation

I.3.1.2 diagramme de classes :

Il montre la structure du système en terme des classes, des relations d'héritages entre classes, des associations, dont les agrégations, des attributs, des opérations et la spécification d'opérations et contraintes au niveau des entités [13].

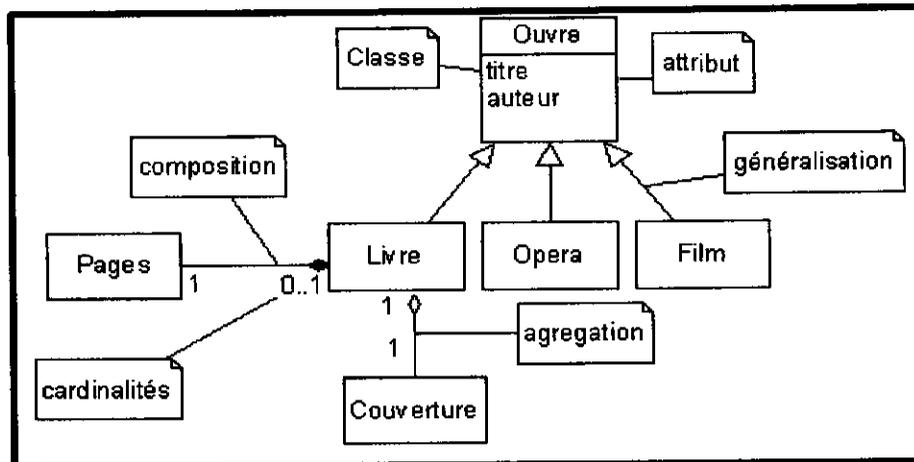


Figure A.6 Exemple de diagramme de classe

I.3.1.3 Diagramme d'objets :

Ce type de diagramme UML montre des objets (instances de classes dans un état particulier) et des liens (relations sémantiques) entre ces objets [10].

Les diagrammes d'objets s'utilisent principalement pour montrer un contexte, par exemple avant ou après une interaction, mais également pour faciliter la compréhension des structures de données complexes [12].

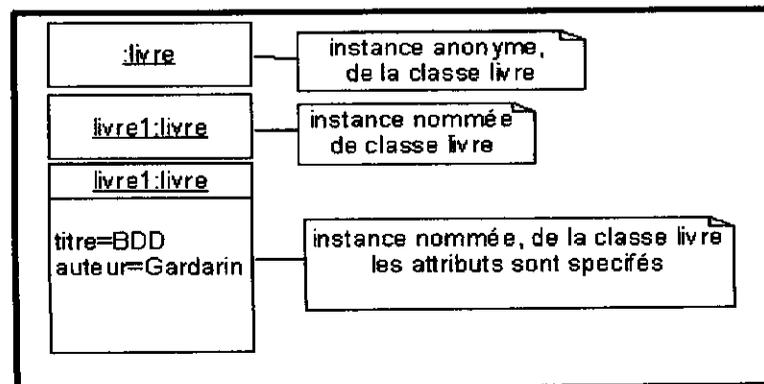


Figure A.7 Représentation graphique d'un objet

I.3.1.4 Diagramme de composants :

Il permet de décrire l'architecture physique et statique d'une application en terme de modules : fichiers sources, librairie, exécutables etc. Ils décrivent les éléments physiques et leurs relations dans l'environnement de réalisation, ils montrent les choix de réalisation [12].

Ils montrent la mise en oeuvre physique des modèles de la vue logique avec l'environnement de développement.

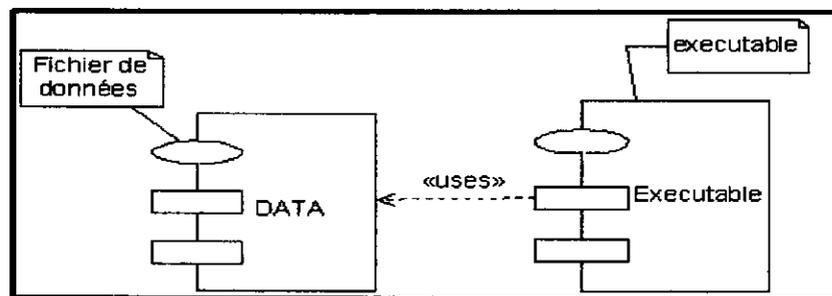


Figure A.8 Représentation d'un composant (fichier)

I.3.1.5 Diagramme de déploiement :

Il montre la disposition physique des matériels qui composent le système et la répartition des composants sur ces matériels. Ils peuvent montrer des classes de nœuds ou des instances de nœuds [14].

Les nœuds sont connectés entre eux, à l'aide d'un support de communication. La nature des lignes de communication et leurs caractéristiques peuvent être précisées. Les diagrammes de déploiement peuvent montrer des instances de nœuds (un matériel précis), ou des classes.

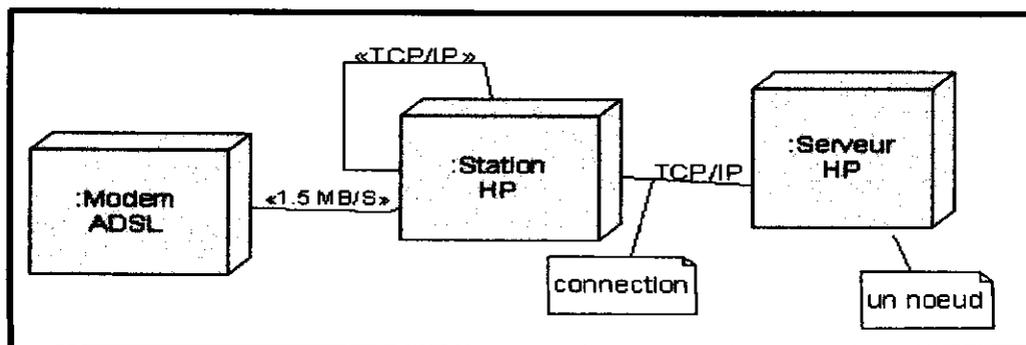


Figure A.9 Représentation de relations entre nœuds

I.3.2 Diagrammes dynamiques :

- **Interactions:** modélisent un comportement dynamique entre objets [6]. Elles se traduisent par l'envoi de messages entre objets.
- **Les messages :** Les messages échangés sont représentés au moyen de flèches horizontales partant de l'émetteur vers le récepteur. L'ordre de l'envoi est donné par la position sur l'axe vertical [15].

I.3.2.1. Diagramme de séquences :

Un diagramme de séquence représente une interaction entre objets, en insistant sur la chronologie des envois de messages [13].

Ils peuvent être utilisés pour représenter les scénarios d'un cas d'utilisation donnée [16].

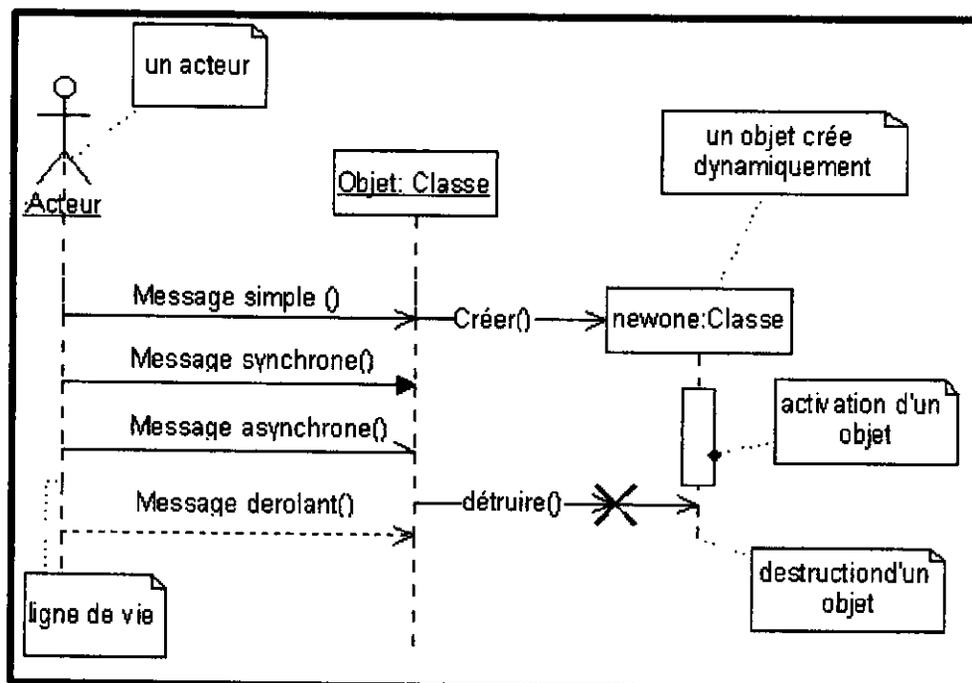


Figure A.10 Agencement de messages

Période d'activation : Correspond au temps pendant lequel un objet effectue une action, soit directement, soit par l'intermédiaire d'un autre objet.

I.3.2.2. Diagrammes de collaboration :

Il montre des interactions entre les objets et les acteurs. Ils permettent de représenter le contexte d'une interaction, car on peut y préciser les états des objets qui interagissent et peuvent être utilisés pour identifier les principaux objets [16].

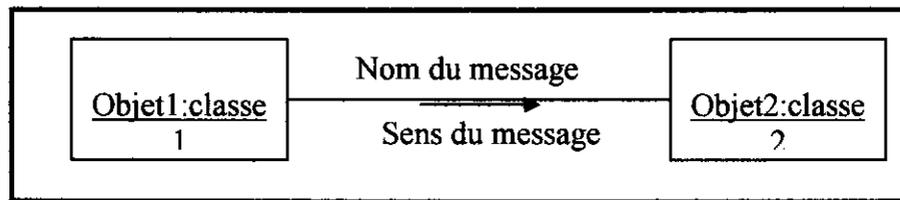


Figure A.11 Formalisme de base du diagramme de collaboration

I.3.2.3 diagramme d'état-transition :

C'est un graphe constitué de nœuds représentant des états ainsi que des flèches représentant des transitions, portant des paramètres et des noms d'événements [10].

Les diagrammes d'états permettent de définir le comportement d'un objet particulier vis-à-vis des sollicitations internes ou externes auxquelles il peut être soumis [16]. Ils permettent aussi de décrire l'évolution dans le temps les états des objets d'une certaine classe, les événements auxquels ils réagissent et les transitions qu'ils effectuent.

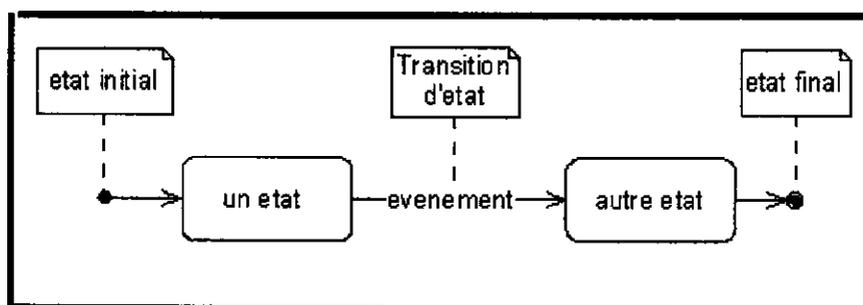


Figure A.12 Diagramme d'état-transition

I.3.2.4. Digramme d'activité :

Il permet de décrire le déroulement d'un cas d'utilisation particulier. Il est possible de décrire les acteurs responsables de chaque activités par l'utilisation des «couloirs d'activités» qui permettent de répartir graphiquement les différentes activités entre les acteurs opérationnels [12]. Chaque activité est placée dans le «couloir» correspondant à l'acteur qui assume cette activité.

II. Méthodologie de conception RUP :

RUP est l'Unique implémentation commerciale d'UP (*Unified Process*), un processus théoriquement « standard » élaboré par Rational .

RUP est une évolution du processus *Objectory* initialement mis au point par la société Rational. RUP est défini par ses créateurs comme un environnement d'exécution de processus génériques destinés à gérer l'ensemble des aspects d'un projet de développement de logiciel. RUP abrège des bonnes pratiques formalisées qui sont destinées aux différents acteurs d'un projet [5].

RUP est néanmoins décrit par ses auteurs comme un méta processus. C'est-à-dire qu'il est possible d'instancier RUP. En clair, cela signifie que RUP peut être adapté à tous les types de projets, pour tenir compte :

- De la taille du projet ;
- Du niveau de compétence initiale des acteurs (par exemple dans le cas où l'entreprise n'utiliserait pas de processus de développement).

Les fondements de RUP:

RUP est, pour citer ses auteurs :

« Piloté par les cas d'utilisation, centré sur l'architecture, itératif et incrémental. » [5]

Piloté par les cas d'utilisation :

Les cas d'utilisation, dans RUP, ne servent pas qu'à améliorer la capture des exigences fonctionnelles des utilisateurs. Ils servent d'intrant aux différentes activités du processus : analyse, conception, implémentation, rédaction des tests et déploiement.

Le modèle de conception est développé à partir des cas d'utilisation ; le modèle d'implémentation reprend le modèle de conception et définit des composants qui groupent les

classes en unités d'implémentations indépendantes.. Chaque modèle est respectivement d'un niveau d'abstraction inférieur à celui dont il descend. C'est-à-dire que le modèle d'implémentation est plus proche de la structure du code source finalement produit que le modèle de conception. RUP contient les guides qui permettent aux développeurs de se familiariser avec ces concepts. Mais rien ne remplace la pratique et l'expérience accumulée en projet.

II.1. La vue logique :

Cette vue de haut niveau se concentre sur l'abstraction et l'encapsulation, elle modélise les éléments et mécanismes principaux du système ; Elle identifie les éléments du domaine, ainsi que les relations et interactions entre ces éléments. Le système est décomposé en ensemble d'abstractions principales, pris (la plupart du temps) du domaine de problème, sous forme *d'objets* ou *de classes d'objet* Ils exploitent les principes de l'abstraction.

Nous employons l'approche d'UML comme nous avons dit précédemment pour représenter l'architecture logique, au moyen de diagrammes de classe et diagrammes d'objets.

II.2. La vue de processus «la vue dynamique »:

Cette vue est très importante dans les environnements multitâches ; elle montre : [6], [7]

- La décomposition du système en terme de processus (tâches).
- Les interactions entre les processus (leur communication).
- La synchronisation et la communication des activités parallèles.

Cette vue est basé sur les diagrammes dynamiques d'UML.

II.3. La vue des composants : « la Vue de Développement »

Elle se concentre sur l'organisation réelle de logiciel, cette vue est de bas niveau (aussi appelée "vue de réalisation"), montre : [6], [7]

- L'allocation des éléments de modélisation dans des modules (fichiers sources, bibliothèques dynamiques, bases de données, exécutables, etc...).
- L'organisation des composants, c'est-à-dire la distribution du code en gestion de configuration, les dépendances entre les composants...
- Les contraintes de développement (bibliothèques externes...).

En d'autres termes, cette vue identifie les modules qui réalisent (physiquement) les classes de la vue logique. Elle est représentée par des diagrammes de composants dans la notation

d'UML.

II.4 La vue physique : « la vue de déploiement »

Cette vue très importante dans les environnements distribués, décrit les ressources matérielles et la répartition du logiciel dans ces ressources : [6], [7]

- La disposition et nature physique des matériels, ainsi que leurs performances.
- L'implantation des modules principaux sur les nœuds du réseau.

II.5 La vue des cas d'utilisation :

La description d'une architecture peut être organisée autour de ces quatre vues, et alors illustré par les cas d'utilisation ou *scénarios* qui deviennent une cinquième vue. [6], [7] Cette vue définit les besoins des clients du système et centre la définition de l'architecture du système sur la satisfaction (la réalisation) de ces besoins.

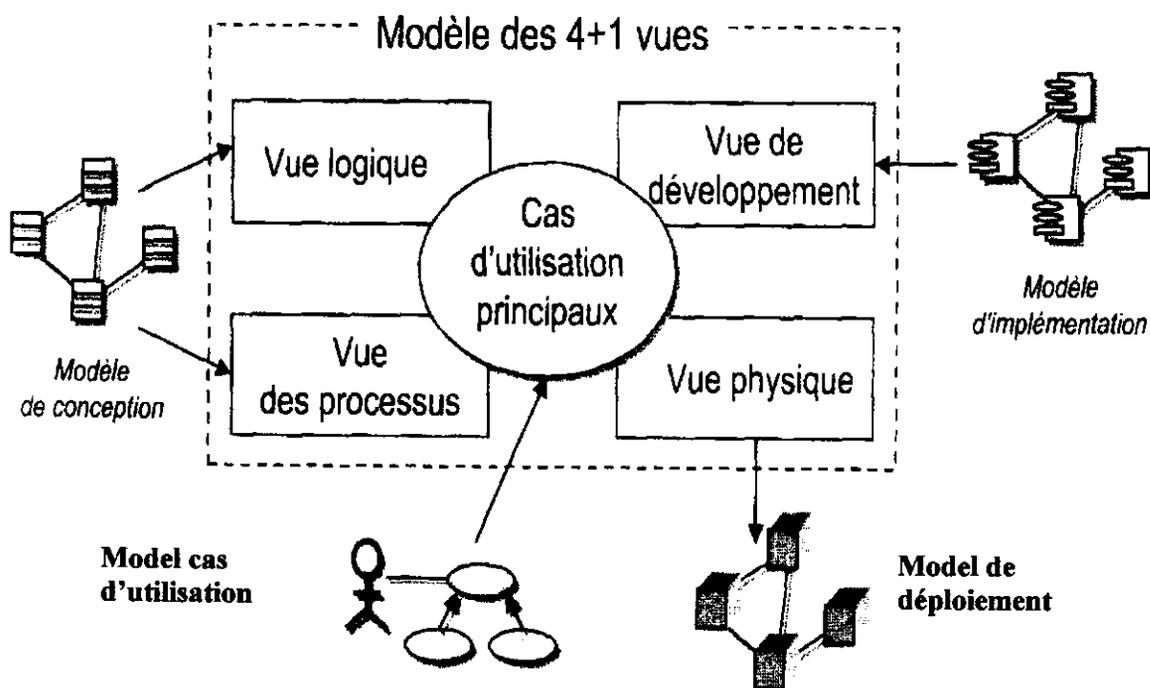


Figure A.13 Modèle 4+1. [7]

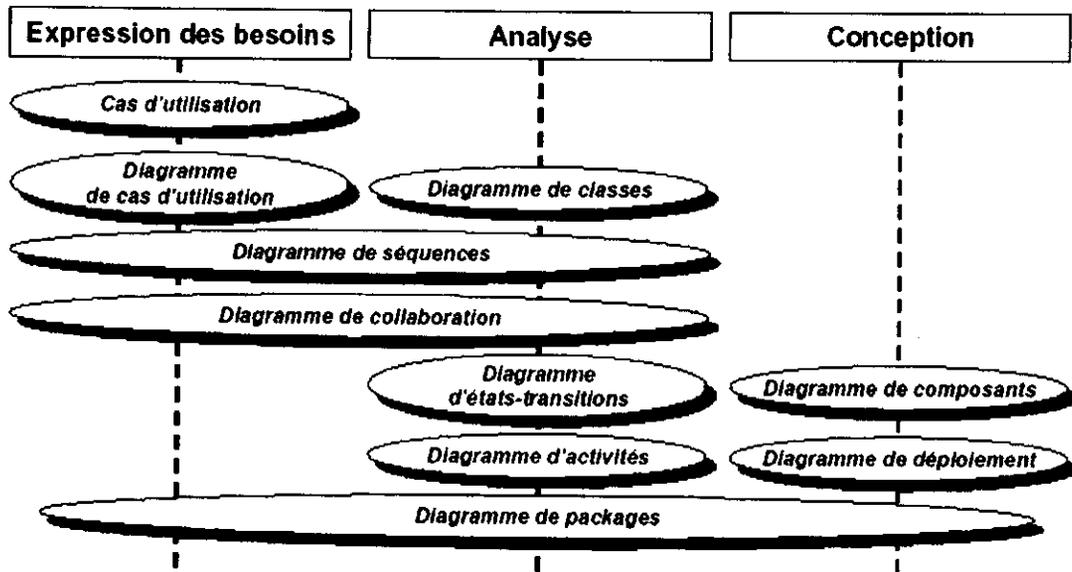


Figure A.15 : Diagrammes d'UML et architecture du logiciel [17]

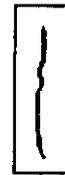


Les films de radiographie et les défauts de soudures associés utilisés dans le calcul d'attributs et la classification qui sont extraits de la base de données des films de radiographie au sein du laboratoire de traitement du signal et de l'image du centre de soudage et contrôle.

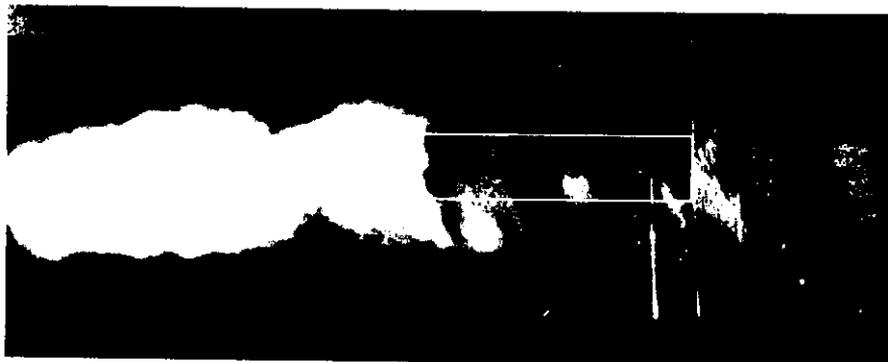
Première classe : On peut associer à cette catégorie les fissures, les caniveaux, les manques de fusion latéraux, etc



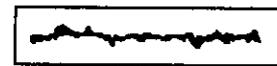
F_1



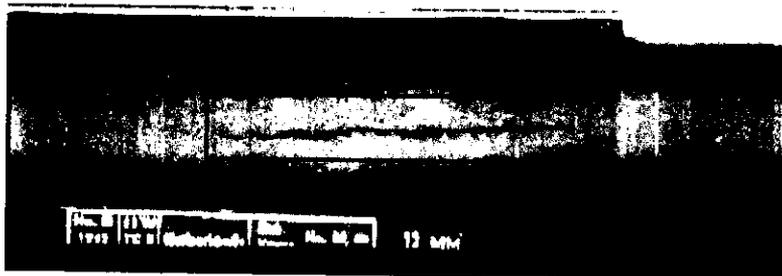
D_F_1



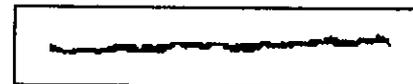
F_2



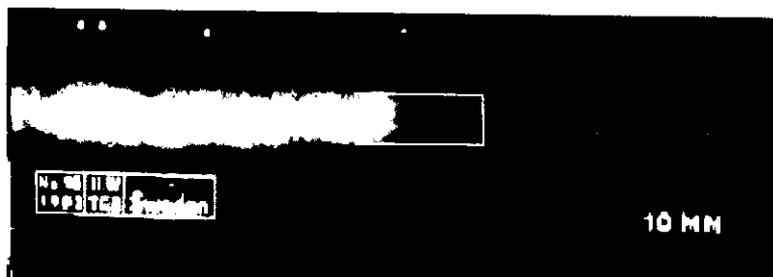
D_F_2



F_3



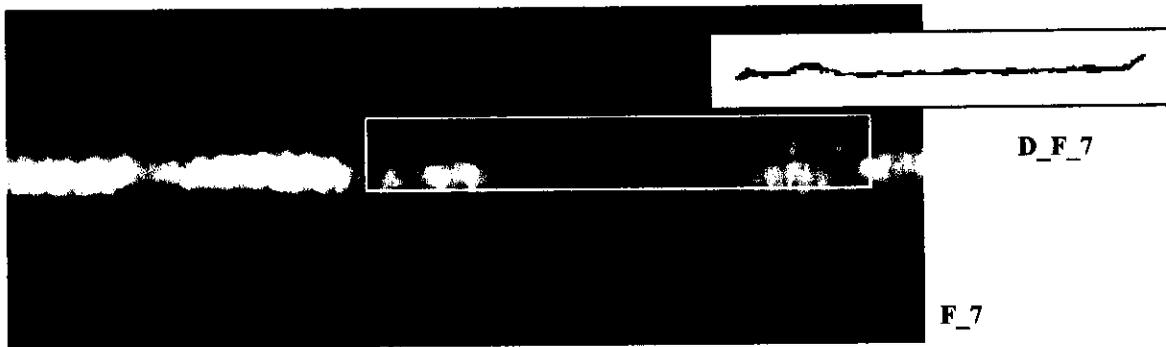
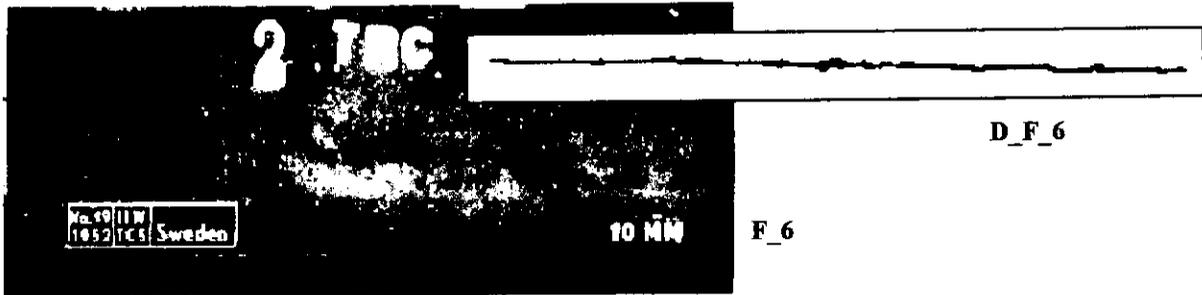
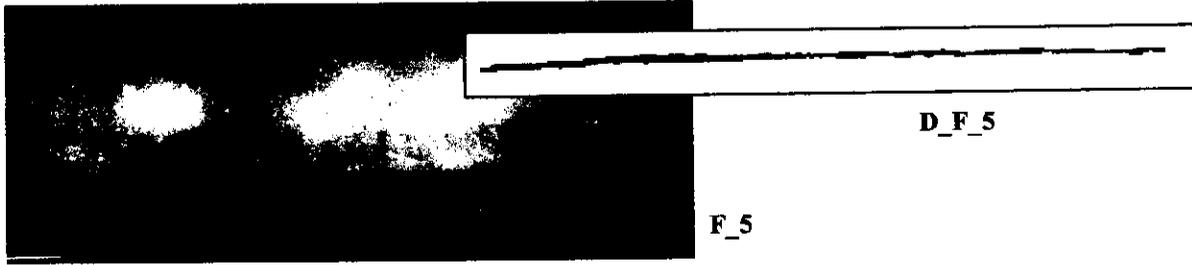
D_F_3



F_4

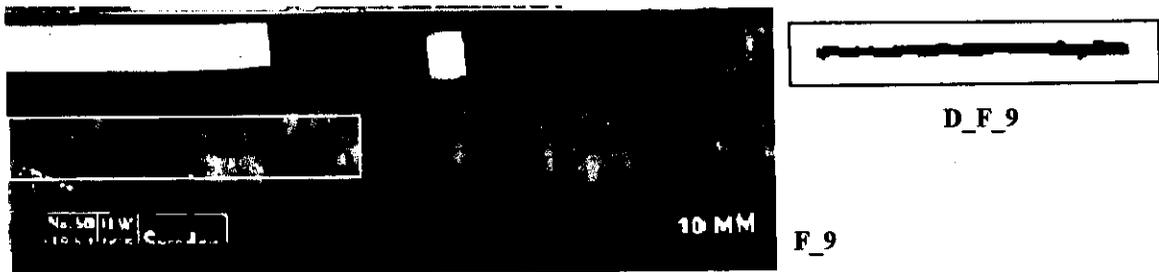


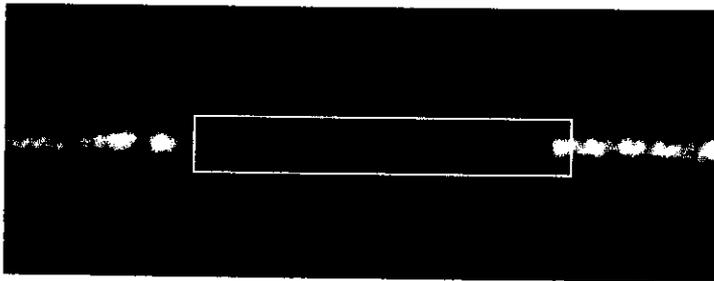
D_F_4



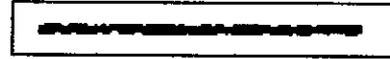
Deuxième classe :

Elle peut comporter le manque de pénétration, les soufflures allongées, etc

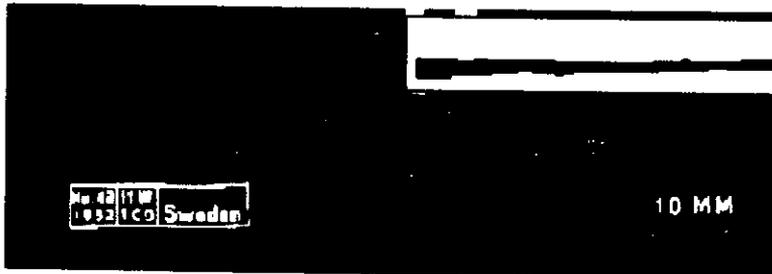




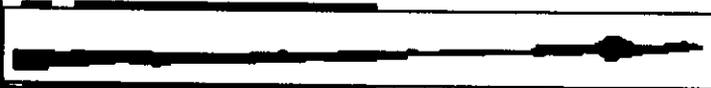
F_10



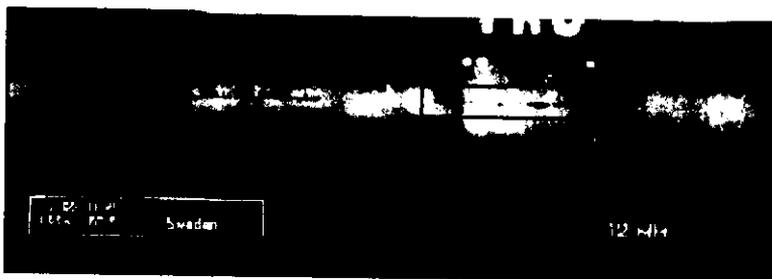
D_F_10



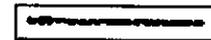
F_11



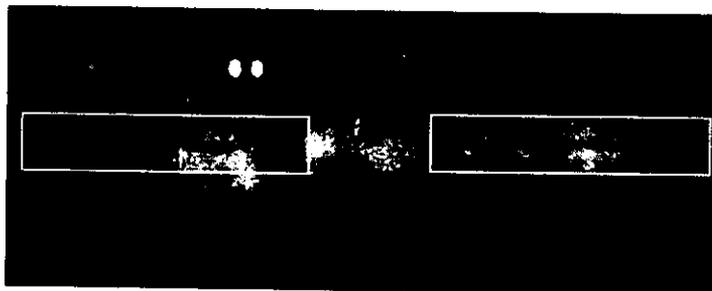
D_F_11



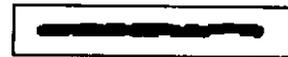
F_12



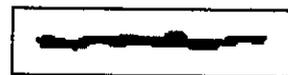
D_F_12



F_13



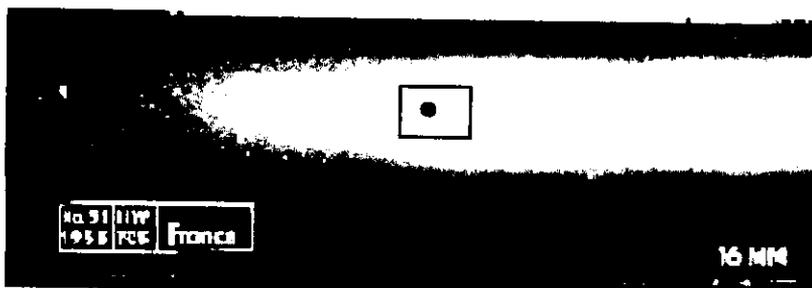
D1_F_13



D2_F_13

Troisième classe :

Elle concerne les porosités, les inclusions de tungstène, etc



F_14



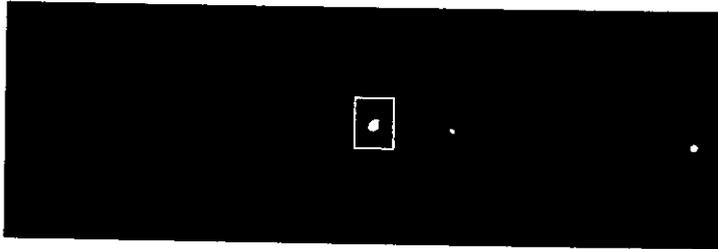
D_F_14



F_15



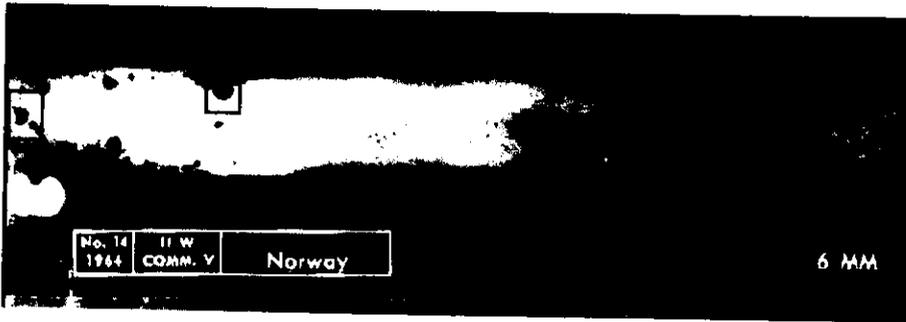
D_F_15



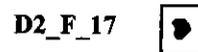
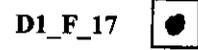
F_16



D_F_16



F_17



D1_F_17

D2_F_17

Quatrième classe :

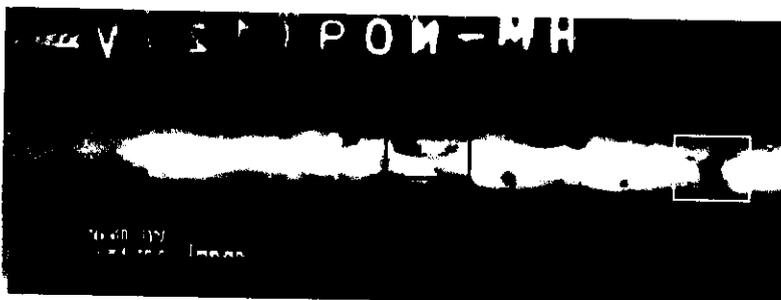
Cette catégorie peut inclure les inclusions de solide, les inclusions de laitier, les soufflures vermiculaires, etc.



F_18



D_F_18

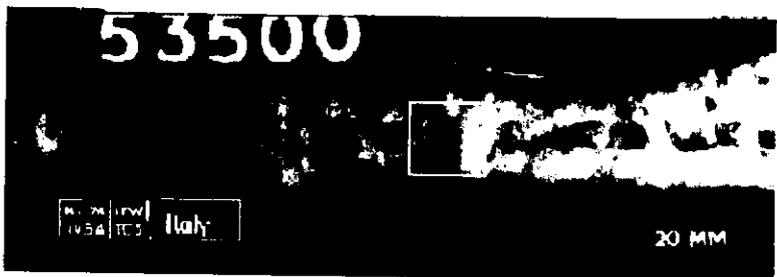


F_19



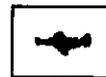
D1_F_19

D2_F_19



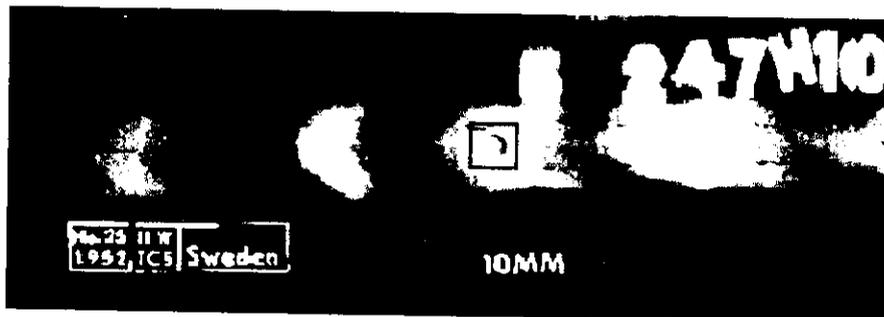
D_F_20

F_20



D_F_21

F_22



D_F_23

F_23



Références bibliographiques :

- [1] L. Mouchat, S. Malek, "Mise en œuvre de classificateurs basés sur la notion de distance et classification statistique. Applications : Identification des défauts de joint de soudure sur les images radiographiques". *Mémoire d'ingéniorat*. Université de Blida. 2002.
- [2] "Classification des défauts dans les soudures par fusion des métaux, avec commentaires explicatifs". *Norme of International Standard Organization*, 6520-1982, 1982
- [3] N. Nacereddine , " Mise au point d'une méthode de détection et classification automatisée des défauts de soudure an radiographie industrielle. *Mémoire de Magistère*. Centre de soudage et Contrôle, Chéraga. 2004
- [4] R. Halmshaw, "Radiographie industrielle". *Agfa-Gevaert N.V.* 1988.
- [5] "Soudages et techniques connexes, Tome 3 : Contrôles et Essais". Recueil des Normes Françaises, *Edition AFNOR*, Paris, 1994.
- [6] S. Castan, "Vision par ordinateur". *Techniques de l'ingénieur, Traité Informatique*, vol.H2, Fasc. H1980, Dec. 1985.
- [7] H. El moussaoui "Analyse Structurale, Etude et Implémentation de l'ACP et l'AFC" , *Mémoire d'Ingéniorat*.
- [8] N. BENAHMED "Optimisation de réseaux de neurones pour la reconnaissance de chiffres manuscrits isolés: sélection et pondération des primitives par algorithme génétiques" *Mémoire d'Ingéniorat*.
- [9] Plamondon Réjean, & Srihari Sargur. (2000). On-Line and Off-Line Handwriting Recognition : A comprehensive Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1), 63-84.
- [10] Abhijit S. Pandya, & Robert B. Macy (1996). *Pattern Recognition with Neural Networks in C++*. Floride: A CRC book published in cooperation with IEEE press.
- [11] Robert Schalkoff (1992). *Pattern Recognition: Statistical, structural and Neural Approaches*. USA: John Wiley and Sons, Inc.
- [12] Miclet L. (1984). *Méthodes structurelles pour la reconnaissance des formes*. Paris : Eyrolles.

- [13] Gaillat G. (1983). *Méthodes statistiques de reconnaissance de formes*. Centre d'édition et de documentation de l'École Nationale de Techniques Avancées.
- [14] Pierre-Yves Cloux, « RUP, XP Architectures et outils 'industrialiser le processus de développement' », Dunod, Paris, 2003.
- [15] Philippe Kruschten, « THE 4+1 VIEW MODEL OF ARCHITECTURE », IEEE, November 1995.
- [16] Rradly Booch, «Software Architecture and the UML», Rational, 1999