HAL
open science

Analyse de formes et de textures : application à l'identification et à l'aide de la gradation des pièces de monnaie

Xingyu Pan

▶ **To cite this version:**

Xingyu Pan. Analyse de formes et de textures : application à l'identification et à l'aide de la gradation des pièces de monnaie. Traitement des images [eess.IV]. Université Lyon 2 Lumière, 2018. Français. NNT : . tel-01916100

HAL Id: tel-01916100

https://hal.science/tel-01916100v1

Submitted on 28 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ LUMIÈRE LYON 2
École Doctorale Informatique et Mathématiques de Lyon

THÈSE

pour obtenir le grade de
DOCTEUR EN INFORMATIQUE

présentée et soutenue par

Xingyu PAN

le 05/06/2018

Analyse de formes et de textures : application à l'identification et à l'aide à la gradation de pièces de monnaie

---

Devant le jury composé de:

Rapportrice : Sylvie TREUILLET, MCF, École Polytechnique de l'Université d'Orléans

Rapporteur : Pascal DESBARATS, PR, Université de Bordeaux

Examinatrice : Véronique EGLIN, PR, INSA-LYON

Examinateur : Alain CROUZIL, MCF, Université Toulouse III

Examinateur : Antoine VACAVANT, MCF, IUT du Puy-en-Velay

Directrice de thèse : Laure TOUGNE, PR, Université Lyon 2

Remerciements

Je tiens à remercier vivement ma directrice de thèse, Professeure Laure Tougne, pour toutes les heures qu'elle a consacrées à me diriger et m'accompagner tout au long de ce travail doctoral. Je lui suis également reconnaissant pour son attention, sa vigueur, sa patience et sa grande gentillesse. Par contre, je ne la remercierai jamais assez d'avoir autant relu et corrigé mon manuscrit en mettant de rouge de partout sur mes phrases originales « trop orales » et « incompréhensibles ». Je n'aurais jamais pu réaliser ce manuscrit sans ses relectures méticuleuses.

Mes remerciements vont aussi à Michel Taillard et Florian Chauvin de la société GENI qui m'ont fait découvrir le monde de la numismatique. Je les remercie amplement pour la confiance qu'ils m'ont accordée, également pour leurs données de qualité et leurs expertises qui m'ont beaucoup aidé pour ce travail de recherche.

Ensuite, j'aimerais remercier tous les membres et les anciens membres du laboratoire LIRIS pour l'ambiance sympathique dans laquelle mon stress du travail a fondu dans la bonne humeur. Mention spéciale à Dominique Maniez, l'un des meilleurs professeurs en français que j'ai jamais eus, qui a relu des chapitres en m'expliquant les règles sophistiquées de l'orthographe, à Catherine pour qui m'a énormément aidé pour la vie au laboratoire ainsi que pendant mon hiver sans chauffage, mon déménagement, *etc.* à Sarah et Mehdi mes potes de bureau qui m'ont aidé, motivé, et appris beaucoup d'expressions françaises que je ne peux pas utiliser dans ce manuscrit. Ce fut un véritable plaisir de travailler avec eux.

Enfin, je remercie mes parents en Chine pour leur soutien, leur amour et leurs encouragements malgré mon éloignement. Ces derniers mots sont pour ma copine Cheng qui remplit ma vie à l'étranger de joie et de bonheur.

Résumé

OBJETS de collection depuis les temps anciens, de nos jours les pièces de monnaie constituent un marché de plus en plus important. L'évaluation par des experts de l'état de conservation des pièces de monnaie, que nous nommons **gradation**, joue un rôle important pour déterminer leur valeur sur le marché. Dans le but de grader des pièces de monnaie de manière efficace et objective, la société GENI collabore avec le laboratoire LIRIS, afin d'automatiser le processus de gradation à partir de photos de pièces de monnaie.

L'objectif principal de cette thèse est de fournir une aide à la gradation des pièces de monnaie à partir des photos de qualité. Le projet est composé de quatre phases : segmentation des monnaies, identification du type monétaire, détection et reconnaissance du millésime et gradation des monnaies.

Dans la première phase, la pièce de monnaie est segmentée de sa photo de manière précise à l'aide d'un modèle paramétrique déformable. Ce dernier permet également d'extraire des caractéristiques de la pièce de monnaie telles que sa taille, son nombre de coins, de pans, *etc.*

Lors de la deuxième phase, nous cherchons dans une base de données le type monétaire de référence correspondant à la pièce de monnaie requête à l'aide de scores de similarité basés sur des graphes. Le premier score se base sur des caractéristiques locales des contours en relief, et le second, qui est semi-global, permet de mettre en évidence des différences de motifs.

La troisième phase concernent la reconnaissance du millésime. Il s'agit d'un sujet difficile car les caractères, dans ce contexte, ont un premier plan de couleur très similaire à l'arrière-plan. Après avoir localisé la zone du millésime et l'avoir découpée en imagettes de chiffres, nous proposons une méthode de reconnaissance de chiffres à l'aide de caractéristiques « topologiques ».

Enfin, concernant la gradation des monnaies, nous proposons une méthode se basant sur une quantification des « éléments inattendus » comme les rayures et les taches. La pièce de monnaie est d'abord recalée sur une monnaie de référence, puis, nous détectons les « éléments inattendus » significatifs sur des zones d'intérêt. Enfin, concernant les « éléments inattendus » ténus difficiles à repérer individuellement, nous détectons les zones granuleuses à l'aide du *Deep Learning*. Le résultat obtenu par cette méthode, proche de

ce que l'expert réalise « à la main », servira d'aide aux numismates.

Mots clefs : pièces de monnaie, segmentation, identification, reconnaissance de chiffres, gradation

Abstract

Coins have been collected and studied since ancient times. Today, coin collection has become a hobby for anyone who wants to participate in. Coin grading is a way of determining the physical condition of a coin to provide an indicator for its market value. For grading coins on a large scale and in a relatively objective way, the GENI company cooperates with the laboratory LIRIS to automate the process by using coin photos.

The main objective of this thesis is to grade coins from well-conditioned photos. The project is composed of four steps : coin segmentation from raw photos, monetary type identification, coin date detection and recognition and, coin grading.

The first step is to extract coins from raw photos with a high precision. With a deformable geometric model, we segment precisely round coins, many-sided coins, wavy-edged coins and holed coins by recognizing their shapes.

The second step consists of coin recognition or monetary type identification. We match the query coin to the most identical type reference by using two similarity scores. The first similarity score is based on local features of relief contours. The second similarity score is a semi-global measure that highlights the difference between relief patterns.

The third step is to detect and recognize coin dates. However, the fact that such characters have the same color as the background makes traditional optical character recognition methods difficult to apply. After extracting the date zone and cropping it into digit images, we propose a learning-free method to recognize those digits by analyzing their “topological” features.

In the last step, the grading process is carried out by quantification of "unexpected elements" such as scratches and dirty marks. The coin to grade is registered to a reference coin. Then, large “unexpected elements” are detected in some regions of interest. However, some “micro-scratches” are difficult to extract individually but all together make a "grainy" surface. To deal with it, we use Deep Learning techniques to classify those grainy zones containing such “micro-scratches”. The result of our system, which is close to the manual expert one, is considered as a useful help for numismatists.

Keywords: coins, segmentation, identification, digit recognition, grading

Table des matières

Introduction générale **1**

1 Contexte numismatique **7**

1.1 Petits lexiques numismatiques 10

1.2 Grade et gradation 14

1.2.1 État de conservation 14

1.2.2 Échelles de gradation 15

1.2.3 Gradation professionnelle 20

1.3 Photographie numismatique 22

1.3.1 Dispositif d'imagerie numismatique 23

1.3.2 Photographies professionnelles 24

1.4 Société GENI 25

1.5 Problématiques de vision par ordinateur appliquées à la numismatique . . 28

2 État de l'art **31**

2.1 Segmentation d'images de monnaies 35

2.1.1 Vue d'ensemble de la segmentation d'images 36

2.1.2 Segmentation des monnaies 39

2.2 Analyse d'images pour l'identification du type monétaire 43

2.2.1 Approches par descripteurs globaux 44

2.2.2 Approches par descripteurs locaux 48

2.2.3 Approches basées sur la lecture des légendes 52

2.3 Analyse d'images pour la gradation des monnaies 54

2.3.1 Brevets et produits industriels de la gradation 54

2.3.2 Travaux de recherche sur la gradation 56

2.4 Conclusion et aspects spécifiques de la thèse 58

3 Bases de données numismatiques **63**

3.1 Base de données d'images de monnaie 65

3.1.1 Bases de données utilisées dans la littérature 66

3.1.2 Bases de données utilisées dans le cadre de cette thèse 69

3.1.3 Comparaison des bases de données numismatiques 74
3.2 Base de données des chiffres extraits des monnaies 76
3.2.1 Méthodes d'extraction 77
3.2.2 Différentes sous-bases de CoinNUMS 82
3.2.3 Comparaison des propriétés visuelles des sous-bases de CoinNUMS 84
3.3 Conclusion 86

4 Segmentation des monnaies 89
4.1 Méthodologie proposée 91
4.1.1 Segmentation « grossière » 92
4.1.2 Estimation des paramètres du modèle 94
4.1.3 Segmentation précise par évolution du modèle actif 100
4.2 Optimisation et adaptation 104
4.2.1 k-estimation 104
4.2.2 Adaptation aux monnaies trouées 105
4.3 Expériences et résultats 106
4.3.1 Expériences sur *LIRIS_CoinPhotos* 107
4.3.2 Expériences sur *GENI_CoinPhotos* 108
4.4 Conclusion 110

5 Identification du type monétaire 115
5.1 Méthodologie proposée 117
5.1.1 Score de similarité à l'aide de correspondances locales 118
5.1.2 Mise en correspondance des points d'intérêt 120
5.1.3 Score de similarité semi-global basé graphe 125
5.2 Expériences et discussion 130
5.2.1 Expériences sur une sous-base de *USA_Grading* 130
5.2.2 Expériences sur USA_Grading 131
5.3 Conclusion 135

6 Reconnaissance des chiffres extraits des monnaies 137
6.1 Méthodologie proposée 140
6.1.1 « Topologies » des chiffres 141
6.1.2 Système de votes 144
6.1.3 Graphes de Reeb 145
6.2 Expériences 149
6.2.1 Bases de données 150
6.2.2 Résultats et discussions 151
6.3 Conclusion 153

7 Gradation des monnaies 157
7.1 Méthodologie proposée 159
7.1.1 Recalage des monnaies 162
7.1.2 Construction des monnaies de référence 163
7.1.3 Détection des « éléments inattendus » significatifs 166
7.1.4 Détection des « éléments inattendus » ténus 167
7.1.5 Gradation résultante 173
7.2 Expériences et discussions 175
7.2.1 Détection des « éléments inattendus » significatifs 175
7.2.2 Classification des imagettes annotées 176
7.3 Conclusion 181

Conclusion générale 183

Bibliographie 193

Introduction générale

PASSE-TEMPS des rois depuis la Renaissance, la collection de pièces de monnaie est devenue de nos jours une activité qui s'est démocratisée, avec l'essor du marché international, même si certains collectionneurs sont prêts à payer un prix considérable pour obtenir leurs pièces de monnaie favorites. Hormis le métal précieux contenu dans la pièce de monnaie et la rareté de ces dernières, le prix d'une pièce de monnaie sur le marché dépend principalement de son état de conservation. Un exemplaire bien conservé coûte souvent plusieurs fois le prix d'un exemplaire usé. Ainsi, les numismates ont créé des systèmes pour évaluer l'état de conservation des pièces de monnaie. Cette évaluation est principalement basée sur les reliefs restants, les rayures, les coups, *etc.* La patine ainsi que l'aspect esthétique global sont également pris en considération. L'état de conservation évalué, appelé grade de la pièce de monnaie, sera résumé par une description, par exemple, TTB (Très Très Beau), ou par un score numérique, par exemple sur une échelle de 1 à 70 dans le cadre de l'échelle de Sheldon. Ce processus de l'évaluation de l'état de conservation des pièces de monnaie s'appelle la gradation. Cependant, malgré des critères généraux expliqués dans les systèmes de gradation utilisés, l'attribution d'un grade précis à une pièce de monnaie reste subjective et varie d'une personne à une autre. Pour résoudre ce problème de standardisation des grades attribués, des sociétés de gradation professionnelles ont émergé dans les années 1980 afin de proposer une soi-disant gradation impartiale. Le grade d'une pièce de monnaie est considéré certifié si plusieurs experts d'une société de gradation donnent les mêmes avis. Cependant, les grades attribués à une même pièce de monnaie par différentes sociétés sont rarement identiques. Par ailleurs, le coût associé à la gradation traditionnelle est élevé en considérant l'intervention des experts et le transport sécurisé, notamment pour grader un grand volume des pièces de monnaie lors d'une session de gradation.

La société Gradations et Expertises Numismatiques Internationales (GENI) est une jeune société numismatique française qui est spécialisée dans l'authentification et l'expertise de pièces de monnaie. Elle est aussi experte dans le domaine de la photographie numismatique. Comparés à d'autres applications qui profitent des derniers développements informatiques, comme la reconnaissance de visages, la plupart des services de la numismatique professionnelle s'effectuent encore de façon « manuelle ». Aussi, l'automatisation de ces processus aura-t-elle un grand l'intérêt pour la société GENI. Pour cela, la société GENI a fait appel à l'expertise du laboratoire LIRIS (Laboratoire d'InfoRmatique en Image et Systèmes d'information) en traitement d'images en finançant cette thèse. Une autre motivation de cette thèse est de réaliser une gradation des pièces de monnaie de manière plus objective et plus efficace en utilisant des photos de qualité.

L'objectif principal de cette thèse est, d'une part, de reconnaître automatiquement les pièces de monnaie par analyse d'images et d'autre part, de fournir une aide à leur gradation. Pour ce faire, à partir de photos, un détourage très précis est réalisé. Ce dé-

tourage n'est pas d'un grand intérêt pour la reconnaissance mais il l'est pour la société GENI qui souhaite utiliser ces images précisément détourées pour réaliser des catalogues. Notons toutefois que cette phase de segmentation permet également d'extraire des caractéristiques de forme de la pièce de monnaie qui seront utilisées pour la reconnaissance. Ensuite la reconnaissance proprement dite s'effectue en deux temps : reconnaissance globale et travail sur les millésimes.

Notons que les algorithmes développés dans cette thèse sont spécifiques aux pièces de monnaie dont les photos sont prises dans des conditions bien contrôlées.

Au niveau scientifique, nos contributions portent donc sur la segmentation d'objets de manière précise, la reconnaissance d'objets ayant des similarités interclasses élevées, la reconnaissance de caractères de monnaies pour particularité d'être de la même couleur que l'arrière-plan et la détection des « éléments inattendus » tels que des rayures, des coups, des taches, *etc.* Cette thèse est composée de sept chapitres.

Le Chapitre 1 aborde le contexte de la gradation des pièces de monnaie, la photographie numismatique et présente la société numismatique GENI qui a financé cette thèse.

Le Chapitre 2 traite de la bibliographie portant sur des approches du traitement d'images, notamment la segmentation et la reconnaissance d'objets, qui prennent les pièces de monnaie comme objets d'intérêt.

Le Chapitre 3 concerne les bases de données numismatiques les plus connues utilisées dans les études précédentes et les bases de données que nous avons construites au cours de cette thèse pour développer et valider les méthodes dans chaque étape de la thèse.

Le Chapitre 4 décrit la première phase de la segmentation précise des pièces de monnaie en reconnaissant leur forme. Cette étape servira pour les suivantes mais répond principalement à une demande de l'entreprise qui souhaite des illustrations parfaites afin de les intégrer dans des catalogues de vente.

Le Chapitre 5 présente la deuxième phase de l'identification automatique du type monétaire par comparaison avec une base de référence à l'aide de scores de similarité robustes pour les types monétaires ayant des similarités interclasses élevées.

Le Chapitre 6 se focalise sur la reconnaissance des chiffres extraits des monnaies qui possèdent un premier plan difficile à segmenter. Cette étape permet de reconnaître les millésimes afin de distinguer les types monétaires ayant des motifs extrêmement similaires.

Le Chapitre 7 expose la dernière phase du travail qui concerne une aide à la gradation des pièces de monnaie en s'appuyant sur une quantification des « éléments inattendus » en comparant la monnaie à grader avec une monnaie de référence.

Cette thèse a fait l'objet de plusieurs publications d'audience internationale. Nos nouveaux algorithmes de segmentation des pièces de monnaie et d'identification du type monétaire ont été présentés lors d'une conférence internationale ISVC (*International Symposium on Visual Computing*) en 2014 (Pan, Puritat et Tougne 2014). Par ailleurs, notre

travail concernant la reconnaissance des chiffres extraits des pièces de monnaie a été présenté à la conférence internationale ICIAP (*International Conference on Image Analysis and Processing*) en 2016 (Pan et Tougne 2016). Nous avons ensuite proposé une étude étendue de ce travail ainsi qu'une nouvelle base de données de chiffres extraits des monnaies, l'ensemble de ce travail ayant été publié dans une revue internationale (*Frontiers in ICT*), puis mis à la disposition de la communauté (Pan et Tougne 2017). Par ailleurs, nous allons présenter notre travail sur la gradation des monnaies à la conférence internationale IVPAI (*International Conference on Image Video Processing and Artificial Intelligence*) qui aura lieu en août 2018 (Pan et Tougne 2018).

Chapitre 1

Contexte numismatique

De tout temps, l'homme a collectionné des pièces de monnaies. Les études archéologiques indiquent que depuis l'époque de la Rome antique et la Mésopotamie, des pièces de monnaie étaient collectionnées et catégorisées à des fins de thésaurisation. D'après Suetonius dans son *De vita Caesarum* (La vie de douze Césars), l'empereur Auguste présentait des pièces de monnaie anciennes et exotiques à ses amis lors d'occasions spéciales. Pendant l'époque de la Renaissance, la collection des pièces de monnaie, considérée comme le passe-temps des rois, devenait la mode pour des classes privilégiées. Parmi les collectionneurs célèbres, nous pouvons citer Pétrarque, les Médicis, le pape Paul III, Henri IV, Louis XIV, *etc.* Aux XVIIe et XVIIIe siècles, ce passe-temps des rois s'est poursuivi et est devenu celui des classes moyennes qui avaient envie de montrer leur richesse et leur sophistication. Du XIXe siècle jusqu'à aujourd'hui, la collection de pièces de monnaie est devenue un loisir de plus en plus populaire.

Avec l'augmentation du nombre de collectionneurs et du marché correspondant, de nombreux intermédiaires ont émergé répondant aux besoins : boutiques de monnaies, enchères, musées et expositions, bourses (*cf.* Figure 1.1), salons de collectionneurs, catalogues et guides de monnaies, *etc.* Par ailleurs, les instituts monétaires officiels (Monnaie de Paris (*cf.* Figure 1.2), *United States Mint*, *Royal Mint*, *Swissmint*, *etc.*) émettent chaque année un nombre plus ou moins important de pièces de monnaie destinées uniquement à la collection. Malgré la popularité et la croissance de la collection de monnaies, le nombre d'experts qui maîtrisent l'évaluation et la protection des monnaies reste faible par rapport à la population de collectionneurs amateurs. Ainsi, des sociétés privées réunissant des experts en monnaies se sont lancées dans des initiatives de commercialisation de leur expertise, en proposant des services standardisés dans l'archivage, l'authentification, la protection, mais également l'évaluation de l'état de conservation des pièces de monnaie que nous allons présenter ultérieurement dans ce chapitre.

Figure 1.1 – Bourse numismatique à Paris, le 12 Octobre 2013

FIGURE 1.2 – Monnaie de Paris, institut monétaire national de la France

La numismatique est une science vaste qui a pour objet l'étude des monnaies, médailles et jetons, mais dans cette thèse nous n'abordons pas les aspects artistiques, culturels, techniques, historiques et archéologiques des études numismatiques. Le terme numismatique ici, pris au sens strict, fait plutôt référence à la collection des monnaies et aux marchés associés. Parmi toutes les formes monétaires (les pièces de monnaie, billets, médailles, jetons, monnaies archaïques, lingots, *etc.*), nous nous concentrons en particulier sur les pièces de monnaie (y compris les jetons et les pièces commémoratives). Par simplicité, nous utiliserons le terme **MONNAIE** à la place des différentes formes de pièces de collection dans cette thèse. Plus précisément, nous nous concentrons plutôt sur les monnaies « modernes », autrement dit celles frappées à la presse monétaire. Nous ne considérons pas les monnaies anciennes frappées manuellement à l'aide d'un marteau, ou celles frappées avec un balancier.

Dans ce chapitre, nous présentons le contexte et les notions numismatiques utiles à la compréhension de cette thèse.

1.1 Petits lexiques numismatiques

La communauté numismatique a développé un vocabulaire bien riche mais souvent inconnu de tout un chacun. Nous commençons ainsi par présenter quelques notions numismatiques fréquemment utilisées dans le reste de cette thèse.

Différents objets de collection relatifs à une pièce de monnaie

Coin : morceau de métal, généralement en acier et cylindrique, sur lequel est gravée en creux les motifs d'une monnaie (*cf.* Figure 1.3).

Essai : monnaie frappée en très faible quantité avec des coins neufs et destinée à présenter un nouveau type monétaire. L'Essai est reconnaissable par la lettre E ou la légende « ESSAI » en majuscule inscrite (*cf.* Figure 1.4 (a)).

Pièce commémorative : monnaie dont la frappe est motivée par un événement spécifique. En général elle sert à des fins de collection, mais elle peut également être mise en circulation (*cf.* Figure 1.4 (b)).

Jeton : pièce monétiforme aux fonctions multiples comme moyen de paiement temporaire, jeton de jeux, jeton de téléphone, *etc.* (*cf.* Figure 1.4 (c)).

FIGURE 1.3 – Coins des monnaies en euro de la Finlande

(a) (b) (c)

FIGURE 1.4 – Différentes pièces de collection : (a) Essai 5 Francs Napoléon II (légende « ESSAI » encadrée en rouge) ; (b) 2 Euros *D-Day* commémorative ; (c) jeton de jeux

Description d'une monnaie (*cf.* Figure 1.5)

Avers : côté de la monnaie sur lequel est placé le motif essentiel, le plus souvent un symbole de l'autorité émettrice (profil, blason, *etc.*). En langage courant, il s'agit du côté « face ».

Revers : côté de la monnaie sur lequel est souvent placée la valeur faciale. En langage courant, il s'agit du côté « pile ».

Tranche : pourtour circulaire constitué par l'épaisseur de la monnaie. Elle est parfois lisse, striée, cannelée ou portant des inscriptions.

Module : diamètre d'une monnaie. Mais en langage courant et scientifique, nous préférons utiliser le terme diamètre.

Type monétaire : ensemble des motifs qui caractérisent chaque côté d'une monnaie et la tranche. Dans les catalogues de monnaies, un type correspond à un numéro. En langage scientifique, un type monétaire est une classe de monnaie.

Valeur faciale : valeur pour laquelle la monnaie a été mise en circulation. Elle est indiquée habituellement au revers mais cette règle n'a pas toujours été respectée.

Millésime : date d'émission d'une monnaie, souvent gravé sur un côté de la monnaie. Il ne correspond pas forcément à l'année de frappe effective notamment pour les monnaies « refrappées ».

Différent : marque personnelle apposée sur une monnaie par le graveur ou le directeur d'un atelier monétaire.

Grènetis : série de petits grains relevés en bosse au bord de la monnaie.

Listel : rebord de l'avers et du revers d'une pièce de monnaie destinée à empêcher le rognage ou l'usure rapide des reliefs.

Champ : partie plate de la monnaie sur laquelle aucun motif n'a été gravé.

Apparence d'une monnaie

Artefact ou fauté : erreur à la fabrication.

Variétés et variantes : dans un type monétaire, les variantes qui correspondent aux différents coins sont intentionnelles alors qu'une variété est une différence non intentionnelle (*cf.* Figure 1.6).

Patine : coloration naturelle des surfaces de la monnaie souvent liée à l'oxydation dans le temps (*cf.* Figure 1.8).

Coups de sac : chocs nombreux d'aspect bien particuliers causés par les chocs entre les monnaies stockées et transportées en sac.

Velours de frappe : la frappe des coins se traduit par l'apparition de ridules à la surface d'une monnaie en lui donnant un aspect « satiné » mat. Il diminue au fur et à mesure par la circulation.

État de conservation (*Grade*) : niveau d'usure d'une monnaie.

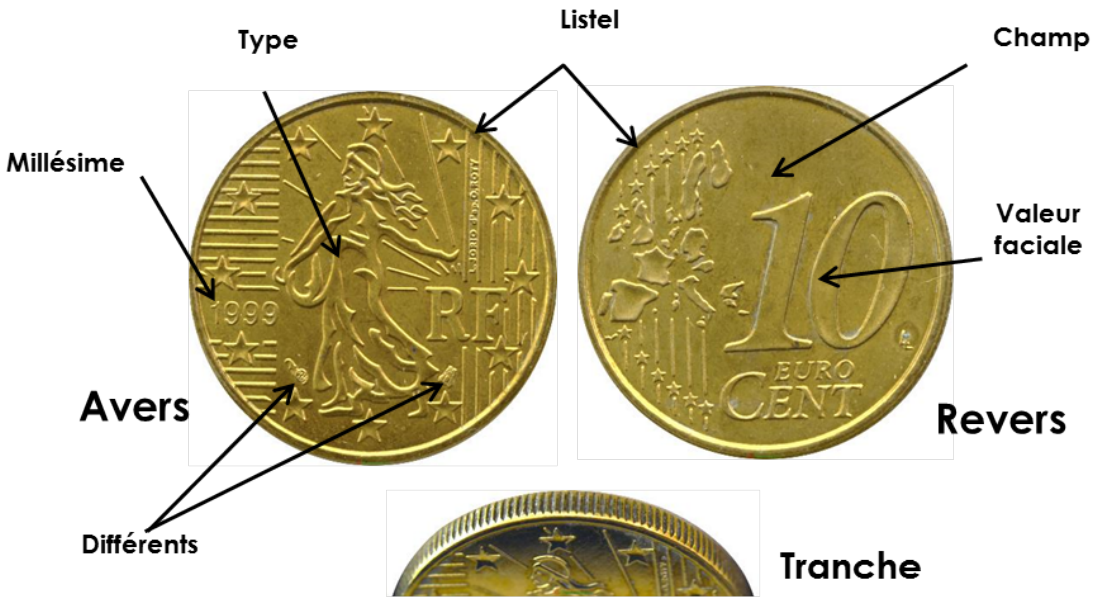


FIGURE 1.5 – Principaux éléments d'une monnaie

Gradation (*Grading*) : évaluation de l'état de conservation d'une monnaie et l'attribution d'un grade.

Fleur de Coin (FDC) : ce terme très ancien définit à la fois la qualité d'un coin qui doit être neuf, et plutôt l'état de conservation d'une monnaie qui doit être parfait et ne présenter aucun défaut visible.

Belle épreuve (BE) ou Flan Bruni (FB) : qualité la plus élevée qui caractérise une monnaie dont les motifs sont mats et le fond est comme un miroir (*cf.* Figure 1.7).

Brillant Universel (BU) : qualité intermédiaire entre la frappe d'une monnaie courante et celle d'une « belle épreuve ».

FIGURE 1.6 – Variantes d'une 2 euros commémorative avec les étoiles de différentes tailles

FIGURE 1.7 – Exemples de monnaies « belles épreuves »

(a) (b)

FIGURE 1.8 – Exemples de monnaies *US Washington Quarter Dollar* : (a) sans patine ; (b) avec patine

1.2 Grade et gradation

1.2.1 État de conservation

La valeur d'une monnaie de collection n'a souvent rien à voir avec sa valeur faciale. Elle est déterminée par de nombreux facteurs comme le cours des métaux dont elle est composée, la rareté ou simplement les fluctuations de la demande sur le marché. Parmi tous les facteurs, l'état de conservation joue un rôle essentiel. L'état de conservation caractérise l'apparence de la monnaie : meilleur est l'état de conservation d'une monnaie, plus cette monnaie se rapproche de son état de frappe originel. Bien évidemment, les monnaies dans un état de conservation parfait sont assurées d'avoir une valeur significative sur le marché. La Table 1.1 présente un exemple des prix estimés de la 10 Francs Cérès extraits d'un ouvrage de numismatique (Desrousseaux et Prieur 2014). Dans cette table, B, TB, TB, SUP, SPL et FDC sont des symboles correspondant à des paliers de l'état de conservation rangés de manière croissante (*cf.* Table 1.2). Nous pouvons facilement trouver une corrélation négative entre la cote et la quantité émise par millésime, et une corrélation positive entre la cote et l'état de conservation. Un exemplaire de 10 Francs Cérès émise en 1878 dans un état de frappe parfait (FDC) se vend une véritable fortune. Récemment, une 10 Francs Cérès évaluée par la société GENI a atteint la mesure de 18100€ + 24 % de frais dans la vente aux enchères MDC qui s'est tenue à Monaco le 1 Décembre 2017. Par ailleurs, les e-boutiques de monnaies de collection témoignent que les

monnaies dans un très bon état de conservation réalisent des prix de vente au-delà des espérances ; au contraire, les monnaies pourtant rares mais dans un état de conservation médiocre atteignent difficilement leur prix estimé. Par conséquent, l'évaluation de l'état de conservation des monnaies est devenue un enjeu majeur sur le marché de la numismatique.

Par simplicité, nous utiliserons les termes **GRADE** et **GRADATION** pour parler de l'état de conservation et de son estimation dans la suite de cette thèse. Ce sont deux termes courants dans la communauté numismatique internationale mais aussi connus par les numismates français.

Table 1.1 – Catalogue des prix (€) pour 10 Francs Cérès

	Date	**Quantité**	**B**	**TB**	**TTB**	**SUP**	**SPL**	**FDC**
1	1878	30	–	–	–	6000	8500	12000
2	1889	100	–	–	–	4800	7000	9000
3	1895	213.999	-	-	140	280	600	1100
4	1896	585.010	-	-	130	230	500	900
5	1899	1.600.000	-	-	120	200	450	800

- : sans autre valeur marchande dans cet état
– : n'est pas supposé exister dans cet état

1.2.2 Échelles de gradation

Le grade est un terme technique qui mesure l'état de conservation, et qui peut être comparé pour classifier l'apparence d'une monnaie. Très souvent, il est interprété comme un niveau d'usure, mais il est difficile de trouver une définition exacte qui comporte tous les éléments visuels contribuant au grade affecté. Certains experts disent que le grade d'une monnaie est déterminé par les facteurs suivants :

- l'intérêt esthétique (*eye appeal*),
- les motifs en relief restants,
- la patine,
- le nombre et profondeur des rayures,
- le nombre de coups,
- la tranche,
- *etc.*

Par ailleurs, certains experts disent que les défauts comme la frappe faible, défauts dans le métal, défaut sur la tranche, *etc.*, doivent être signalés séparément mais ne participent pas à l'estimation du grade (Desrousseaux et Prieur 2014).

La gradation d'une monnaie est plus un art qu'une science, au cours des années, la gradation des monnaies a évolué vers un système avec des échelles fixées et des critères de plus en plus standardisés.

Premiers systèmes de gradation

Dans les premiers temps, l'état de conservation d'une monnaie était simplement catégorisé par « neuf » et « usé ». Pour différencier la qualité des monnaies, les collectionneurs commençaient à utiliser trois termes pour décrire l'état de conservation d'une monnaie de manière générale :

- *Good* : la surface s'est usée pendant la circulation, mais certains détails des motifs en relief restent visibles.
- *Fine* : le motif en relief est peu usé.
- *Uncirculated* : le motif en relief est net et la frappe s'approche de l'état d'origine.

Ce système contenant seulement trois échelles est considéré comme un système basique de la gradation. Dans ce système, le grade est uniquement déterminé par les motifs en relief. Les défauts éventuels sur la surface ne sont pas considérés. Le manque de grade inférieur à *good* est probablement causé par le fait que les collectionneurs ne s'intéressaient pas aux monnaies trop usées.

Systèmes européens

En Europe, les numismates utilisent souvent une échelle de gradation de 5 à 8 niveaux. Bien que ces systèmes varient sur chaque marché national, la plupart d'entre eux utilisent des critères et une terminologie similaire. Dans les systèmes européens, les échelles sont purement descriptives comme dans les premiers systèmes que nous avons présentés.

La Table 1.2 présente le système de gradation en France. Il est exprimé en 7 niveaux : Assez Beau (AB), Beau (B), Très Beau (TB), Très Très Beau (TB), Superbe (SUP), Splendide (SPL) et Fleur de Coin (FDC) pour la plupart des monnaies. Des monnaies gradées suivant ces 7 niveaux sont illustrées à la Figure 1.9. Belle épreuve (BE), absents de cette table, correspondent au huitième niveau qui est destiné aux pièces épreuves. L'attribution d'un grade est principalement basée sur le pourcentage des motifs en relief restants, mais également en considérant d'autres facteurs comme la patine, les légendes, les chocs, les rayures, les velours de frappe, *etc.* La Table 1.3 récapitule les termes équivalents employés dans les différents pays européens. Le grade est présenté par un terme adjectif. Par exemple, nous utilisons AB en Français, équivalent à *Good* en anglais et GE (*Gut Erhalten*) en allemand, pour décrire une monnaies ayant 10% motifs restants.

Deux constats supplémentaires semblent intéressants. Premièrement, FDC, l'abréviation de Fleur de Coin (ou *Fior di Conio* en italien ou *Flor de Cuño* en espagnol) est utilisée presque comme un terme universel pour décrire une monnaie « parfaite », y compris pour les pays parlant d'autres langues. Secondement, les numismates en Scandinavie préfèrent des échelles numériques plutôt que des échelles descriptives, bien que le nombre d'échelles et les critères soient presque identiques à ceux de leurs voisins européens.

Malgré le fait que les systèmes européens soient encore largement utilisés aujourd'hui, les numismates trouvent souvent que le nombre d'échelles n'est pas suffisant, surtout lorsque les critères sont en contradiction. Une solution consiste à créer plus d'échelles que dans le système classique, par exemple utiliser TB+ ou TTB+. Au Royaume-Uni, la tendance est à développer une échelle à 100 niveaux.

Table 1.2 – Système de gradation français à 7 échelles

État	**Relief**	**Identif.**	**Légendes**	**Chocs**	**Rayures**	**Velours de frappe**
AB	10%	difficile	illisibles	oui	visibles (œil)	non
B	25%	oui	difficiles	oui	visibles (œil)	non
TB	50%	oui	lisibles	oui	visibles (œil)	non
TTB	75%	oui	lisibles	oui	visibles (œil)	non
SUP	90%	oui	lisibles	non	visibles (loupe)	léger
SPL	100%	oui	lisibles	non	absentes	sauf points hauts
FDC	100%	oui	lisibles	non	absentes	oui

Table 1.3 – Termes équivalents dans les systèmes de gradation européens

France	**U.K.**	**Espagne**	**Portugal**	**Italie**	**Allemagne**	**Scandinavie**
AB	G	RC	REG	M	GE	2
B	VG	BC	MREG	B	SGE	1-
TB	F	BC+	BC	MB	S	1
TTB	VF	MBC	MBC	BB	SS	1+
SUP	EF/XF	EBC	Bela	SPL	VZ	01
	UNC				UNZ	01/0
SPL	UNC	SC	Soberba	SPL/FDC	UNZ	0
FDC	FDC	FDC	FDC	FDC	STGL	0

Figure 1.9 – Exemples de monnaies gradées suivant l'échelle française

Système américain

À la fin du XIXe siècle et au début du XXe siècle, la communauté de la numismatique nord-américaine s'est rendue compte de la nécessité de développer un système de gradation

plus précis. En 1948, un numismate célèbre, Dr. William H. Sheldon (*cf.* Figure 1.10) a mis au point un nouveau système de gradation en utilisant une échelle à 70 niveaux dans son ouvrage *Penny Whimsy* (Sheldon, Paschal et Breen 1958), lors d'une étude des *early large cents*. Ce système numérique est maintenant reconnu par son nom, l'échelle de Sheldon.

Dans la version originelle de l'échelle de Sheldon, une monnaie *large cent* à peine identifiable, « *identifiable and unmutilated, but so badly worn that only a portion of the legend or inscription is legible* », est décrite par le terme état*Basal*, ou *Poor* dans un langage courant, et se voit attribuer le grade 1 (*cf.* Figure 1.11).

À partir de cela, une monnaie dont la valeur mérite deux fois plus qu'une monnaie à grade 1 se voit attribuer le grade 2. Théoriquement, une monnaie au grade 70, le grade le plus haut pouvant être attribué, a un état 70 fois meilleur qu'une monnaie dans l'état *Basal*.

Il faut noter que dans cette échelle, nous n'utilisons pas tous les nombres entre 1 et 70, mais seuls 30 niveaux numériques sont attribués (*cf.* Table 1.4). Le fait d'utiliser des valeurs numériques rend l'échelle de Sheldon plus « scientifique » que les autres échelles de gradation. Le grade complet est alors donné par l'abréviation d'un grade traditionnel descriptif plus un grade numérique, par exemple, PO (*Poor*)-1, F (*Fine*)-12 ou MS (*Mint State*) 70.

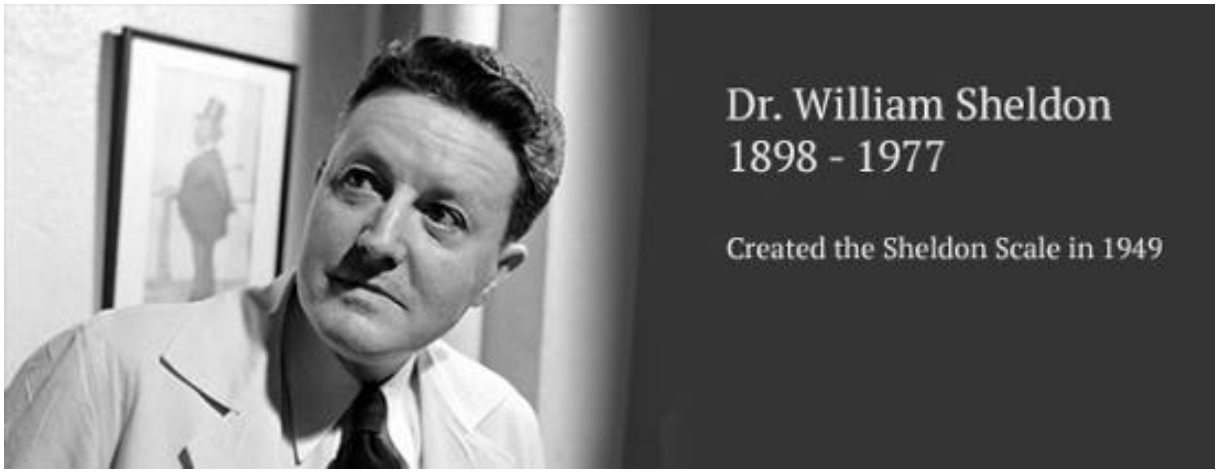


FIGURE 1.10 – Dr. Willam H. Sheldon, inventeur de l'échelle de Sheldon

À partir de 1970, dans le but de s'adapter à toutes les monnaies américaines, l'échelle de Sheldon a été amélioré par *American Numismatic Association* (ANA). Dans l'échelle de Sheldon d'aujourd'hui, les niveaux inférieurs à 60 correspondent à des monnaies en circulation, et les niveaux supérieurs entre 60 et 70 correspondent à des monnaies de collection dans l'état de frappe (*Mint State* (MS)), autrement dit, celles qui n'ont pas été mises en circulation. Des monnaies « belles épreuves » sont émises à des fins de collection. Une échelle distincte destinée à cette catégorie de monnaies de collection n'utilise que certains niveaux de 60 à 70.

Concernant la rareté d'une monnaie, Dr. Sheldon a mis en place une autre échelle qui

FIGURE 1.11 – Monnaie *Flowing hair 1 Cent* au grade PO-1 (*basal state*)

mesure la rareté d'un type monétaire. Cette échelle de rareté comporte huit niveaux de R1 à R8, dans lesquels R1 signifie les monnaies à gros tirage (supérieur à 10000 exemplaires), alors que R8 signifie les millésimes rassismes (inférieur ou égal à 3 exemplaires) (*cf.* Table 1.5). Cependant, la rareté de la monnaie, bien qu'elle soit déterminante pour la valeur d'une monnaie de collection sur le marché, n'est pas importante dans cette thèse qui se concentre principalement sur l'apparence de la monnaie.

Grâce à ces nombreux avantages, l'échelle de Sheldon a été rapidement acceptée partout dans le monde. Elle est considérée comme une échelle universelle sur le marché international. Cependant, habitués à leurs systèmes nationaux, les numismates européens utilisent souvent un nombre réduit de niveaux de l'échelle de Sheldon dans la pratique. Les grades numériques de Sheldon sont rajoutés aux grades descriptifs européens par correspondance approximative.

TABLE 1.4 – Echelle de Sheldon de gradation

Grade descriptif	**Grade numérique**
Mint State (MS), *Uncirculated* (UNC)	60-70
Almost Uncirculated (AU)	50, 53, 55, 58
Extremely Fine (XF or EF)	40, 45
Very Fine (VF)	20, 25, 30, 35
Fine (F)	12, 15
Very Good (VG)	8, 10
Good (G)	4, 6
About Good (AG)	3
Fair (Fr)	2
Poor (P)	1

TABLE 1.5 – Echelle de Sheldon de rareté

Indice	État de rareté	Effectif connu
R1	Courante	>10.000
R2	Assez courante	1000-9999
R3	Peu courante	200-999
R4	Très peu courante	76-199
R5	Rare	31-75
R6	Très Rare	13-30
R7	Extrêmement Rare	4-12
R8	Insigne rareté	1-3

Critères spécifiques à certains type monétaires

Certains numismates n'ont pas été convaincus par un système de gradation universel qui peut s'appliquer à toutes les monnaies. Pour ces personnes, chaque type monétaire est un cas spécifique qui doit être analysé individuellement. Concernant le système de gradation basé sur le pourcentage de motifs en relief préservés, le motif essentiel par exemple le portrait du roi, est beaucoup plus important que la légende et d'autres motifs décoratifs. Dans cette situation, certains numismates proposent des critères spécifiques à certains types monétaires qu'ils connaissent très bien, au lieu d'utiliser des critères généraux (*cf.* Table 1.6).

1.2.3 Gradation professionnelle

Même basée sur une échelle affinée et détaillée comme celle de Sheldon, les critères descriptifs associés à un grade restent plutôt un guide indicatif qu'un standard. En effet, pour une monnaie donnée, les grades donnés par deux numismates expérimentés peuvent varier. Lors des bourses numismatiques, les revendeurs et acheteurs ne sont pas toujours d'accord sur le grade d'une monnaie. Ils ont besoin d'un expert qui peut donner un grade accepté par les deux parties.

Dans les années 1980, cette problématique a enfin conduit à l'apparition de sociétés de gradation. Après plus de 30 ans de compétitions, les quatre sociétés américaines,

- *Professional Coin Grading Service* (PCGS),
- *Numismatic Guaranty Corporation* (NGC),
- *Independent Coin Graders* (ICG),
- *American Numismatic Association Certification Service* (ANACS),

sont considérées comme les sociétés de gradation professionnelle les plus dignes de confiance sur le marché aux États-Unis, voire dans le monde entier (*cf.* Figure 1.12).

Dans ces sociétés, une monnaie à grader doit être évaluée et validée par plusieurs gradeurs expérimentés jusqu'à ce que leurs avis s'accordent. Par ailleurs, la gradation profes-

Table 1.6 – Critères de gradation destinés aux monnaies spécifiques

Grade	5 Francs Hercule	10 Francs Coq
TB	La léonté d'Hercule complètement visible et le tissu horizontal séparant le haut du bas de la robe de l'Égalité se distingue encore.	Les cheveux doivent être séparés du front.
TTB	Tout le dessin est complet, sauf la cuisse et l'épaule de l'Egalité.	Le contour de la feuille de chêne verticale sur la tempe ne doit pas se confondre avec le bonnet phrygien.
SUP	Très légères traces de circulation sur le sein de l'Égalité, les genoux et les têtes des trois personnages.	Très faibles traces d'usure sur les feuilles de chêne et du coq.
SPL	Légères traces de manipulation mais la plupart du velours de frappe est encore présente.	Très faibles traces de manipulation, mais la plupart du velours de frappe est encore présente.
FDC	Monnaie pratiquement parfaite avec tout son velours de frappe.	Monnaie pratiquement parfaite avec tout son velours de frappe.

sionnelle est toujours précédée par une étape d'authentification pour éviter les contrefaçons, et suivie par une mise sous coque de protection (*cf.* Figure 1.13). Les propriétaires soumettent leurs monnaies et paient le service. La société donne une certification d'authenticité et du grade des monnaies soumises. Voici un processus général pratiqué dans les sociétés qui proposent les services de gradations professionnelles :

- vérification de l'authenticité de la monnaie soumise,
- gradation par le premier expert,
- gradation par le deuxième expert,
- gradation par le troisième expert,
- mise en coque de la monnaie gradée,
- dernière vérification de la monnaie sous coque par un expert.

Ainsi, une confiance relative est donnée par la fusion des avis des experts. Cependant, le fait que le grade d'une monnaie soit un élément subjectif n'a pas été changé même par les sociétés leaders comme PCGS qui déclarent l'impartialité et l'exactitude de leur gradation. Il arrive souvent le cas où une monnaie peut recevoir un grade différent donné par deux sociétés. Pour avoir un grade certifié maximum, les collectionneurs font parfois grader leurs monnaies plusieurs fois ou par plusieurs sociétés. Telles sociétés de gradation proposent également le service de « regradation » destiné aux clients qui cherchent des grades supérieurs à leurs monnaies déjà gradées.

FIGURE 1.12 – Sociétés de gradation professionnelle : PCGS, NGC, ICG et ANACS

FIGURE 1.13 – Monnaies gradées scellées dans des coques de protection

1.3 Photographie numismatique

Durant la seconde moitié du XIXe siècle, l'adoption de la photographie dans les revues numismatiques et les catalogues de ventes permettent aux collectionneurs d'observer et d'évaluer des monnaies de manière plus rapide et plus sûre (Hollard 1991). Comme nous l'avons déjà dit, la valeur d'une monnaie est largement définie par son apparence. Ainsi, la qualité des photos compte énormément dans la valeur marchande des monnaies. Aujourd'hui, les images de monnaie aident considérablement les collectionneurs à présenter, discuter et revendre leurs monnaies sur Internet. Idéalement, une photo de monnaie de qualité devrait être présentée en haute résolution pour en faire ressortir tous les détails. Cependant, la notion de qualité est elle-même liée à son objectif. Pour la documentation, l'image doit faire nettement apparaître les reliefs, en minimisant les bruits comme les reflets et les ombres. Pour attirer les collectionneurs, la photographique peut au contraire bénéficier d'ombres et de reflets afin qu'elle soit plus artistique. Pourtant, la plupart des photos de monnaies en ligne, prises par des amateurs, sont loin d'être de qualité.

1.3.1 Dispositif d'imagerie numismatique

Il existe deux méthodes pour réaliser des photos numismatiques de haute résolution : à l'aide d'un scanner ou à l'aide d'un appareil photo.

- Scanner :

Grâce à la faible épaisseur d'une monnaie, un scanner à plat peut être utilisé pour réaliser le rendu d'une monnaie en haute résolution de manière simple. Par rapport à utiliser un l'appareil photo, il n'existe presque aucun problème technique et le prix du matériel est faible. Pour avoir une bonne qualité d'image, il suffit d'enlever les poussières sur le scanner et sur la monnaie. Ainsi, c'est la solution préférée par les revendeurs particuliers et les amateurs en numismatique. Par contre, les inconvénients majeurs sont incontournables, le scanner ne convient pas aux monnaies à fort relief car il a une très faible profondeur de champ. Par ailleurs, il y a peu de paramètres à régler pour augmenter la qualité de l'image. Un exemple d'une monnaie scannée est montré à la Figure 1.14 (a).

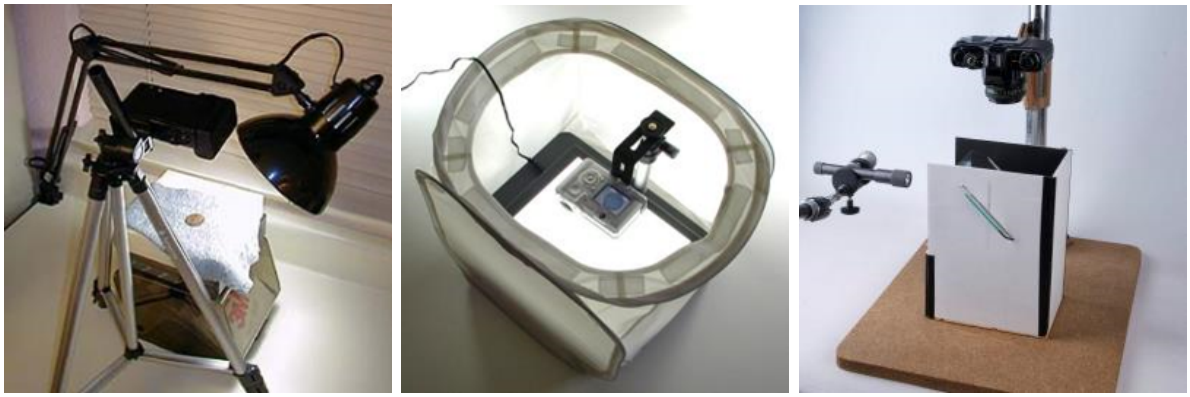
- Appareil photo :

Il s'agit d'un appareil photo numérique. Cependant les principes restent les mêmes pour les appareils photos numériques et argentiques, l'appareil numérique répond à l'utilisation multimédia avec archivage numérique. En outre, l'écran de visualisation dans d'un appareil numérique permet de se rendre compte du résultat sans devoir attendre. Même si un appareil haut de gamme est nécessaire, il faut savoir que le prix de l'appareil photo ne garantit pas une qualité suffisante de photos numismatiques. En effet, une monnaie est tout d'abord un tout petit objet qui est souvent en métal réfléchissant. Ainsi, la problématique de prendre une photo de qualité d'une monnaie reste toujours un sujet difficile à l'heure actuelle, ce qui incite les numismates et les photographes à chercher ensemble la meilleure solution. La Figure 1.14 (b) et (c) montre la différence entre une photo amateur et une photo de qualité. Les photos doivent être prises dans des conditions bien définies. Même si ces conditions peuvent varier selon l'appareil utilisé, l'arrière-plan est souvent choisi quasi-uniforme et l'éclairage est soigneusement contrôlé. Ainsi, il est nécessaire de construire des dispositifs sophistiqués composés d'au moins un appareil photo, un support de monnaie, un pied, et une ou plusieurs sources de lumières. Parmi les différents dispositifs optiques destinés à la photographie numismatique, les solutions retenues sont principalement la tente de lumière, le support illuminé (*cf.* Figure 1.15 (b)) et la lumière axiale (*cf.* Figure 1.15 (c)).

Pour réaliser une photo de monnaie avec un contraste élevé et minimiser l'ombre, une solution est d'utiliser une source de lumière axiale. À l'aide d'une installation optique spécifique, la lumière se déplace du même axe que la direction de la lentille. Le reste de la lumière non axiale est absorbé afin de minimiser la réflexion diffuse. Grâce à cette installation, la partie de la monnaie perpendiculaire au trajet de la lumière reflétée, par

(a) (b) (c)

FIGURE 1.14 – Illustrations des monnaies *Liberty Bell Half Dollar* : (a) issue d'un scanner ; (b) photo amateur ; (c) photo de qualité réalisée par la société NGC



(a) (b) (c)

FIGURE 1.15 – Installations destinées à la photographie numismatique : (a) installation amateur ; (b) tente de lumière et support illuminé ; (c) lumière axiale

exemple le champ, apparaît brillamment, et la partie non-perpendiculaire de la monnaie, par exemple les contours des reliefs, reste toute sombre (*cf.* Figure 1.16 (a)). Les monnaies « belles épreuves » sont difficiles à photographier à cause de leur surface miroir. La lumière axiale permet également de photographier les monnaies « belles épreuves » sans réflexion sur leur surface miroir (*cf.* Figure 1.16 (b)). Théoriquement, il existe une configuration permettant d'obtenir la qualité indispensable à l'identification et à la gradation (seule qualité qui nous intéresse dans cette thèse) : avec le champ et les motifs clairement séparés par un fort contraste. Mais dans la pratique, la moindre différence dans les paramètres optiques, par exemple la distance du diffuseur, donnera un rendu photo bien différent.

1.3.2 Photographies professionnelles

La notion de photographie professionnelle est utilisée pour différencier les photos réalisées par des amateurs, souvent de qualité médiocre. En particulier, il s'agit plutôt de photos numismatiques prises sous un protocole standardisé avec lequel nous pouvons re-

(a)

(b)

FIGURE 1.16 – Monnaies photographiées à l'aide d'une lumière axiale (encadrées en rouge) et non-axiale : (a) monnaie courante ; (b) belle épreuve

produire des photos de même qualité. La standardisation est un critère important pour obtenir des photos professionnelles. Par ailleurs, certaines grandes sociétés numismatiques comme PCGS et NCG fournissent ce type de service aux clients (*cf.* Figure 1.17).

Récemment les chercheurs en traitement d'images se sont mis à explorer les possibilités de la vision par ordinateur dans des applications numismatiques. Les photos professionnelles, prises dans des conditions bien définies et standardisées laissent envisager la possibilité de développer des algorithmes efficaces de reconnaissance des monnaies voire d'aide à la gradation.

1.4 Société GENI

La société **Gradations et Expertises Numismatiques Internationales (GENI)** (*cf.* Figure 1.18), fondée en 2013, est une jeune société de gradation qui a son siège social à

(a) (b)

FIGURE 1.17 – Photos numismatiques professionnelles : (a) PCGS TrueView ; (b) NGC Photo Vision (monnaie sous coque)

Paris, France. Alors que les géants numismatiques américains sont en train d'étendre leurs territoire, la société GENI porte l'ambition de contester leur monopole, et elle s'engage à fournir des services de qualité et des produits plus adaptés au marché de niche en Europe, notamment en France. Les activités principales de la société GENI comportent :

- gradations, expertises et authentifications de monnaies de collection,
- photographies numismatiques,
- retouches d'images et création de catalogues de vente pour des tiers,
- commercialisation de coques de protection,
- *etc.*

Les clients peuvent accéder aux services proposés et consulter leurs monnaies expertisées par la société GENI sur le site www.geni.expert (*cf.* Figure 1.19).

FIGURE 1.18 – Logo de la société GENI

Coque de protection innovante

La coque de protection qui scelle la monnaie gradée est maintenant un standard pour les sociétés faisant de la gradation. Outre la protection contre l'environnement et le temps, et l'aisance du transport, la coque permet également de présenter le logo de la société de gradation et les informations sur la monnaie gradée. Par rapport aux coques américaines basées sur un vieux modèle datant des années 1980 (*cf.* Figure 1.20), la société GENI a conçu un nouveau modèle plus compact et stylé qui facilite énormément le stockage. Pour les monnaies scellées dans les coques américaines, nous ne pouvons pas observer la

FIGURE 1.19 – Site web de la société GENI (www.geni.expert)

tranche. Cependant, pour certaines monnaies européennes, la tranche qui contient souvent des motifs ou des légendes, ne doit pas être occultée par la coque. La coque GENI permet d'observer la tranche des monnaies ainsi que le listel (*cf.* Figure 1.21).

Gradation à domicile

La gradation à domicile est le service ayant le plus de valeur que la société GENI apportera aux clients. En effet, le transport et l'assurance des monnaies sont relativement coûteux. Les monnaies soumises au site de gradation risquent d'être perdues, volées, changées ou endommagées lors de leur transport. Cela est d'autant plus risqué qu'elles ont une grande valeur ou sont en grand nombre. Pour les clients qui ne veulent pas déplacer leurs monnaies précieuses hors de leur maison, la société GENI propose tous les services concernés chez les clients à l'aide de dispositifs spécialement conçus pour être transportables.

Numismatique assistée par ordinateur

Dans le cadre du projet de la gradation à domicile, la société GENI a fait appel aux compétences du laboratoire LIRIS pour tenter d'automatiser certains services à partir de photos professionnelles de monnaies. Il s'agit de pouvoir extraire de manière automatique la monnaie de la photo, de caractériser sa forme, d'identifier son type monétaire mais

FIGURE 1.20 – Coques de protection du modèle « américain »

surtout, de fournir une aide à son identification et à sa gradation. Plus précisément, l'identification de la monnaie se fait en comparant des caractéristiques visuelles, celles liées à la taille et à la forme avec celles de monnaies « connues », et la détection du millésime. La gradation, quant à elle, est principalement donnée par la quantification des motifs en relief et des éventuelles rayures et taches présentes dans la photo.

1.5 Problématiques de vision par ordinateur appliquées à la numismatique

De nos jours, même si les développements informatiques ont abouti à la mise en place de systèmes automatiques dans de nombreux domaine, la numismatique profite peu de ces développements afin d'automatiser ces services clefs. Au début des années 1990, la société PCGS a pourtant lancé des tentatives de gradations assistées par ordinateur et a signalé un taux de précision autour de 90% (Haperin 1990 ; Crain 1993). Malgré tout, 27 ans plus tard, la gradation professionnelle dans la pratique reste encore un travail laborieux fait par l'œil d'experts. À noter toutefois, la nouvelle application *Photograde*^{TM} développée par la société PCGS qui encourage une gradation grossière par des non experts. Dans cette application, chaque type monétaire de dollar américain de chaque grade est représenté par une photo sélectionnée. Il est demandé à l'utilisateur de trouver une photo parmi ces dernières qui est visuellement la plus similaire à la monnaie qu'il souhaite estimer. Certains numismates restent sceptiques quant à la possibilité de remplacer un travail très minutieux s'appuyant sur des arguments souvent subjectifs par celui d'une machine. Cependant, il s'agit-là du défi de cette thèse.

FIGURE 1.21 – Conception de la coque de protection par la société GENI

Ci-dessus sont répertoriés des verrous scientifiques et technologiques pour la vision par ordinateur concernant la numismatique :

Le premier verrou porte sur la segmentation ou le détourage des monnaies avec une extrême précision. En effet, contrairement à certains travaux de la littérature qui segmentent grossièrement des monnaies rondes, il est très important de pouvoir présenter des images de monnaies parfaitement détourées dans les catalogues de vente. Même si la prise de photo est relativement maîtrisée pour les photos professionnelles, il reste des éléments comme l'ombre portée, la texture du fond et l'imperfection de la forme réelle de la monnaie qui gênent une segmentation parfaite. Par ailleurs, le fait que des monnaies rares de formes particulières (monnaies à pans, à dents, striées, cannelées, *etc.*) pourront apparaître dans la pratique multiplie la complexité.

Le second verrou porte sur la reconnaissance des « éléments attendus » présents sur les deux côtés de la monnaie à savoir, sa valeur faciale, son unité, son millésime, *etc.* Il s'agit d'un véritable défi puisque contrairement aux travaux de reconnaissance de texte que nous pouvons trouver dans la littérature (sur des documents scannés, sur des plaques d'immatriculation, dans des vidéos, *etc.*), les caractères sont de la même couleur que le fond de la monnaie. Seule une petite surépaisseur, qu'il faudra arriver à appréhender par l'intermédiaire de l'image, fait la différence. Par ailleurs, en fonction de l'usure de la monnaie, cette différence est souvent très faible pour les monnaies usées. Même si quelques travaux existent dans la littérature à ce sujet, la plupart se focalisent sur certains types monétaires réduisant fortement le champ des possibles.

Le troisième verrou porte sur la reconnaissance des monnaies ou l'identification du type monétaire par analyse d'images. Le type monétaire de la monnaie traitée est reconnu principalement par ses « éléments attendus » qui sont difficile à repérer par ordinateur selon les contraintes mentionnées. La difficulté se situe également au niveau de la variabilité intraclasse et la similarité interclasse. D'une part, l'apparence des monnaies appartenant au même type monétaire varie en fonction de l'usure. D'autre part, certaines monnaies de différents types monétaires se ressemblent globalement et leurs nuances sont subtiles et locales. Par ailleurs, le nombre de classes est important et des nouvelles classes seront incluses au fur et à mesure. Une grande base d'images d'apprentissage est en général indispensable pour tout sujet de reconnaissance d'objets, mais dans le cadre de cette application le problème crucial est le peu de données d'apprentissage, c'est-à-dire le peu de photos professionnelles pour tous les types monétaires.

Enfin, le dernier verrou, concernant la gradation, porte sur la détection des « éléments inattendus » de la monnaie. Plus précisément, il s'agit de repérer d'éventuels petits coups ou d'éventuelles rayures sur la monnaie, de même que d'éventuelles traces d'usures. Il s'agit-là également d'un point assez délicat et peu abordé dans la littérature. En effet, les algorithmes de reconnaissance d'objets dans la littérature, visent plutôt à la reconnaissance d'éléments connus en essayant justement d'être robuste aux bruits, aux imperfections, *etc.* L'objectif est ici exactement le contraire : repérer et quantifier les « éléments inattendus » sur la monnaie à traiter qui n'apparaissent pas sur la monnaie de référence « parfaite ». Cependant, de telles monnaies de référence sans aucun défaut existent rarement dans la réalité, et sont donc inexistantes dans la base de données.

Chapitre 2
État de l'art

ÉTANT parmi les objets les plus utilisés dans la vie quotidienne, les monnaies apparaissent souvent dans des images de test pour différents algorithmes de traitement d'images. Par exemple, des images contenant plusieurs monnaies qui se touchent ou se superposent sont souvent utilisées dans les tutoriels de mathématique morphologique. Cependant, les monnaies elles-mêmes ne sont pas des sujets populaires comme les visages ou les plaques d'immatriculation qui font l'objet sans cesse de nouvelles études en analyse d'images.

Dans la vie quotidienne, les systèmes automatiques de reconnaissance de monnaies, existants comme dans les distributeurs de boissons et les machines à sous, semblent suffire par mesurer les propriétés physiques (diamètre, poids, largeur et propriétés électromagnétiques). En ce qui concerne les monnaies de collection, la reconnaissance se fait, comme indiqué dans le Chapitre 1, principalement via des interventions manuelles des numismates professionnels.

Dans le cadre de cette thèse, nous nous concentrons uniquement sur les travaux parallèles destinés aux monnaies dont les données utilisées sont des images qui présentent les côtés d'une monnaie. D'autres problématiques qui utilisent des images de test contenant une ou plusieurs monnaies ne sont pas abordées. Parmi les travaux en traitement et analyse d'images qui prennent les monnaies comme des objets à étudier, l'identification automatique de types monétaires est le sujet le plus populaire. Il attire l'intérêt des chercheurs depuis une dizaine d'années. Les contributions principales à la littérature sont dans le cadre de la compétition MUSCLE CIS (Nölle, Rubik et Hanbury 2006), le projet COINS (Zaharieva, Kampel et Zambanini 2007) et le projet ILAC (Kavelar et al. 2013) :

- Compétition *MUSCLE Coin Images Seibersdorf* (CIS) : c'est une compétition d'algorithmes pour l'identification de types monétaires. Les organisateurs fournissent une grande base de données composée de 60000 images de monnaies « modernes » européennes décomposées en 692 types monétaires. Cette base de données (*MUSCLE CIS*) est considérée comme la première et la plus grande base de données d'images de monnaie standardisées.

- Projet COINS (*COmbatting Illicit Numismatic Sale*) : il s'agit d'un projet qui vise à contribuer à la lutte contre les commercialisations illégales de monnaies anciennes. L'objectif de ce projet est de développer des inventaires standardisés de monnaies anciennes accompagnés d'outils de gestion de données et d'un moteur de recherche en ligne. C'est dans le cadre de ce projet que des algorithmes de reconnaissance d'objet ont été développés pour la première fois pour des monnaies anciennes.

- Projet ILAC (*Image-based cLassification of Ancient Coins*) : il s'agit d'un projet interdisciplinaire qui a pour objectif de construire une application numismatique de classification des monnaies anciennes par traitement d'images sur une grande base de données. De nouveaux problèmes conceptuels concernant les bases de données et les méthodes de reconnaissance d'objets dédiées aux monnaies anciennes sont étudiées dans

le cadre du projet.

Selon la nature des monnaies étudiées, il est possible de diviser les études en deux groupes : celles traitant des monnaies « modernes », autrement dit celles frappées à la machine, et celles traitant des monnaies anciennes frappées manuellement à l'aide d'un marteau ou d'un balancier. Plus précisément, les monnaies « modernes » possèdent toujours une forme géométrique régulière, très souvent une forme de disque, et des motifs fixes pour un type monétaire, alors que les monnaies anciennes sont irrégulières (*cf.* Figure 2.1). Par ailleurs, les monnaies « modernes » appartenant au même type monétaire portent des motifs identiques, alors que les monnaies anciennes sont plus ou moins « personnalisées » (*cf.* Figure 2.2). Les caractéristiques différentes entre les monnaies « modernes » et les monnaies anciennes conduisent à des stratégies de reconnaissance automatique bien différentes. Dans le cadre de cette thèse, nous nous focalisons sur les monnaies « modernes » et adaptons un processus classique de reconnaissance tel que présenté à la Figure 2.3. Il faut noter que la gradation est effectuée après l'identification du type monétaire.

Ce chapitre examine les différents travaux qui sont en lien avec le sujet de cette thèse. Dans la section 2.1, nous étudions les techniques de la segmentation d'images appliquées à l'extraction des monnaies. La section 2.2 aborde les approches visant l'identification du type monétaire. La section 2.3 présente les premières tentatives de gradation de monnaies que nous trouvons dans la littérature. Enfin la section 2.4 présente le contexte spécifique de cette thèse.

Figure 2.1 – Différence de formes entre des monnaies anciennes (en haut) et des monnaies « modernes » (en bas)

FIGURE 2.2 – Motifs non-standardisés dans les monnaies anciennes appartenant au même type monétaire (Zambanini et Kampel 2012)

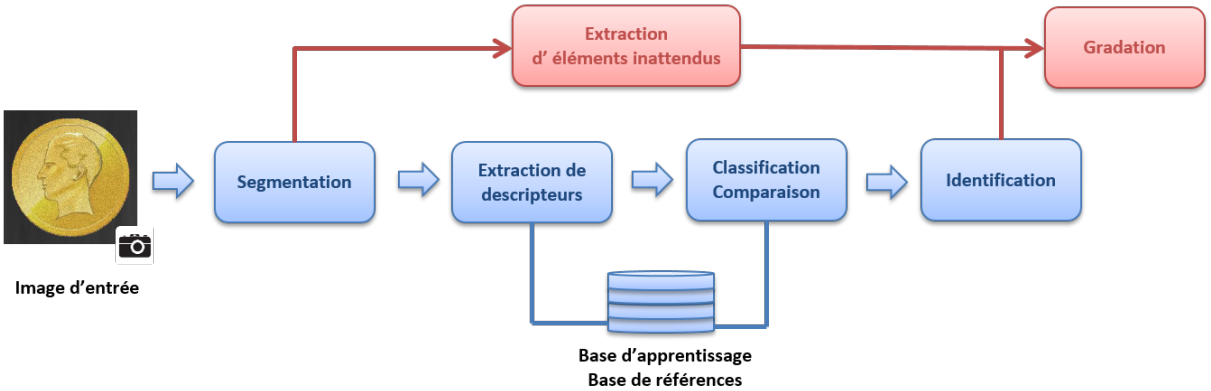


FIGURE 2.3 – Schéma général de la reconnaissance d'objets (en bleu), les boîtes rouges représentent le schéma de gradation des monnaies utilisé pour cette thèse.

2.1 Segmentation d'images de monnaies

La segmentation d'images a pour objectif de regrouper les pixels appartenant à différentes régions disjointes selon certains critères comme l'intensité, la couleur, la texture, des propriétés sémantiques, *etc.* Si le nombre de régions à séparer est égal à deux, ce qui est le cas pour l'extraction d'objets, la segmentation peut être considérée comme une binarisation des pixels appartenant à l'objet d'intérêt et des pixels appartenant à l'arrière-plan. En analyse d'images, la segmentation sert souvent d'étape de prétraitement afin d'extraire l'objet d'intérêt. Par exemple, dans le cadre de cette thèse, l'extraction de la monnaie à partir d'une photo de manière précise est la première phase de toutes les analyses suivantes. Elle est indispensable pour les numismates qui souhaitent utiliser ces monnaies bien segmentées pour présenter dans les catalogues de vente.

Avant de présenter les approches de segmentation destinées à extraire les monnaies, nous préférons d'abord présenter une vue d'ensemble sur les techniques de segmentation de la littérature. Des exemples de chaque catégorie de techniques sont illustrés à la Figure 2.4.

2.1.1 Vue d'ensemble de la segmentation d'images

À ce jour, les techniques de segmentation peuvent être catégorisées en cinq groupes :
- segmentation basée seuillage,
- segmentation basée régions,
- segmentation basée contours,
- segmentation basée graphe,
- segmentation à l'aide d'approches plus évoluées.

Segmentation basée seuillage

Ce genre de méthodes se base sur l'hypothèse qu'il existe une différence significative en termes d'intensité entre les régions que nous souhaitons séparer. Dans le cas où nous utilisons un seuil qui permet de séparer les deux zones de l'image (par exemple l'objet d'intérêt et le fond), ce seuil est alors soit défini manuellement comme un paramètre d'entrée, soit calculé automatiquement. Les techniques de seuillage peuvent être appliquées de manière globale, et dans ce cas nous utilisons un seul seuil pour l'ensemble de l'image, ou de manière locale, et alors, différents seuils sont appliqués suivant les régions locales de l'image. Nous pouvons citer, par exemple le seuillage adaptatif proposé par Yanowitz et Bruckstein (1989).

La complexité du problème se situe dans l'obtention du seuil optimal. Différents travaux de la littérature se sont penchés sur ce problème en utilisant, par exemple, l'information d'histogramme (Otsu 1975 ; Glasbey 1993 ; Dong et al. 2008), l'entropie de la répartition des intensités (Shanbhag 1994), ou l'information sur la forme (Shi, Ray et Zhang 2012). Une revue de ces différentes techniques de seuillage est résumée dans Sezgin et al. (2004).

Segmentation basée régions

Les méthodes basées régions manipulent directement des régions en fonction des critères d'homogénéité. Parmi les algorithmes les plus répandus, nous pouvons citer les techniques de croissance de régions qui partent de points sources, ou graines (*seeds*), qui vont grandir ensuite par incorporation des pixels voisins les plus similaires suivant des critères d'homogénéités donnés (Adams et Bischof 1994). De telles graines peuvent être calculées automatiquement (par exemple, les minima de l'image), ou fournies par un utilisateur de manière interactive (Mivcuvsik et Hanbury 2006). Les méthodes de lignes de partage des eaux (*watershed*), basées sur la morphologie mathématique, forment aussi une croissance de régions qui cherchent les bassins versants du relief formé par l'image (Vincent et Soille 1991 ; Beucher et Meyer 1992). Également basée sur les critères d'homogénéité, les approches *split & merge* subdivisent progressivement les régions non-homogènes et fusionnent les régions voisines similaires (Pavlidis et Horowitz 1974). En outre, les méthodes

de *clustering* telles que *k-means*, *meanshift* (Cheng 1995 ; Comaniciu et Meer 2002), *etc.*, visent à séparer les classes de pixels en minimisant la différence intraclasse.

Segmentation basée contours

La segmentation basée contours cherche une transition détectable entre les zones, par exemple, des changements abrupts d'intensité entre régions connexes. Les contours dans l'image sont trouvés par détecteurs de contours tels que *Canny*, *Sobel*, *Laplacien of Gaussian* (LoG), *Difference of Gaussian* (DoG), *etc.* Heath et al. (1996) proposent une revue de ces détecteurs. La construction des partitions par contours fermés est l'un des problèmes complexes dans ce genre de méthodes. Pour résoudre ce problème, il est possible d'utiliser des modèles déformables de courbes paramétriques (courbe de Bézier, spline, *etc.*) Les contours actifs sont des approches populaires basées contours, dans lesquelles le contour initial évolue itérativement de façon à minimiser une énergie basée sur certains critères d'homogénéité (Kass, Witkin et Terzopoulos 1988). Dans les algorithmes plus récents de contours actifs, le contour est considéré comme une courbe de niveau (*level-set*) (Chan et Vese 2001). Très souvent, les approches sophistiquées basées contours requièrent l'initialisation des contours de départ. Une telle approche a été développée dans le travail de Zhang, Stoecker et Moss (2000). L'intérêt principal des méthodes de segmentation basées contours est de minimiser le nombre d'opérations. Malgré des itérations nécessaires, si les contours peuvent être initialisés en étant proches de l'objet à segmenter, l'efficacité et la précision de la segmentation seront assurées.

Segmentation basée graphe

Les méthodes basées graphe modélisent l'image sous forme de graphe dont les sommets sont des pixels dans l'image et les arêtes sont pondérées souvent par la distance de couleur. La segmentation est ainsi réalisée par la recherche d'une coupe optimale de graphe, qui s'appelle coupure de graphe (*graph-cuts*) (Shi et Malik 2000). Les techniques de segmentation basées graphe s'appuient sur les notions de plus court chemin (*shortest path*) (Boykov et Jolly 2001 ; Falcão, Stolfi et Alencar Lotufo 2004), arbre recouvrant de poids minimal (*minimum spanning tree*) (Felzenszwalb et Huttenlocher 2004), et de marche aléatoire (*random walks*) (Grady 2006), *etc.*

Segmentation basée sur des approches plus évoluées

Dans des contextes où une seule approche n'arrive pas à aboutir une segmentation idéale, la coopération de plusieurs des techniques mentionnées ci-dessus est une stratégie qui peut être utilisée dans des applications spécifiques. Par exemple, pour segmenter les feuilles d'arbres sur un fond naturel, Cerutti et al. (2011) ont proposé une segmentation sophistiquée combinant un modèle paramétrique, des contours actifs et une croissance de

régions basée couleur. Certaines approches semi-automatiques telles que *Intelligent Sissor* (Mortensen et Barrett 1995) et *GrabCut* (Rother, Kolmogorov et Blake 2004) requièrent une interaction de l'utilisateur pour arriver à une segmentation relativement précise. Les méthodes de co-segmentation traitent simultanément plusieurs images similaires en profitant des informations discriminantes communes à toutes les images traitées (Joulin, Bach et Ponce 2010). La segmentation basée apprentissage peut quant à elle modéliser des informations de haut niveau pour diviser l'image en régions sémantiques, par exemple, à l'aide de réseaux de neurones convolutifs (Long, Shelhamer et Darrell 2015).

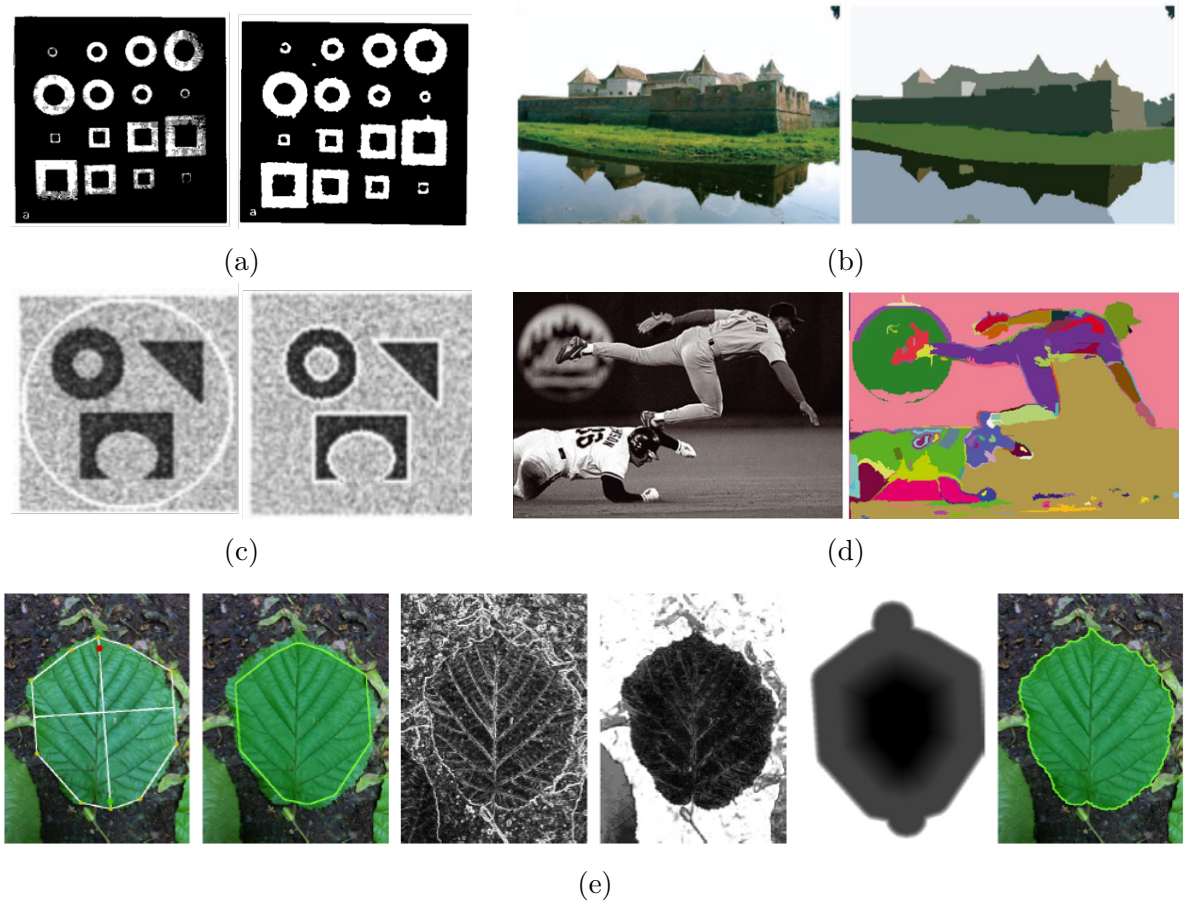


Figure 2.4 – Exemples de techniques de segmentation d'image : (a) par seuillage : seuil adaptatif (Yanowitz et Bruckstein 1989) ; (b) par régions : *meanshift* (Comaniciu et Meer 2002) ; (c) par contours : contours actifs *level-set* ; (d) par graphe : de plus court chemin (Falcão, Stolfi et Alencar Lotufo 2004) (e) par approche évoluée : modèle paramétrique + carte de dissimilarité de couleur + contours actifs guidés (Cerutti et al. 2011)

.

Le résumé général ci-dessus montre la variété des techniques de segmentation d'images. Les approches simples, par exemple, celles utilisant une valeur de seuil globale, sont faciles à implémenter et à calculer, mais elles ne peuvent pas être appliquées à des situations compliquées. En revanche, les méthodes avancées, comme la segmentation sémantique par réseau convolutif, sont puissantes mais nécessitent un coût de calcul important et une grosse base d'apprentissage. Par conséquent, il n'existe pas de méthode de segmentation

qui produise de meilleurs résultats, quelle que soit l'application. Pour la segmentation des monnaies, la méthode choisie devrait permettre une segmentation très précise de l'objet d'intérêt, être obtenue en temps réel et être robuste aux monnaies de formes différentes.

2.1.2 Segmentation des monnaies

La segmentation des monnaies est souvent considérée comme un problème facile pour les chercheurs, puisque les monnaies à traiter sont supposées être sur un fond quasi-uniforme avec une couleur différente de celle de la monnaie. Cette condition dans l'image est assurée dans les photos numismatiques professionnelles, voire dans certaines photos amateurs. L'extraction de la monnaie dans un environnement complexe présente peu d'intérêt dans les véritables applications numismatiques.

En conséquence, il est facile de réaliser une segmentation « grossière » à l'aide de la plupart des méthodes de segmentation de la littérature. La simplicité d'implémentation et la vitesse de calcul sont les seuls éléments à prendre en compte.

En donnant des photos de monnaie comme données d'entrée, la séparation de la monnaie de son arrière-plan est la première phase pour toutes les analyses numismatiques telles que l'identification du type monétaire.

Seuillage global

Dans un contexte où la monnaie se sépare de façon évidente de son arrière-plan en utilisant l'intensité ou la couleur, il est possible de réaliser une segmentation très simple à l'aide d'un seuil global.

Si l'arrière-plan est constant, la valeur de seuil est souvent considérée comme un paramètre d'entrée expérimental, fixé avant le déroulement du processus (Nölle et al. 2003 ; Van Der Maaten et Poon 2006). Par exemple, pour les monnaies photographiées sur un convoyeur noir, les intensités des pixels appartenant à la monnaie sont supposées être supérieures à une valeur fixe. Dans un cas plus général, le seuillage global peut être également déterminé automatiquement par la méthode d'Otsu (Otsu 1975). C'est une méthode de segmentation globale affectant les pixels aux deux régions en une seule passe, basée sur la définition d'un seuil optimal sur l'histogramme d'intensités.

Cependant, il reste souvent des pixels mal affectés, par exemple, dans la Figure 2.5 des pixels sombres autour des reliefs sont affectés au fond. Pour supprimer ces mauvaises affectations, des techniques de détection de contours et de morphologie mathématique sont souvent utilisées dans le processus de segmentation (Van Der Maaten et Poon 2006).

L'intérêt de la segmentation de monnaie par seuillage global est la simplicité et la vitesse de calcul. L'inconvénient majeur de ce genre de méthodes est qu'elles n'arrivent pas à segmenter des monnaies dont l'intensité ou la couleur est proche de l'arrière-plan, par exemple des monnaies « sombres » sur un convoyeur noir (Reisert, Ronneberger et

Burkhardt 2007). Par ailleurs, pour certaines monnaies bimétalliques ou celles riches de patines multicolores, la segmentation par seuillage automatique peut obtenir seulement une partie de la monnaie car le reste est proche de l'arrière-plan en intensité (*cf.* Figure 2.5 (b)).

La suppression de l'ombre est un sujet peu abordé dans la segmentation des monnaies dans ce contexte, parce que les ombres sont soit minimisées par des conditions de prise de photos contrôlées, soit atténuées par le fait que le fond est très souvent relativement sombre.

(a) (b)

Figure 2.5 – Exemples de monnaies bimétalliques segmentées à l'aide de la méthode d'Otsu : (a) monnaie globale avec une intensité différente de l'arrière-plan ; (b) monnaie avec une partie dont l'intensité est proche de l'arrière-plan

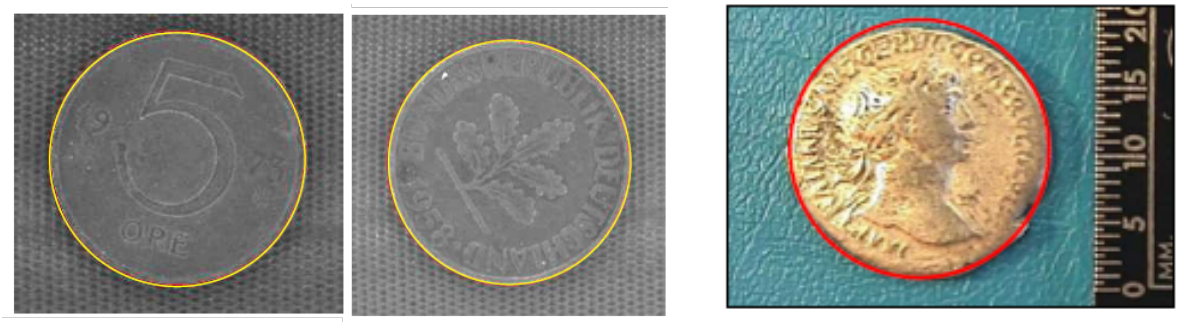
Reconnaissance de forme circulaire

Dans un contexte où les monnaies à segmenter sont de forme circulaire, les techniques de reconnaissance de formes comme la transformée de Hough généralisée (Ballard 1981) sont utilisées (Reisert, Ronneberger et Burkhardt 2007 ; Shen et al. 2011). Le principe de la transformée de Hough consiste à rechercher la présence d'une courbe caractérisée par un jeu de paramètres ayant un score élevé. La segmentation résultante est toujours présentée par une forme circulaire parfaite.

La segmentation des monnaies rondes par transformée de Hough est plus robuste que celle par seuillage, notamment dans le cas où le contraste entre la monnaie et l'arrière-plan est faible (*cf.* Figure 2.6 (a)). Cependant, ces techniques risquent de « couper » le bord de la monnaie. L'autre inconvénient est que cette technique ne s'adapte pas à des monnaies avec une autre forme, à pans par exemple.

Méthodes destinées à des monnaies anciennes

Par rapport à la segmentation d'une monnaie « moderne » qui est souvent considérée comme étant ronde, la segmentation d'une monnaie ancienne est plus compliquée. En effet, la forme irrégulière empêche d'utiliser des méthodes basées sur une forme prédéfinie. Par ailleurs, en raison de l'irrégularité de l'épaisseur de la monnaie ancienne, il est plus difficile



(a) (b)

FIGURE 2.6 – Exemples de monnaies segmentées par transformée de Hough : (a) monnaies « modernes » (Reisert, Ronneberger et Burkhardt 2007) ; (b) monnaies anciennes (Arandjelovic 2010)

.

d'avoir des photos de bonne qualité du point de vue de l'éclairage. En effet, l'éclairage non-uniforme fait qu'un seuillage global sur l'image n'est pas suffisant pour avoir des résultats satisfaisants.

Zaharieva et al. (2007) ont adapté la méthode de seuillage adaptatif (Yanowitz et Bruckstein 1989) aux monnaies anciennes. Dans ce contexte, une surface de seuillage est calculée par l'interpolation des points échantillonnés sur l'image de contours, obtenue par l'algorithme Marr-Hildreth (Marr et Hildreth 1980).

En supposant que les monnaies anciennes qui ont une forme irrégulière sont tout de même plus ou moins proches d'une forme circulaire, Zambanini et Kampel (2009) ont proposé une méthode spécifique qui s'appelle segmentation par forme contrôlée (*Shape-controlled segmentation*). Pour ce faire, ils appliquent une série de valeurs de seuil sur une carte des probabilités et, déterminent la segmentation optimale grâce à un facteur de forme (FF) décrivant la circularité.

Cette méthode est composée de deux phases. Plus précisément, dans la première phase, une carte des probabilités est calculée par combinaison d'un filtre d'entropie locale et d'un filtre de gamme d'intensité locale. Dans la deuxième phase, le Facteur de Forme (FF) est calculé pour toutes les valeurs de seuil t et le seuillage optimal t^* est celui qui maximise FF sur la segmentation résultante C_{t^*}, c'est-à-dire celui qui permet d'obtenir un objet segmenté le plus proche d'une forme circulaire (*cf.* Figure 2.7). Il faut noter que cette méthode ne peut pas exclure les ombres si elles englobent une région quasi-ronde (*cf.* Figure 2.8).

$$t^* = argmax\ FF(C_t) \tag{2.1}$$

$$FF(C_t) = \frac{4\pi A_t}{P_t^2} \tag{2.2}$$

où C_t est la composante connexe binaire obtenue par seuillage t , A_t et P_t sont l'aire et le périmètre de C_t.

Quant à Arandjelovic (2010), il s'est aussi basé sur l'hypothèse que la forme à segmenter est une approximation d'une forme ronde, mais il a utilisé d'abord une transformée de Hough circulaire modifiée (Atherton et Kerbyson 1999) à effectuer une segmentation « grossière » afin d'enlever la plupart des pixels de l'arrière-plan « inutiles » (*cf.* Figure 2.6 (b)). Ensuite, le bord précis est détouré dans des coordonnées polaires par les distances radicales dans les directions discrètes, présentées par un modèle markovien caché.

En résumé, les méthodes proposées pour segmenter les monnaies anciennes en forme irrégulière sont basées sur la forme segmentée attendue et composée d'une segmentation « grossière » et d'une segmentation « fine ».

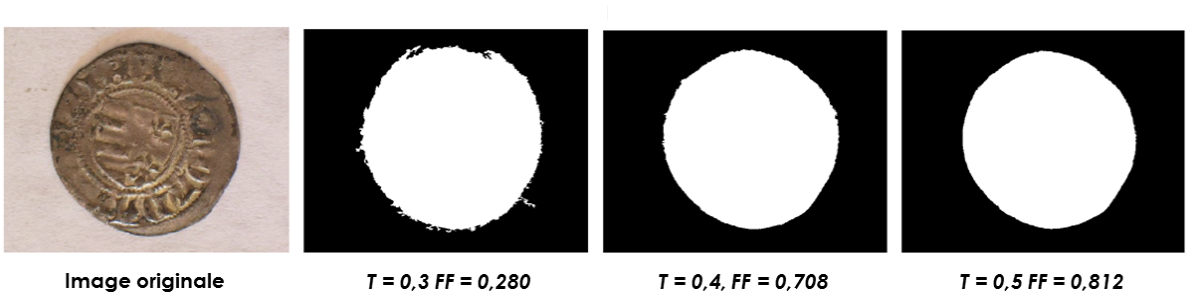


Figure 2.7 – Facteur de Forme (FF) des régions segmentées obtenues à l'aide de différents seuils T (Zambanini et Kampel 2009)

(a) (b)

Figure 2.8 – Exemples de monnaies anciennes segmentées par la méthode de Zambanini et Kampel (2009) : (a) monnaie bien segmentée ; (b) monnaie mal segmentée à cause des ombres

2.2 Analyse d'images pour l'identification du type monétaire

En numismatique nous employons le terme **type monétaire** pour décrire la classe. Il s'agit d'un problème de reconnaissance ou de classification supervisée des types monétaires. Ainsi, nous appellerons reconnaissance des monnaies l'identification du type monétaire dans la suite de cette thèse. Bien qu'il existe des systèmes automatiques de reconnaissance de monnaies dans la vie quotidienne, ces systèmes sont principalement basés sur les propriétés physiques des monnaies comme le diamètre, l'épaisseur ou les propriétés électromagnétiques. De ce fait, ils ne peuvent que traiter un nombre limité de monnaies. Ces systèmes existants peuvent être trompés par des monnaies étrangères ou des jetons ayant des propriétés physiques similaires (*cf.* Figure 2.9). Les motifs des monnaies sont ainsi plus discriminants et informatifs que les propriétés physiques.

(a) (b)

FIGURE 2.9 – Monnaies physiquement similaires : (a) Thaï 10 Baht (environ 20 centimes en euro) ; (b) 2 euros

L'identification automatique du type monétaire par analyse d'images intéresse des chercheurs depuis une dizaine d'années. L'étape essentielle consiste à trouver une description discriminante, globale ou locale, de leurs motifs. Cette description est utilisée ultérieurement pour entraîner un classifieur, ou est comparée avec celle de monnaies de référence via une mesure de similarité. Dans certains systèmes, si la similarité la plus élevée est inférieure à un seuil empiriquement fixé, la monnaie requête est considérée comme un type monétaire inconnu (Nölle et al. 2003 ; Van Der Maaten et Poon 2006 ; Reisert, Ronneberger et Burkhardt 2007). Normalement, l'entraînement des classifieurs nécessite une quantité importante de données d'apprentissage pour chaque type monétaire. Ainsi le temps d'apprentissage peut être non négligeable. De même, la reconnaissance peut prendre du temps. Cependant, des informations *a priori* disponibles comme les propriétés physiques, peuvent être incluses pour réduire le temps de calcul. Au contraire, elles peuvent être utilisées *a posteriori* pour améliorer le résultat. Dans la suite de ce paragraphe, nous nous concentrons sur les travaux portant sur l'extraction de caractéristiques sur des monnaies.

- Approches basées sur des caractéristiques globales : dans ces techniques, la monnaie segmentée est à l'origine d'une description globale, souvent sous forme d'un vecteur de caractéristiques. Ce sont en général des méthodes classiques que choisissent les auteurs qui travaillent notamment sur les monnaies « modernes ». Cependant, ces approches globales ne s'adaptent pas aux monnaies anciennes à cause d'une variabilité intraclasse extrêmement élevée. Des comparaisons de plusieurs approches globales entre les monnaies « modernes » et les monnaies anciennes ont été conduites par Van Der Maaten et Postma (2006) et Zaharieva et al. (2007). Nous détaillons ces méthodes dans le paragraphe 2.2.1.

- Approches basées sur des caractéristiques locales : ces techniques utilisent l'appariement de caractéristiques locales extraites sur différentes régions des monnaies, ou l'apprentissage de caractéristiques locales en employant des sacs de mots visuels. De nombreuses études sur les monnaies anciennes ont utilisé ce genre de méthodes car les caractéristiques locales sont robustes à la déformation ou à la dégradation partielle des motifs. Nous détaillons ces méthodes dans le paragraphe 2.2.2.

- Approches basées sur la lecture des légendes : il s'agit de l'identification du type monétaire en reconnaissant l'information textuelle, constituées par les légendes inscrites. Dans certains cas particuliers, par exemples les monnaies anciennes comportant une variabilité intraclasse élevée dans leurs motifs, la discrimination peut principalement se faire par l'information textuelle des légendes plutôt que par l'information visuelle de ses motifs. Pourtant, la détection et la reconnaissance des caractères de la monnaie est un véritable défi. Le fait que les caractères de la monnaie soient souvent de la même couleur que le fond empêche d'adopter les techniques classiques de la reconnaissance de caractères. Dans la pratique, la lecture des légendes est souvent combinée avec d'autres méthodes basées sur des caractéristiques extraites des autres motifs pour avoir de meilleurs résultats par fusion, comme nous le verrons dans le paragraphe 2.2.3.

Dans les paragraphes suivants, nous allons détailler les études incluses dans les trois catégories mentionnées précédemment.

2.2.1 Approches par descripteurs globaux

Deux monnaies appartenant au même type monétaire devraient avoir des apparences similaires, principalement données par leurs motifs, contrairement à celles appartenant à des types monétaires différents. L'utilisation des descripteurs globaux est basée sur l'idée de trouver une description discriminante sur l'ensemble des motifs caractérisant un type monétaire.

Distribution de contours

Les caractéristiques globales basées contours présentent directement les informations visuelles du type monétaire car les contours obtenus à l'aide de détecteurs de contours correspondent directement à l'ensemble des motifs en relief de la monnaie. Ce sont les premières caractéristiques qui ont intéressé les chercheurs.

Dans le système *COIN-O-MATIC*, Van Der Maaten et Poon (2006) ont fait des expériences sur la manière d'encoder les contours. Avant l'extraction des caractéristiques, l'image de la monnaie est convertie en une carte des contours à l'aide de détecteurs de contours comme *Sobel*. Ensuite, la carte des contours est divisée en sous-régions (*bins*) et le nombre de pixels des contours dans chaque sous-région est compté pour construire un histogramme. Cet histogramme normalisé est ainsi un vecteur de caractéristiques. Les auteurs ont proposé trois variantes de ce descripteur global en fonction de la répartition des sous-régions (*bins*) dans la monnaie :

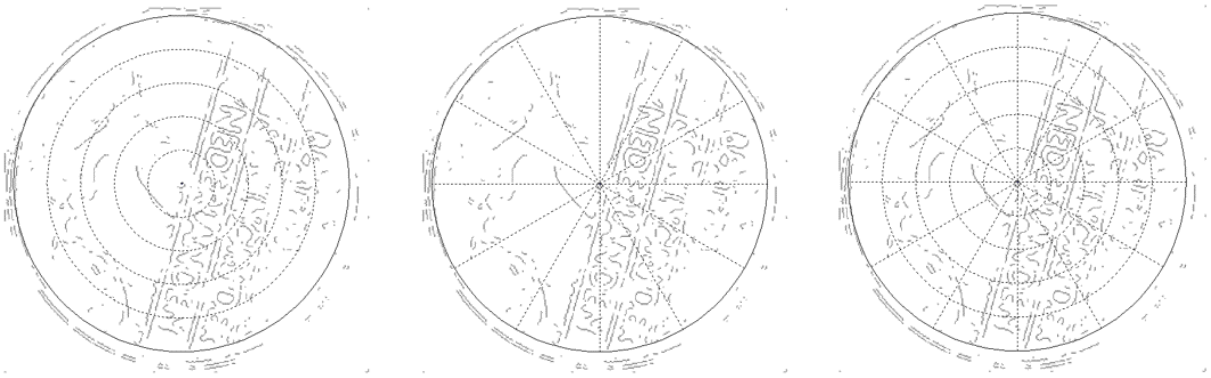
- *Edge-distance-distribution* (EDD) : la carte des contours est divisée par des anneaux concentriques. Le descripteur résultant caractérise ainsi la distribution des contours par rapport au centre de la monnaie (*cf.* Figure 2.10 (a)). Il s'agit d'un descripteur invariant par rotation.

- *Angle-distance-distribution* (ADD) : la carte des contours est divisée en secteurs égaux. Le descripteur caractérise ainsi la distribution des contours par angle (*cf.* Figure 2.10 (b)). Ce descripteur n'est pas invariant par rotation mais nous pouvons utiliser la transformée de Fourier pour atteindre cette invariance.

- *Edge-angle-distance-distribution* (EADD) :la carte des contours est divisée à l'aide d'une grille log-polaire (*cf.* Figure 2.10 (c)). Comme pour ADD, la transformée de Fourier permet de rendre ce descripteur invariant à la rotation. Il s'agit au final de la combinaison de EDD et ADD (Van Der Maaten et Poon 2006).

Les trois descripteurs ci-dessus peuvent être obtenus à l'aide d'un nombre de *bins* variable. Selon les expériences menées par les auteurs, le descripteur EADD multi-échelles utilisant plusieurs nombres de *bins* permet d'obtenir des meilleurs résultats qui sont intéressés au système *COIN-O-MATIC*. Il faut bien noter que ces trois types de répartition de l'image de la monnaie segmentée, résumés dans cette étude, sont largement utilisés pour l'encodage des caractéristiques globales des monnaies dans des travaux parallèles.

Ainsi, par exemple, la classification de deux types monétaires (500 yen japonais et 500 won coréen) est effectuée par Fukumi et al. 1992. Les auteurs utilisant des caractéristiques de contours échantillonnés à l'aide d'une répartition similaire que celle de EADD pour entraîner des réseaux de neurones.



(a) (b) (c)

Figure 2.10 – Répartition de la carte des contours (Van Der Maaten et Poon 2006) : (a) EDD ; (b) ADD et (c) EADD

.

De même, pour trier automatiquement les types monétaires pré-euro de plus de 30 pays européen, le système *Dagobert* a été conçu au sein du *ARC Seibersdorf research GmbH* en adoptant des techniques de traitement et d'analyse d'images. Dans ce système, Nölle et al. (2003) ont utilisé des descripteurs similaires à EDD et ADD, avec une autre caractéristique qui encode la distribution des contours autour de chaque pixel de contours. À l'aide d'une présélection sur les propriétés physiques mesurées par les détecteurs du système, les auteurs ont reporté un taux d'identification global élevé de 96% sur 12949 images réparties avec 913 types monétaires de référence.

Cependant, la détection des contours reste sensible au bruit et l'encodage par un simple comptage des contours par région n'utilise pas pleinement l'information des monnaies « modernes ».

Directions des gradients

Reisert, Ronneberger et Burkhardt (2007) ont construit un descripteur en utilisant l'information sur les directions des gradients. Pour simplifier, les cartes des gradients radiales et angulaires sont présentées dans des coordonnées polaires et normalisées de taille de $(r \times \varphi)$. Les directions des gradients sont ensuite normalisées, discrétisées et encodées par *Fast Fourier Transform* (FFT) sous forme de M représentations binaires (M étant le nombre de discrétisations angulaires). Le descripteur final est ainsi un vecteur contenant $(M \times r \times \varphi)$ éléments. Le processus de cette méthode est illustré à la Figure 2.11. En utilisant ce genre de descripteur, Maaten et Boon (2006) ont effectué des tests sur les monnaies synthétiquement occultées. Contrairement aux descripteurs basés contours, ce descripteur global basé sur les directions des gradients montre que même les régions sans motif comportent des informations structurées.

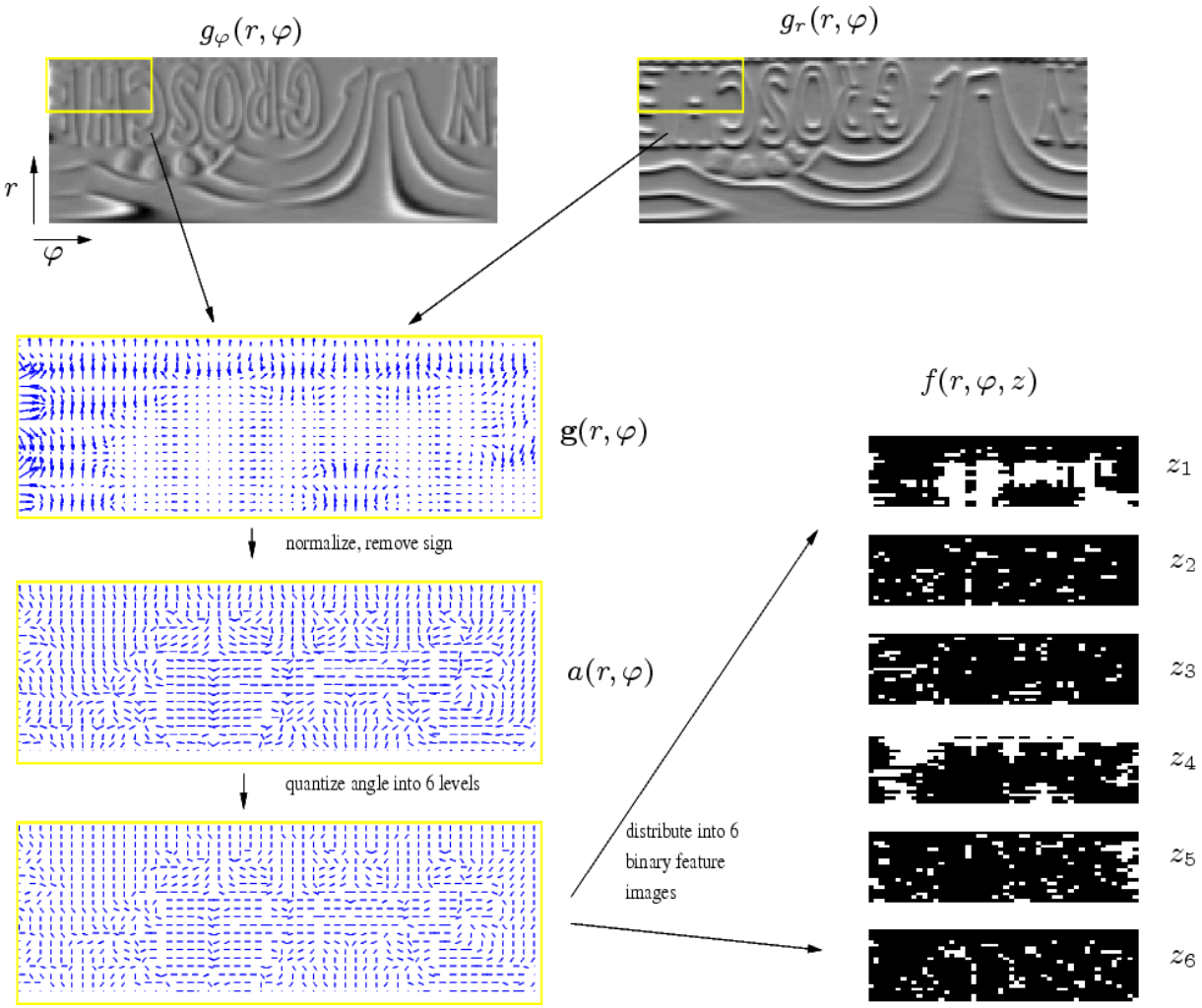


FIGURE 2.11 – Processus de construction du descripteur fondé sur les directions des gradients (cartes des directions des gradients et représentations binaires des descripteurs correspondant uniquement aux boîtes englobantes en haut à gauche) (Reisert, Ronneberger et Burkhardt 2007)

Approche basée eigenspace

Dans l'objectif de surmonter les difficultés causées par des illuminations variables dans les photos de monnaies, Huber et al. (2005) ont choisi les eigenimages comme description. Pour obtenir une invariance à la rotation, les monnaies sont recalées par corrélation croisée avec les monnaies de référence. Pour faire cela, la carte des contours présentée dans des coordonnées polaires est divisée en bandes binaires encodées par *Fast Fourier Transform* (FFT). Ensuite, un eigenspace invariant par illumination est construit à l'aide de la carte des contours lissés en prenant compte le diamètre et l'épaisseur. La Figure 2.12 illustre sur un exemple les huit premières eigenimages reconstruites. La décision finale est obtenue par fusion Bayésienne des probabilités de deux surfaces de la monnaie requête avec un mécanisme de rejet.

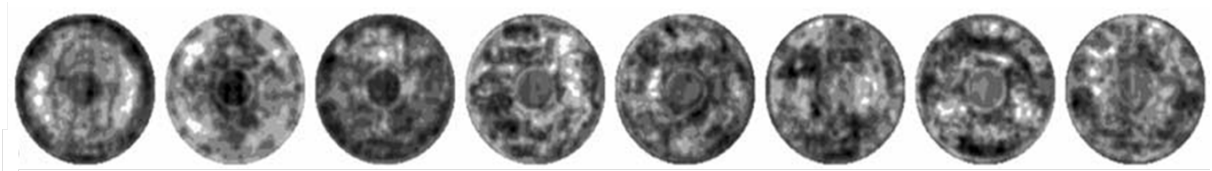


Figure 2.12 – Exemples d'eigenimages construites à partir des contours lissés (Huber et al. 2005)

Descripteurs de texture

L'ensemble des motifs en relief de la monnaie peut être considéré comme une texture, c'est la raison pour laquelle certains chercheurs pensent que les descripteurs de textures représentent mieux les détails de la monnaie que les descripteurs basés uniquement sur les contours. La description de la texture passe le plus souvent par une phase de filtrage de l'image d'intensité, suite à la convolution par un filtre d'ondelettes. Trois descripteurs de texture sont les plus souvent utilisés dans les applications de reconnaissance de monnaies : ondelettes de Gabor (Van Der Maaten et Postma 2006 ; Shen et al. 2009, 2011), ondelettes de Dauchechies (Van Der Maaten et Postma 2006), et LBP (*Local Binary Pattern*) (Shen et al. 2011). Plus précisément, Van Der Maaten et Postma (2006) ont sélectionné SPCA (*Sparce Principal Component Analysis*) pour la réduction de la dimension des vecteurs de caractéristiques ; Shen et al. (2011) ont extrait les caractéristiques sur une répartition concentrique de l'image semblable à celle de ADD. Selon les résultats reportés par les auteurs, les descripteurs de texture ne sont pas performants. En effet, les motifs des monnaies ne sont pas toujours « texturés », autrement dit présentés de manière répétitive.

Descripteurs de forme

Les descripteurs de forme ont été d'abord utilisés pour les monnaies anciennes non-rondes. Zaharieva et al. (2007) ont adopté le *shape context* (Belongie, Malik et Puzicha 2002) alors que Huber-Mörk et al. (2011) ont utilisé une approche qui s'appelle *Deviation from Circular Shape Matching* (DCSM), illustrée à la Figure 2.13. Cette approche s'adapte mieux que *shape context* aux cas des monnaies anciennes. En effet, la forme de chaque monnaie ancienne non ronde présente une courbe particulière. Il faut mentionner que le descripteur de forme tout seul est performant pour trouver une monnaie ancienne individuelle, mais n'est pas discriminant pour l'identification du type monétaire.

2.2.2 Approches par descripteurs locaux

En cas de forte dégradation et/ou de forte déformation, ce qui est souvent le cas pour les monnaies anciennes, effectuer une mesure de similarité globale n'est pas très efficace. Cependant, les monnaies appartenant au même type monétaire devraient avoir plus de régions locales similaires que celles appartenant à des types monétaires différents.

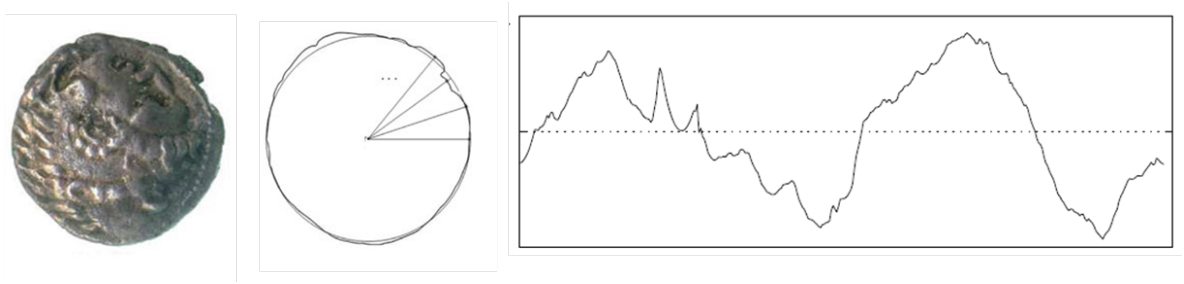


FIGURE 2.13 – Description de forme d'une monnaie ancienne à l'aide d'une courbe normalisée (Huber-Mörk et al. 2011)

Dans ce cas-là, une représentation de la monnaie en capturant les caractéristiques locales discriminantes semble plus avantageux qu'une représentation globale sensible aux bruits. Les approches basées sur des descripteurs locaux sont ainsi très souvent utilisées pour ce contexte compliqué.

L'un des descripteurs locaux les plus populaires est *Scale-Invariant Feature Transform* (SIFT) et ses variantes. Ils sont les plus utilisés dans la littérature grâce à la robustesse et l'invariance à la rotation. Le descripteur SIFT repose sur un histogramme des orientations du gradient (HoG). La description du point de caractéristique est donnée sous forme d'un vecteur. L'orientation principale du point de caractéristique a l'intérêt pour assurer l'invariance à la rotation du descripteur (Lowe 2004).

De nombreux autres descripteurs locaux sont largement utilisés dans la littérature de reconnaissance d'autres types d'objets, mais ils sont rarement utilisés pour identifier les types monétaires dans la littérature.

Concernant le contexte de l'identification du type monétaire, l'éclairage inhomogène est dominant dans certaines bases de données des monnaies anciennes. Pour mieux traiter cette situation, Zambanini et Kampel (2013) ont proposé un nouveau descripteur local, *local image descriptor robust to illumination changes* (LIDRIC), basé sur des filtres Gabor multi-échelles et multi-orientations.

Appariement de descripteurs locaux

Kampel et Zaharieva (2008) ont fait les premières tentatives pour reconnaître les monnaies anciennes en utilisant des descripteurs locaux. L'identification est donnée par une mesure de similarité entre les monnaies requêtes et les monnaies de référence, en comptant le nombre de points de caractéristiques SIFT appariés sans aucune contrainte de spatialité. Le taux de réussite obtenue est de 90% sur seulement trois types monétaires (*cf.* Figure 2.14).

Basée sur les descripteurs SIFT, Zambanini et Kampel (2011) ont proposé le flux SIFT (*SIFT-flow*) (Liu, Yuen et Torralba 2011) qui aligne les images en minimisant une fonction d'énergie basée sur les descripteurs SIFT extraits sur une grille dense. L'énergie

nécessaire pour aligner deux images similaires devrait être plus faible que celle pour aligner deux images différentes. Cette technique qui prend en compte l'arrangement spatial de l'ensemble de descripteurs locaux permet d'obtenir des taux de réussite beaucoup plus élevés que par le seul appariement de descripteurs locaux, mais bien évidemment, avec un coût de calcul beaucoup plus élevé. Pour réduire le temps de calcul, Zambanini et Kampel (2012) ont proposé un schéma de *corse-to-fine* sur une pyramide de l'image où les correspondances entre descripteurs locaux sont calculées d'abord sur une couche à résolution faible, puis affinées sur des couches à résolutions de plus en plus élevées. Grâce à cette optimisation, la complexité de l'algorithme est réduite de $O(h_4)$ à $O(h^2 logh)$, où h est la largeur de l'image.

(a) (b)

Figure 2.14 – Points d'intérêt SIFT appariés (Kampel et Zaharieva 2008) : (a) sur deux monnaies appartenant au même type monétaire ; (b) sur deux monnaies appartenant à des types monétaires différents

Pour améliorer encore la performance, Zambanini, Kavelar et Kampel (2014) ont utilisé un nouveau descripteur LIDRIC qui s'adapte mieux à leurs données, et ont évalué la similarité par consistance géométrique entre les descripteurs locaux appariés.

Sac de mots visuels

Dans le cadre des approches « sac de mots visuels », la monnaie est représentée par une distribution des descriptions locales, appelées mots visuels. Ce genre de technique, introduit à la base pour le traitement du langage naturel, permet de décrire de façon robuste des photos de monnaies par descripteurs locaux. L'idée principale est que ces descripteurs contiennent des informations locales extraites d'un nombre important d'images qui permettent de générer un vocabulaire visuel de taille fixe, à l'aide d'une méthode de *clustering*. Pour une nouvelle image, chaque descripteur local détecté est associé à un mot visuel du dictionnaire. La description de l'image est ainsi donnée par l'histogramme des occurrences de chaque mot visuel, dont la taille est celle du dictionnaire.

Dans les travaux de Anwar, Zambanini et Kampel (2013, 2015a, 2015b) et Anwar et al. (2015), le vocabulaire visuel est construit à l'aide des descripteurs SIFT extraits d'une grille dense, en utilisant une approche k-plus proches voisins (*cf.* Figure 2.15). Pour

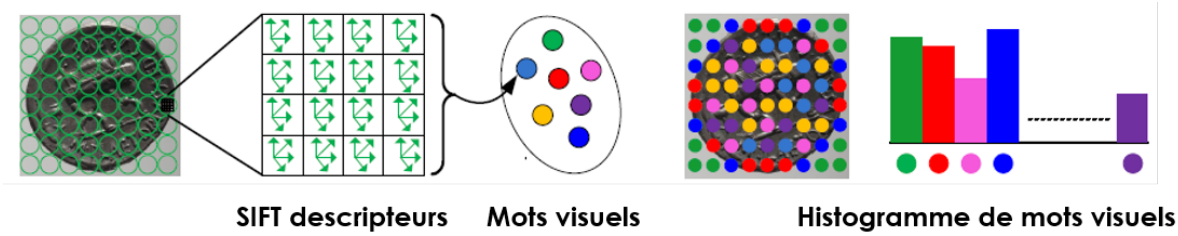


FIGURE 2.15 – Schéma général du fonctionnement d'un sac de mots visuels (Anwar, Zambanini et Kampel 2013)

capturer l'information concernant les relations spatiales entre les mots visuels, différentes répartitions (rectangulaire, circulaire, log-polaire) de l'image sont évaluées. Ainsi, l'histogramme des occurrences de mots visuels est obtenu par concaténation des histogrammes de chaque partie de l'image. Avec la répartition rectangulaire, cette méthode permet d'obtenir 90% de taux de réussite, pour classifier 8 motifs principaux (griffon, louve, *etc.*) sur le revers des monnaies anciennes. Pour la répartition circulaire, une amélioration a été proposée en introduisant *paire-wise identicial visual words*, afin de mieux profiter des relations spatiales (Anwar, Zambanini et Kampel 2014).

Kim et Pavlovic (2014) ont construit un vocabulaire visuel basé sur une répartition log-polaire centrée sur le portrait de l'empereur au lieu du centre physique de la monnaie. Pour faire cela, les monnaies sont alignées en localisant le portrait de l'empereur à l'aide de *Deformable Part Model* (DPM) (Felzenszwalb et al. 2010). Pour encoder directement les informations spatiales, une autre approche est proposée par les mêmes auteur en utilisant un vecteur de Fisher décrivant les changements de gradients.

Arandjelovic (2010) a exploité la configuration spatiale des mots visuels autrement. Dans sa méthode, l'apparence de l'image est considérée dans chaque point d'intérêt détecté. Ainsi, un *locally-biased Directional Histogram* (LBDH) est proposé pour encoder la distribution des mots visuels SIFT selon huit directions autour du point. Les caractéristiques LBDH sont ensuite utilisées pour reconstruire le vocabulaire visuel.

Parmi toutes les techniques basées sur les descripteurs locaux que nous avons présentées, la spatialité des descripteurs locaux joue un rôle important pour assurer la robustesse des algorithmes. Grâce à une rigidité relative des motifs d'un même type monétaire, y compris sur les monnaies anciennes, les relations spatiales des descripteurs locaux établissent également une transition entre les approches locales et les approches globales, en se basant sur les caractéristiques locales, mais en même temps en considérant une configuration globale. La principale difficulté est de parvenir à convertir des données brutes en une représentation en capturant les caractéristiques locales pour décrire les propriétés de l'image dans les régions spécifiques. Par opposition, les caractéristiques globales représentent souvent l'objet par un seul vecteur de caractéristiques.

2.2.3 Approches basées sur la lecture des légendes

Dans la plupart des approches globales ou locales mentionnées précédemment, les légendes ne sont pas traitées différemment des autres motifs de la monnaie. Parmi les types monétaires d'un pays issus d'une même période, la puissance discriminante des légendes peut être utilisée. Plus que leur contenu, c'est aussi leur apparence (la taille, la position, l'inclinaison, la police, *etc.*) qui peut être intéressant à analyser. Par exemple, nous pouvons trouver *UNITED STATES OF AMERICA* et *E PLURIBUS UNUM* sur différents types monétaires américains (*cf.* Figure 2.16 (a)). Par conséquent, les informations textuelles obtenues par détection et reconnaissance de ces légendes communes ne permettent pas d'identifier les types monétaires. D'un côté, il est donc plus intéressant et rentable de décrire les apparences visuelles de ces légendes en les considérant comme des motifs plutôt que comme du texte ; de l'autre côté, la détection et la reconnaissance des légendes de la monnaie sont un véritable défi peu abordé dans la littérature car les caractères sont souvent de la même couleur que le fond, avec la présence de bruits liés au matériau (la dégradation, la patine, *etc.*) et de défauts dans les photos (reflets et ombres). Cela peut expliquer pourquoi la lecture des légendes n'est pas un sujet très traité par les chercheurs qui étudient les monnaies « modernes » notamment.

Cependant, la lecture des légendes a un intérêt pour reconnaître des monnaies anciennes pour lesquelles le texte des légendes diffère plus que d'autres motifs (*cf.* Figure 2.16 (b)). Par exemple, les noms des empereurs inscrits sont uniques alors que leurs portraits subissent des variations intraclasses (Arandjelović 2012). Cependant, les méthodes classiques de reconnaissance de caractères exigent très souvent une bonne segmentation des caractères, et sont donc difficilement applicables dans ce contexte. Dans la littérature, les travaux concernant la reconnaissance de caractères, la détection et la reconnaissance des légendes de monnaies sont souvent signalées comme des situations difficiles conduisant à un échec. Il en est d'ailleurs de même avec la reconnaissance des lettrines qui se trouvent dans des ouvrages anciens. Une lettrine est considérée comme un objet graphique composé d'une couche de caractères, une couche de motifs aux alentours et une couche de fond (Coustaty et al. 2011). Dans ce cadre, la difficulté réside en l'extraction des différentes couches. Dans le cas des caractères extraits des monnaies, la principale difficulté est le fait que le premier plan et l'arrière-plan aient des propriétés visuelles similaires (couleur, texture, *etc.*). Il n'est donc pas envisageable, dans ce contexte, d'utiliser des techniques classiques pour la reconnaissance de caractères.

Pour éviter une extraction compliquée du premier plan, une solution proposée par les auteurs est de considérer l'imagette entière contenant un caractère comme un motif spécifique, et donc la légende comme une série d'imagettes ordonnées. L'idée générale est alors basée sur l'apprentissage d'une base d'imagettes de motifs spécifiques étiquetés. Certaines informations *a priori*, par exemple des zones susceptibles de contenir les légendes

FIGURE 2.16 – (a) Légendes comportant le même texte (*E PLURIBUS UNUM*) avec des apparences variées sur de différents types monétaires « modernes » ; (b) légendes avec des textes différents qui différencient les types monétaires similaires.

et un lexique, servent à améliorer les performances.

Arandjelović (2012) a encodé l'apparence d'un caractère individuel par un *HoG-like* descripteur. Un modèle spatial qui limite les positions relatives des caractères voisins est utilisé pour achever le placement optimal des caractères sur la légende de l'avers. La légende extraite est ensuite recherchée à l'aide d'un lexique des légendes connues grâce à un appariement de motifs visuels et à une analyse des méta-données. Dans son étude, les légendes sont supposées être placées de façon circulaire à l'intérieur de la bordure (listel) de la monnaie, donc horizontalement en coordonnées polaires. L'auteur précise qu'il a réussi à retrouver les identifiants de 24 des 25 types monétaires *Denarii*.

De manière similaire, Kavelar, Zambanini et Kampel (2012, 2014) ont encodé l'apparence de l'imagette de caractère par un seul descripteur SIFT et ont utilisé un SVM pour la classification. Par rapport au travail de Arandjelović (2012), l'orientation des légendes n'est pas donnée *a priori* mais détectée à l'aide de SIFT. L'hypothèse faite est qu'une légende est linaire et composée de caractères avec des orientations homogènes quelle que soit la rotation de la monnaie. Dans une étape supplémentaire, les légendes détectées sur d'autres motifs sont rejetées. Pour un lexique de taille de 35, le taux de reconnaissance des légendes est de 29% à 67% selon la paramétrisation du descripteur SIFT.

En conclusion, la description de l'imagette de caractère comme motif spécifique et la configuration spatiale des imagettes de caractères sont deux points essentiels dans les solutions proposées. Elles peuvent être considérées comme des approches adaptées des caractéristiques locales.

Avec la forte présence de bruits, les algorithmes d'apprentissage même spécifiques n'arrivent pas à obtenir des taux de reconnaissance très élevés. Les résultats de la reconnaissance des légendes sont donc souvent intégrés dans un pipeline d'identification du type monétaire utilisant également d'autres caractéristiques. Soit ils servent à présélectionner des monnaies de référence, soit ils sont pris en compte pour la classification par fusion. À l'aide de la lecture des légendes, Zambanini, Kavelar et Kampel (2013) ont augmenté leurs résultats de 78,9% à 81% ; Arandjelović (2012) a obtenu 24 sur 25 types monétaires correctement identifiés par rapport à 22 en utilisant uniquement l'appariement des motifs du revers.

Il faut noter toutefois que, bien que les apparences des caractères sur les monnaies anciennes étudiées soient différentes en fonction de l'illumination et de la dégradation, les polices et les tailles des caractères utilisées ne sont pas très variées. Par ailleurs, le nombre de types monétaires pris en compte est limité. Dans le cas où les apparences d'un caractère varient significativement sur un nombre important de types monétaires, une base d'apprentissage conséquente est nécessaire et peut s'avérer difficile à exploiter.

Dans le paragraphe qui suit, nous nous intéresserons à une tâche de gradation qui intègre non seulement la tâche d'identification du type monétaire, mais également la détection « d'éléments inattendus » ainsi que de quantification de l'usure.

2.3 Analyse d'images pour la gradation des monnaies

Pour grader une monnaie, il s'agit non seulement de l'identifier mais également d'évaluer son état de conservation. En ce qui concerne le dernier point, l'objectif principal est de détecter les « éléments inattendus » comme les coups et les rayures, par opposition aux « éléments attendus » comme les légendes et les motifs. Une monnaie ayant un grade élevé devrait comporter la totalité des « éléments attendus » et le moins possible « d'éléments inattendus ». La localisation de ces « éléments attendus » est spécifique à chaque type monétaire, donnée par une monnaie de référence « parfaite ». La faisabilité de l'évaluation d'autres facteurs comme la patine, voire l'intérêt esthétique (*eye appeal*) par analyse d'images, exigeant une véritable expertise numismatique, reste encore à étudier.

2.3.1 Brevets et produits industriels de la gradation

Au début des années 1990, des brevets liés à des processus de gradation assistés par ordinateur ont été déposés (Merton 1990 ; Crain 1993). Il s'agit de systèmes semi-

automatiques comportant l'ensemble des configurations mécaniques, électroniques et informatiques, concernant une série d'opérations de traitement d'images bas niveau.

Dans le système de Merton (1990), la gradation est principalement basée sur la quantification des « éléments inattendus ». Chaque « élément inattendu », nommé *detracting mark*, est détecté, localisé, et ensuite pondéré par sa position et son aire. La somme pondérée des « éléments inattendus » sur chaque face de la monnaie donne un score qui sera associé à un grade à l'aide d'une base de données score-grade préétablie. Les scores des deux faces de la monnaie avec les identifiants des positions des « éléments inattendus » sont considérés comme l'empreinte numérique de la monnaie à grader. La patine, analysée par microscope, est incluse dans le grade estimé.

Le système interactif conçu par PCGS (Crain 1993) a pour objectif la reproduction du processus effectué par un gradeur professionnel. Pour rendre compte de la visibilité des différents « éléments inattendus » selon les conditions d'éclairage et réaliser une analyse fine de la patine et de la nature des défauts, plusieurs photos de la monnaie à grader sont prises dans différentes conditions d'éclairage, puis alignées sur l'image de référence. Normalement une monnaie comporte des régions de hautes fréquences, par exemple, les motifs et les légendes, et des régions de basses fréquences, par exemple, le champ. Comme un « élément inattendu » dans une région de basses fréquences est plus facile à détecter, y compris par l'œil humain, que dans une région de hautes fréquences, il semble logique d'étudier séparément ces deux types de régions. Ainsi, les « éléments inattendus » sont tout d'abord recherchés sur les régions des basses fréquences, puis sur les régions des hautes fréquences. Sur les régions qui comportent également des « éléments attendus » fins, par exemple les cheveux. La patine est analysée à l'aide de la transition de couleurs. Le grade final est enfin obtenu par combinaison des différentes analyses avec un grand nombre de paramètres. L'expertise de gradeurs professionnels est essentielle pour fixer ces paramètres.

La gradation peut aussi être considérée comme un problème de reconnaissance d'objets où les monnaies de même grade correspondent à une classe. Ainsi, ce problème serait traité soit par l'apprentissage via des classifieurs, soit par appariement à une référence. En considérant le nombre de données d'apprentissage nécessaire pour la première méthode, la seconde approche semble plus réaliste. En effet, pour des monnaies de collection par exemple, il est très difficile d'avoir des données suffisantes pour l'entraînement des classifieurs. Il n'en reste pas moins qu'il faut avoir quand même une image de référence par type monétaire et par grade. Basée sur cette idée, la société américaine PCGS a récemment proposé une application libre *Photograde*^{TM}, en ligne (*cf.* Figure 2.17) et sur mobile, qui permet aux utilisateurs d'estimer approximativement le grade d'une monnaie américaine en comparant à l'œil nu la monnaie à grader avec les images de référence. L'utilisateur est supposé connaître le type monétaire de sa monnaie à grader, sinon il doit chercher manuellement le type monétaire dans une centaine de répertoires. Bien que *Photograde*^{TM}ne soit

qu'une galerie d'images sans intelligence artificielle, une collection des photos professionnelles pour tous les grades d'une centaine de types monétaires américains, photographiées et prétraitées par un même prototype, elle n'existe nulle part ailleurs. Il faut noter que *Photograde*^{TM} est gratuite et qu'elle ne fournit que des images d'une résolution faible d'environ 500 × 500 pixels. La comparaison à l'œil nu est plutôt basée sur l'apparence globale que sur les détails.

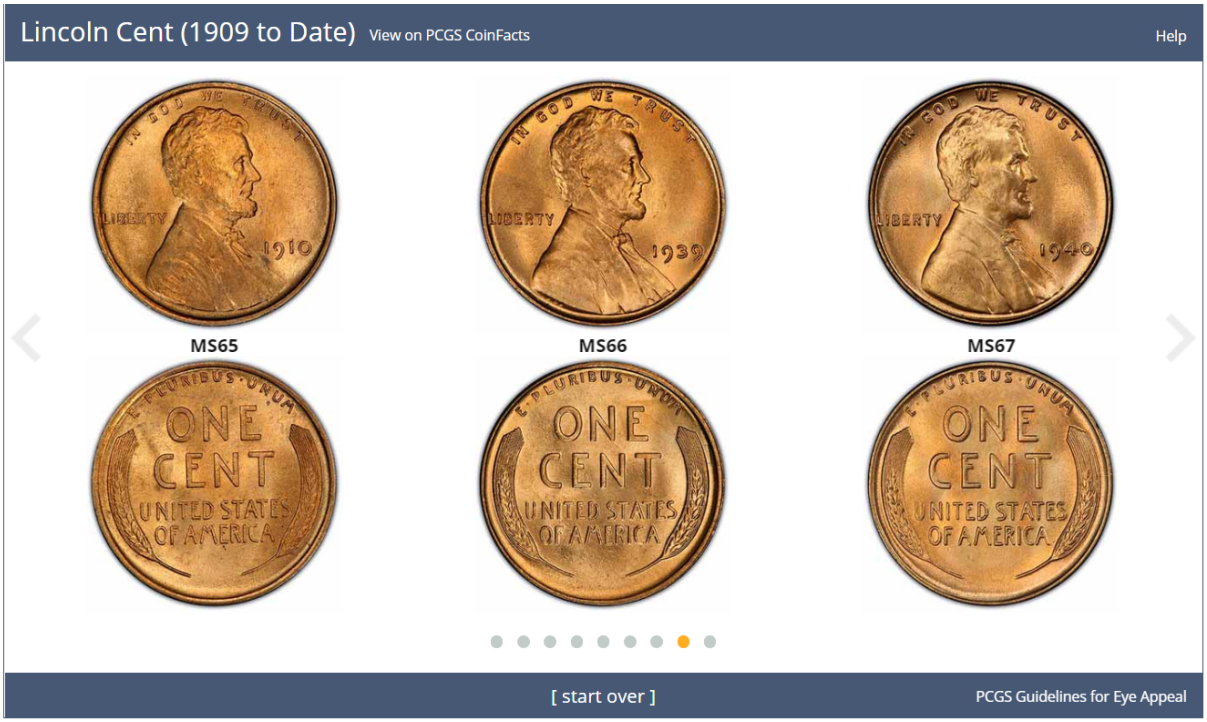


Figure 2.17 – Capture d'écran de *Photograde*^{TM} en ligne : images de référence représentant un *Lincoln Cent* avec le grade MS-65, MS-66 et MS-67.

2.3.2 Travaux de recherche sur la gradation

Dans la communauté des chercheurs en traitement d'images, la gradation de monnaie est un sujet peu traité en raison, probablement, du manque de données de qualité et du manque d'expertise numismatique. Concernant les données, tous les détails de la monnaie qui peuvent être analysés par l'œil humain, même avec une loupe, devraient être clairement présentés dans les images. Cependant, presque toutes des images de monnaies collectées et utilisées dans les travaux précédents ne sont pas éligibles par leur faible résolution ou à cause de conditions d'éclairage inadéquates. Concernant l'expertise numismatique, il faut que les données soient annotées attentivement par des gradeurs professionnels, ce qui n'est pas facile à obtenir.

La patine de la monnaie peut être étudiée via la couleur de l'image. Dans un workshop d'étudiants, Bassett et al. (2002) ont essayé de grader les monnaies *Lincoln Cent* en utilisant l'histogramme de couleur, basé sur l'hypothèse que les monnaies à l'état de frappe

sont brillantes et les monnaies usées sont sombres. Cette hypothèse n'est pas toujours exacte car il existe aussi des monnaies sombres en état Fleur de coin. Ils ont utilisé de petites images scannées sans tenir compte d'aucun motif. Sur la Figure 2.17, nous voyons bien que la couleur n'est pas une caractéristique discriminante des monnaies avec un grade relativement élevé. Par ailleurs, la couleur changera de manière significative en cas de présence de patine. Même si ce rapport technique met en avant beaucoup de limitations, il semble que ce soit l'une des premières tentatives visant à grader des monnaies par analyse d'images.

Dans l'algorithme de reconnaissance du système *Dagobert* (Nölle et al. 2003), le concept des « éléments attendus » et des « éléments inattendus » a été pour la première fois introduit. ils sont donnés par l'énergie d'abrasion et l'énergie de saleté, calculées par comparaison des cartes des contours. Celles-ci sont utilisées comme deux seuils pour présélectionner les monnaies de référence associées.

L'influence du grade sur le problème de l'identification du type monétaire a été examinée dans une étude récente (Fare et Arandjelović 2017). Plusieurs méthodes basiques de la littérature (distance d'histogrammes, sac de mots visuels basés SIFT, LBP, *etc.*) sont évaluées respectivement sur les monnaies anciennes avec des grades différents (TB, TB et SUP, selon le système de gradation européen (*cf.* Table 1.3). À cause de la complexité des données et de la simplicité des méthodes évaluées, les taux d'identification sont faibles (moins de 10%) par rapport à ce que nous avons présenté dans la section 2.3. Cependant, l'objectif de cette étude est d'analyser la corrélation entre le taux d'identification et le grade de la monnaie requête. Selon les expériences conduites par les auteurs, les résultats sur les monnaies avec le grade le plus faible (TB) sont, sans surprise, moins bons, mais les taux d'identification sur les monnaies avec le grade le plus élevé (SUP) sont généralement plus faibles que sur les monnaies avec un grade moyen (TB). Les auteurs expliquent que les monnaies avec un grade élevé possèdent probablement plus de détails particuliers, ce qui accentue la variabilité intraclasse dans les monnaies anciennes.

Pour conclure, la gradation des monnaies par analyse d'images n'est encore pas vraiment étudiée par les chercheurs. Les sociétés numismatiques, quant à elles, en dépit de quelques brevets déposés il y a longtemps, font toujours appel à l'intervention coûteuse des gradeurs professionnels. D'un côté, il est difficile de quantifier exactement tous les facteurs contributifs ; d'un autre côté, la gradation coûteuse et subjective correspond bien à un marché de niche. Cependant, il semblerait que la communauté numismatique dans sa globalité ne soit pas motivée pour une sophistication du processus. La société GENI, au contraire, a bien compris l'intérêt d'une telle automatisation.

2.4 Conclusion et aspects spécifiques de la thèse

Dans ce chapitre, nous avons dressé un état de l'art des diverses étapes de la numismatique par analyse d'images : segmentation des monnaies, identification du type monétaire et premières tentatives de gradation des monnaies.

En fonction des caractéristiques des monnaies « modernes » et des monnaies anciennes, les stratégies choisies sont complètement différentes. Il faut noter que les monnaies « modernes » traitées dans les travaux précédents sont généralement des monnaies de circulation, et non pas des monnaies de collection qui sont l'objet de cette thèse. La segmentation est réalisée par les techniques basiques en supposant que les monnaies à traiter sont toutes rondes et bien différentes de l'arrière-plan. Le type monétaire est identifié plutôt par des approches basées sur une description globale de l'ensemble des motifs. La plupart des images, utilisées pour ce cas, sont de faible résolution. Comme dans ces études la priorité est l'efficacité et la facilité d'implémentation, l'influence de la qualité des photos sur la performance n'est pas discutée.

Concernant les monnaies anciennes, les méthodes de segmentation sont plus compliquées étant donné leur forme irrégulière, souvent composées de plusieurs étapes. Pour identifier le type monétaire, les approches sont principalement basées sur des caractéristiques locales, parfois exploitant les légendes. Ces méthodes sont plus robustes que celles exploitant une description globale. Bien que certaines études soient motivées à des fins de collection, les photos utilisées n'ont pas une qualité professionnelle telle qu'indiquée en section 1.3. Par conséquent, certaines études s'intéressent aux bruits liés à la prise de photos, ce qui ne correspond pas au contexte de cette thèse.

L'objectif de cette thèse est de reproduire, à terme, le processus humain utilisé par les sociétés numismatiques. Il s'agit de développer une méthode automatique, en profitant de leur expertise en photographie numismatique et en gradation. À la différence des monnaies « modernes » de circulation et des monnaies anciennes étudiées dans la littérature, nous nous concentrons sur les **monnaies « modernes » de collection**, en utilisant des **photos professionnelles numismatiques**. Ce genre de données est exploré pour la première fois.

Grâce à la qualité des photos professionnelles, les bruits dans l'image sont minimisés et l'orientation des monnaies est relativement proche. Dans ce contexte, l'objectif de notre travail est ainsi orienté vers le développement d'approches qui permettent de segmenter très précisément la monnaie et d'en extraire les détails les plus fins. Nous terminons ce chapitre par lister les défis apportés par le contexte de cette thèse par rapport à la littérature :

- Détourage précis des monnaies de différentes formes

Les négociants en monnaies de collection ont pour habitude de présenter leurs produits dans des catalogues de vente dans lesquels la monnaie est présentée dans une photo segmentée à la main. Dans ce contexte, la segmentation est faite très précisément (au pixel près) de manière manuelle. Donc, un premier challenge consiste à reproduire de manière automatique cette tâche qui prend beaucoup de temps. Si la segmentation « grossière » est aisée, la plupart des méthodes utilisées dans la littérature n'arrivent cependant pas à atteindre la très grande précision exigée.

Il s'agit en effet d'extraire uniquement les pixels de la monnaie et l'ensemble de ces pixels doit produire une forme parfaite. De même, pour une analyse fine des détails de la bordure de la monnaie, il s'agit d'extraire les pixels qui représentent le moindre détail de la forme de la monnaie comme le listel.

En numismatique, la segmentation des monnaies s'appelle le détourage. Il est important que les monnaies détourées présentées dans le catalogue de vente aient des formes parfaites, en particulier, pour les monnaies de forme régulière. Même si la prise de photos est maîtrisée, il reste toujours des éléments comme de petites ombres portées et de petites déformations causées par le fait que le plan focal n'est pas parfaitement parallèle au plan de l'image. Cette imperfection de la forme réelle générera des difficultés pour obtenir un détourage « parfait », au pixel près. De même, les monnaies à pans, les monnaies à dents, les monnaies avec une bordure cannelée ou striée font partie des monnaies de collection. Il est également important de les détourer et, pour cela, de reconnaître leur forme de manière précise, mais cela reste très souvent un challenge. Par exemple, une monnaie à 24 pans ressemble vraiment à une monnaie ronde y compris à l'œil humain sur une image zoomée (*cf.* Figure 2.18). La nouvelle monnaie de 1 livre d'après « Brexit » est celle à 12 pans et avec 6 pans sinusoïdaux (*cf.* Figure 2.19). De tels détails expliquent la nécessité d'utiliser des photos professionnelles. Comme nous l'avons mentionné, dans la pratique les numismates utilisent des outils comme GIMP ou Photoshop pour réaliser un détourage précis qui correspond aux besoins, mais avec un coût de temps phénoménal.

- Très grande similarité interclasse

Dans le cadre de cette thèse, la grande similarité entre différents types monétaires est due au fait qu'il existe des « sous-types » (variétés et variantes) avec uniquement des modifications subtiles. Les motifs quasi-identiques entre monnaies de différents types monétaires sont très difficiles à distinguer même à l'œil nu. Dans la plupart des cas, ces motifs quasi-identiques, ou identiques apparaissent uniquement sur un des côtés de la monnaie, mais ils peuvent également apparaître sur les deux côtés. Dans certaines études précédentes sur les monnaies « modernes », les types monétaires avec une similarité interclasse élevée sont simplement considérés comme une seule classe. Cependant, pour les monnaies considérées dans cette thèse, il est important de les distinguer.

- Variabilité intraclasse en fonction de grade

Une déformation non-rigide intraclasse dans les monnaies anciennes est mise en évidence par Kampel et Zaharieva (2008) (*cf.* Figure 2.2). En ce qui concerne les monnaies « modernes », la frappe d'un même type monétaire est standardisée, mais la variabilité intraclasse provient principalement des différents niveaux d'état de conservation de la monnaie, autrement dit son grade. Les motifs en relief sur une monnaie usée sont évidemment moins saillants que ceux qui sont sur une monnaie bien conservée du même type monétaire. Par ailleurs, à cause de la patine, les monnaies d'un même type monétaire peuvent avoir des variations dans la couleur globale ou locale. Par conséquent, la similarité de deux monnaies appartenant au même type monétaire avec des grades bien différents peut être plus faible qu'entre deux monnaies de grades proches mais appartenant à différents types monétaires, notamment lorsque leurs motifs sont similaires.

- Manque de données

Une base de données numismatique idéale serait composée de nombreux échantillons pour chaque type monétaire et avec des exemples de différents grades. Mais dans la réalité, il n'est vraiment pas facile de construire une base de données contenant suffisamment de données éligibles même avec l'aide de sociétés numismatiques. Cela est non seulement dû à un problème de disponibilité des monnaies rares mais aussi à la complexité de la prise de photos et de l'annotation qui rendent la construction d'une telle de base de données difficile, voire impossible. Par ailleurs, le nombre total de types monétaires considérés est très important. Nous l'estimons à plusieurs dizaines de milliers selon *Standard Catalog of World Coins* (Cuhaj et Micheal 2015 ; Judkins et Micheal 2016a, 2016b). D'autant plus que de nouveaux types monétaires sont inclus sans arrêt (découverte de monnaies rares jamais répertoriées, types monétaires nouvellement proposés), il est impossible d'utiliser des techniques qui se basent uniquement sur l'apprentissage machine. Il est au contraire nécessaire d'apporter de « l'intelligence manuelle » comme nous le verrons dans la suite de cette thèse.

FIGURE 2.18 – Avers de l'Essai 10 centimes Merley à 24 pans (bord zoomé encadré en rouge)

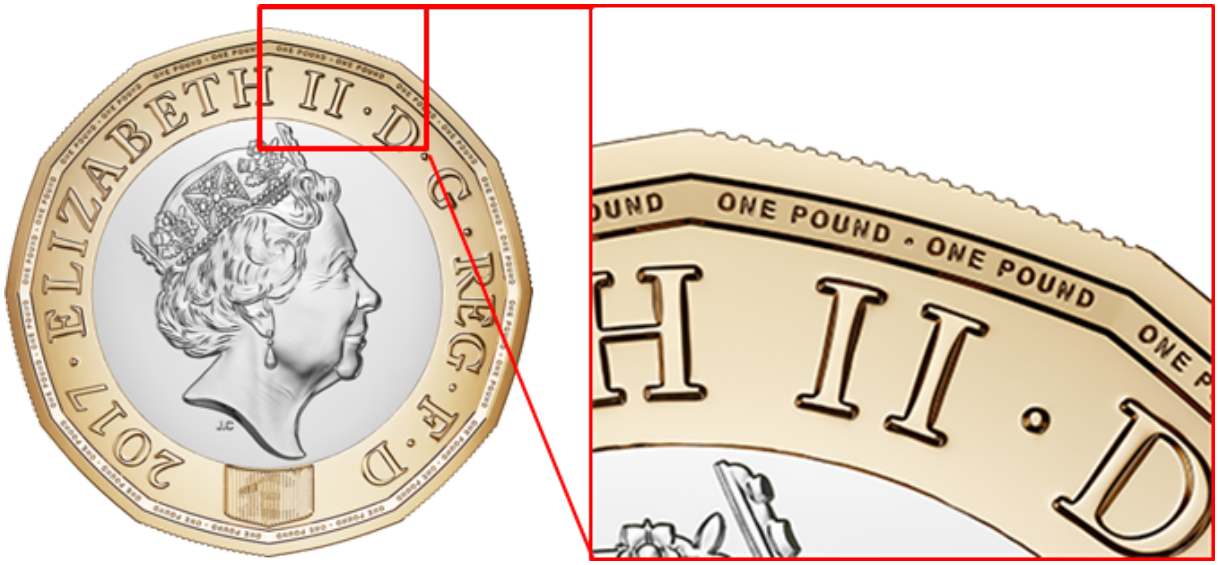


FIGURE 2.19 – Avers de l'Essai 1 livre à 12 pans (bord zoomé encadré en rouge)

Chapitre 3

Bases de données numismatiques

Ce chapitre présente différentes bases de données utilisées et proposées dans le cadre de cette thèse. Dans la section 3.1, nous introduisons et comparons les bases de données fréquemment mentionnées dans la littérature et celles construites pour les différentes phases de notre travail, la segmentation, l'identification du type monétaire et la gradation. Dans la section 3.2, nous présentons la base de données de chiffres extraits de monnaies que nous avons construite à l'aide de différentes approches d'extraction. La section 3.3 contient la conclusion et les perspectives concernant les bases de données présentées.

3.1 Base de données d'images de monnaie

Comme nous l'avons déjà mentionné dans le Chapitre 2, la construction d'une base de données de photos de monnaies de qualité n'est pas une tâche facile. Une grande collection de monnaies, des techniques avancées et des conditions contrôlées de prises de vues graphiques sont nécessaires. Par ailleurs, une expertise numismatique est indispensable pour annoter les données. Avec toutes ces contraintes, les images de monnaies peuvent être obtenues de manière générale selon trois approches :

- Images réalisées sous conditions de laboratoire : certains chercheurs se satisfont avec leur propre base de données obtenue à l'aide d'un scanner ou d'un appareil photo. Nous qualifierons les photos prises par les chercheurs eux-mêmes comme étant des photos amateurs. Le nombre d'images et de types monétaires sont en général limités. L'un des avantages est que la base de données peut être construite en fonction des besoins de chaque étude si les conditions le permettent.

- Téléchargement via Internet : De nombreuses images de monnaies sont sans cesse déposées sur Internet par des amateurs numismates, des commerçants de monnaies, *etc.* Grâce aux moteurs de recherche, tout le monde peut accéder aux images de monnaies même rares. Nous pouvons également trouver des exemples de photos professionnelles réalisées par des musées ou des sociétés numismatiques. Les images de monnaies téléchargées sont souvent utilisées pour construire des bases de données qui contiennent notamment des monnaies que les chercheurs ne possèdent pas. Parfois, les données viennent de plusieurs sources. Par conséquent, la qualité des images n'est pas standardisée. Par exemple, la base de données des monnaies anciennes de l'Empire romain utilisée dans le travail de Arandjelovic (2010) correspond à cette catégorie (*cf.* Figure 3.1).

- Photos professionnelles ou quasi-professionnelles : pour avoir des bases de données de taille importante avec des photos de qualité supérieure, il est idéal d'utiliser des photos professionnelles ou quasi-professionnelles, fournies par des musées ou des sociétés numismatiques.

Figure 3.1 – Exemples de 65 types monétaires extraits de la base de données constituées par Arandjelovic (2010)

3.1.1 Bases de données utilisées dans la littérature

Suite à l'intérêt croissant porté à la reconnaissance de monnaies, plusieurs bases de données de photos de monnaies ont été construites. Cependant, certaines bases de données ne contiennent que peu d'images de test de qualité médiocre. Ces images ne correspondent pas au contexte de notre travail. Par ailleurs, la plupart des bases de données construites par les chercheurs ne sont pas publiées. Dans les paragraphes suivants, nous allons présenter deux bases de données des monnaies présentées dans les études sur les monnaies « modernes » et les monnaies anciennes qui ont attiré notre attention.

MUSCLE Coin Image Seiberdorf (CIS)

La base *MUSCLE CIS* est la base de données numismatique la plus largement utilisée concernant les monnaies « modernes ». La création de *MUSCLE CIS* a été motivée par le changement de monnaies de circulation en euro dans les 12 pays européens en janvier 2002. Dans ce contexte, le système d'images *Dagobert* a été proposé pour trier des monnaies de manière automatique. À partir de la base d'apprentissage collectionnée dans le système *Dagobert*, la base de données *MUSCLE CIS* a été construite avec 60000 images acquises à partir de 30000 monnaies. Les deux faces d'une monnaie font deux images différentes. Toutes les images dans cette base sont divisées en 692 types monétaires (définis par deux images) et 2270 classes « faciales ». Il y a plus de deux fois plus de classes « faciales » que de types monétaires parce qu'un type monétaire peut contenir plusieurs variantes (ou « sous-types ») avec des motifs variés. Les images sont en niveaux de gris, avec une taille normalisée à 640×576 pixels. L'image est composée de deux parties : une partie principale, en haut, comportant la monnaie sur un fond sombre et quasi-uniforme, et une partie en bas, qui contient un texte noir sur un fond blanc. Le texte indique le diamètre, l'épaisseur et la vérité terrain du type monétaire de la monnaie dans l'image. Les monnaies

concernées sont principalement des monnaies « modernes » pré-euro.

Par rapport à d'autres bases de données, le nombre total d'images et le nombre total de classes sont significatifs. C'est la première base dans laquelle autant de monnaies ont été acquises et prétraitées de la même manière. En ce qui concerne la qualité des photos, les conditions d'éclairage sont bien contrôlées et le fond est quasi-uniforme. La résolution des images est cependant faible. En considérant la taille originale de l'image et un ratio objet-image relativement petit, la monnaie segmentée a un diamètre uniquement d'une centaine de pixels. Cela ne permet pas une analyse fine des détails de la monnaie. Un autre inconvénient est que les images sont en niveaux de gris. Certaines informations que nous pourrions déduire à partir de la couleur, comme le matériau, la patine, *etc.* ne sont ainsi pas disponibles dans cette base de données.

Il s'agit donc d'une base de données possible pour tester les algorithmes d'identification du type monétaire à l'aide de descripteurs globaux. Étant donné le nombre important de données contenues dans cette base, les chercheurs utilisent souvent une sous-partie de celle-ci dans leurs études. En s'appuyant sur cette base de données, deux compétitions ont été lancées en 2006 et 2007 pour chercher le meilleur algorithme d'identification du type monétaire. Les images utilisées dans la compétition en 2007 sont obtenues à partir des mêmes monnaies que celles en 2006 mais avec des occultations synthétiques. La Figure 3.2 donne des exemples de la base *MUSCLE CIS* 2006.

FIGURE 3.2 – Exemples d'images de la base *MUSCLE CIS* 2006

Monnaies anciennes de la République romaine

Évidemment, les monnaies anciennes sont plus difficiles à collecter que les monnaies « modernes ». Les monnaies anciennes les plus utilisées dans la littérature sont les monnaies romaines (République romaine et Empire romain) et les monnaies mérovingiennes. Une des bases de données de la République romaine fréquemment utilisée dans les premières études sur les monnaies anciennes est la base de (Kampel et Zaharieva 2008). Cette base de données comporte 180 monnaies de la République romaine divisée en 60 types monétaires, c'est-à-dire avec trois échantillons par type monétaire. Elle ne contient que

des images du revers de la monnaie. Un type monétaire est ainsi défini par une seule face de la monnaie. Les monnaies, fournies par le *Museum of Fine Arts Vienna*, sont toutes photographiées de la même manière. En général, la monnaie est sombre dans la photo et le fond est en couleur claire. La taille de l'image est entre 1000 × 1000 et 2000 × 2000 pixels en fonction de la taille de la monnaie contenue. Le ratio objet-image est grand, c'est-à-dire que la majorité des pixels appartient à la région de la monnaie. La rotation de la monnaie est aléatoire. La direction de l'éclairage est non-uniforme. Par conséquent, les motifs en relief d'un même type monétaire peuvent être visuellement différents non seulement à cause de la dégradation et de la frappe non-standardisée, mais aussi à cause de l'éclairage.

C'est une base de données utilisée pour l'identification du type monétaire ancien. À notre connaissance, elle est également la première base de données utilisée pour la reconnaissance de caractères dans les monnaies. La Figure 3.3 illustre des exemples de monnaies de la République romaine issus de cette base.

Figure 3.3 – Exemples d'images de monnaies de la République romaine appartenant à deux types monétaires (un type sur chaque ligne)

3.1.2 Bases de données utilisées dans le cadre de cette thèse

Les deux bases de données de monnaies que nous avons présentées dans le paragraphe précédent ne correspondent pas au contexte spécifique de cette thèse, et ce sont pourtant les deux plus intéressantes que nous ayons trouvées dans la littérature. Cependant, il est important d'avoir des bases de données de test correspondant bien aux besoins spécifiques de chaque phase du travail de cette thèse pour valider les différents apports. Concernant la segmentation des monnaies, nous devons avoir des monnaies de formes différentes. Quant à l'identification du type monétaire, nous devons avoir des monnaies appartenant aux nombreux types monétaires qu'il conviendra d'identifier. Il serait mieux que chaque type monétaire possède de nombreux échantillons avec des apparences différentes. Concernant la gradation de monnaie, l'idéal est d'avoir des monnaies appartenant au même type monétaire mais avec des grades différents et en particulier des monnaies de référence « parfaite », par exemple, en état « Fleur de coin ». Pour cette dernière phase, il est essentiel d'avoir des images avec une résolution élevée pour pouvoir analyser le moindre détail sur la monnaie.

Pour évaluer les algorithmes développés dans un véritable environnement numismatique, il aurait été souhaitable d'avoir une base de données pour le processus complet. Cependant, la société GENI n'a pas pu fournir une telle base pour le moment. En conséquence, plusieurs bases de données ont été construites correspondant aux études respectives. Suite aux activités de la société GENI, il est probable que, dans le futur, nous pourrons construire une base de données de photos professionnelles pour tout le processus de gradation.

USA_Grading

USA_Grading est une base de données provenant de *Photograde*^{TM} de la société américaine PCGS. Elle contient les séries principales des monnaies américaines des XVIIIe, XIXe et XXe siècles. Toutes les monnaies sont rondes. Elle a été initialement utilisée pour estimer grossièrement le grade d'une monnaie américaine en comparant avec les références proposées. Les images dans *Photograde*^{TM} sont des monnaies segmentées normalisées à la même taille sur un arrière-plan blanc dont la taille est de 500 $\times$ 500 pixels. Chaque image est annotée par son type monétaire, son côté (avers ou revers) et son grade. Des exemples de la base de données *USA_Grading* sont illustrés à la Figure 3.4.

Les principales caractéristiques de cette base de données sont les suivantes : une similarité interclasse élevée pour certains types monétaires, une variabilité intraclasse en fonction de l'état de conservation, et une distribution de classes quasi-égale. Plus précisément, chaque classe « faciale » possède une dizaine d'images. Les images contiennent des monnaies complètement usées jusqu'aux monnaies toutes neuves et brillantes suivant l'échelle de Sheldon. En outre, des types monétaires différents peuvent avoir des motifs

en relief extrêmement similaires sur certaines zones. Les orientations des monnaies sont similaires puisqu'elles sont alignées manuellement lors de la prise de photos.

Pour les monnaies de grade très faible, sur lesquelles les motifs qui caractérisent leur type monétaire sont presque effacés, nous pouvons à peine distinguer les motifs même à l'œil nu (*cf.* Figure 3.5). Comme ces monnaies « sans motifs » sont évidemment des bruits pour l'identification du type monétaire, nous ne considérons dans la suite que les monnaies avec les motifs « visibles ». Après avoir enlevé les monnaies « sans motifs », la base *USA_Grading* comporte 2598 images de monnaies divisées en 148 classes « faciales ». Chaque classe est unique mais certaines se ressemblent vraiment. Il faut mentionner que l'éclairage n'est pas uniforme pour toutes les monnaies dans cette base parce que la société PCGS souhaite jouer sur le reflet pour que les monnaies soient plus belles (*cf.* Figure 3.6).

Figure 3.4 – Exemples de monnaies de la base *USA_Grading* de grades différents appartenant à deux types monétaires (un type sur chaque ligne)

Figure 3.5 – Exemples de monnaies trop usées « sans motif » dans *Photograde*^{TM}, exclues de la base de données *USA_Grading*

FIGURE 3.6 – Éclairage non-uniforme sur des monnaies de grades proches du même type monétaire

GENI_CoinPhotos

Lors du déroulement de cette thèse, la société GENI était en train de développer de nouvelles techniques pour obtenir des photos professionnelles de haute qualité notamment pour la gradation par l'analyse d'images. Dans ce but, elle tend à minimiser tous les effets « externes » comme les ombres, les reflets et les points lumineux sur la photo ; et en même temps, à renforcer les éléments intrinsèques comme les motifs, les rayures, la patine, *etc.*

Grâce à la collaboration entre la société GENI et le laboratoire LIRIS, nous avons reçu des photos de test de la part de GENI. La base de données *GENI_CoinPhotos* est construite à partir de ces photos de test. Les images originales ont une résolution supérieure à 5000 × 3000 pixels. Les monnaies segmentées sont alors de 1000 × 1000 à 2000 × 2000 pixels selon leur taille réelle. L'arrière-plan est soit quasi-blanc, soit quasi-noir, parce que le photographe de GENI a testé l'influence de la couleur de l'arrière-plan sur le rendu couleur de la monnaie. La rotation de la monnaie est contrôlée manuellement. Autrement dit, les motifs principaux de la monnaie sont bien orientés dans le cadre de l'image. Des exemples de la base de données *GENI_CoinPhotos* sont illustrés à la Figure 3.7. Cette base de données contient actuellement 217 monnaies « modernes » européennes, y compris des monnaies Essai et des jetons, mais elle sera étendue par la suite. La majorité des monnaies sont rondes, mais elle comporte également des monnaies quasi-rondes avec un bord strié ou cannelé, et quelques monnaies à pans ou à dents.

LIRIS_CoinPhotos

Dans le cadre de cette thèse, nous envisageons le détourage automatique de monnaies rondes, mais également avec d'autres formes régulières comme les monnaies à pans ou à dents. Par ailleurs, certaines monnaies possèdent un trou au milieu dont la forme est ronde ou carrée.

D'après les numismates de la société GENI, pour les monnaies françaises, le nombre de pans pour les monnaies polygonales est le plus souvent de 8, 10, 11, 12, 16, 18, 20, 22 et

Figure 3.7 – Exemples de monnaies de la base de données *GENI_CoinPhotos*

24, et le nombre de crêtes pour les monnaies sinusoïdales est toujours de 12. N'ayant pas assez de photos professionnelles, pour constituer une telle base, nous avons construit une base de données qui contient des images semi-synthétiques. Pour cela, nous avons d'abord pris des photos de monnaies au laboratoire car ne possédant pas beaucoup de photos de la société GENI. Les monnaies collectées sont multinationales, de différentes tailles et de différentes couleurs. Toutes les monnaies sont rondes parce que c'est la seule forme que nous avions sous la main. Pour minimiser les ombres, les monnaies sont posées sur un papier noir, éclairées par deux sources de lumière. À cause des lampes jaunes utilisées, la couleur des monnaies photographiées est altérée par rapport à une lumière naturelle. Cependant, la modification des couleurs n'influence pas la segmentation. Les monnaies photographiées sont illustrées à la Figure 3.8. À partir de ces monnaies photographiées au laboratoire, nous avons ensuite fabriqué des « fausses monnaies » correspondant aux monnaies à pans et à dents avec un nombre de côtés ou un nombre de crêtes variable, et aux monnaies avec un trou au milieu (*cf.* Figure 3.9).

La base de données *LIRIS_CoinPhotos* comprend 250 images non-segmentées (vraies photos et images semi-synthétiques), comportant 19 formes différentes (monnaies polygonales à 8, 10, 11, 12, 16, 18, 20, 22, 24 côtés, sinusoïdales avec le même nombre de crêtes et rondes), certaines possèdent un trou rond ou carré au milieu. La taille d'image est de 1400 × 1400 pixels.

En dépit des conditions de laboratoire, la qualité des images est supérieure à la plu-

part des photos amateurs utilisées dans la littérature. À notre connaissance, ces photos constituent également la première base de données utilisée pour segmenter des monnaies « modernes » multiformes.

FIGURE 3.8 – Exemples de monnaies photographiées au laboratoire LIRIS

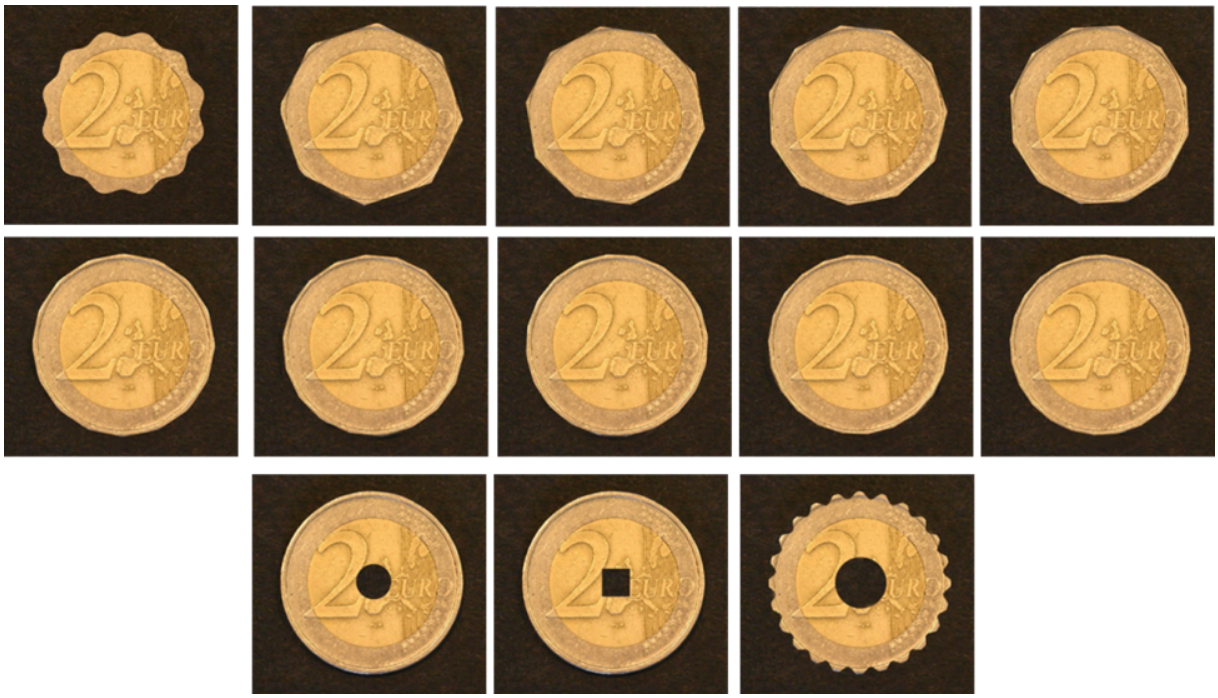


FIGURE 3.9 – Exemples de monnaies semi-synthétiques de formes variées à partir d'une vraie photo

20FrancsCOQ_Photos

La base de données *20FrancsCOQ_Photos* contient 286 images de monnaies du type monétaire 20 Francs COQ. Parmi celles-ci, 143 images contiennent l'avers dont le motif principal est le portrait de Marianne, et 143 images sont pour le revers dont le motif principal est le coq. C'est la raison pour laquelle ce type monétaire s'appelle « 20 Francs COQ » ou « 20 Francs Marianne ». Plus précisément, il existe deux variétés de 20 Francs COQ qui se différencient par la légende inscrite sur la tranche : « DIEU PROTEGE LA

FRANCE » ou « LIBERTE EGALITE FRATERNITE ». Les motifs sur l'avers et le revers sont quasiment identiques pour ces deux variétés (sauf le millésime). Comme la tranche discriminante n'est pas dans la photo, il est intéressant de reconnaître le millésime pour les distinguer.

Les photos originales sont réalisées par la société GENI, mais nous les avons segmentées et recalées (*cf.* Figure 3.10). La taille de la monnaie segmentée est normalisée à 2048×2048 pixels. Le grade n'est pas encore disponible pour cette base de données, mais les gros défauts (rayures, coups, taches, *etc.*) et les zones usées qui contiennent des micro-défauts ont été marqués de manière indicative par un numismate de la société GENI à l'aide de Photoshop (*cf.* Figure 3.11).

Figure 3.10 – Exemples de monnaies segmentées et recalées de *20FrancsCOQ_Photos*

3.1.3 Comparaison des bases de données numismatiques

Après avoir présenté les bases de données numismatiques de la littérature et celles proposées pour les différentes phases du travail de cette thèse, nous souhaitons faire un tableau comparatif (*cf.* Table 3.1) selon les critères suivants :

- Quantité : la quantité de données dépend largement de la collection. *MUSCLE CIS* est la base qui possède le plus de données.

- Résolution : dans la plupart des études sur l'identification du type monétaire, les chercheurs se satisfont d'images avec une résolution relativement faible. Par exemple, le

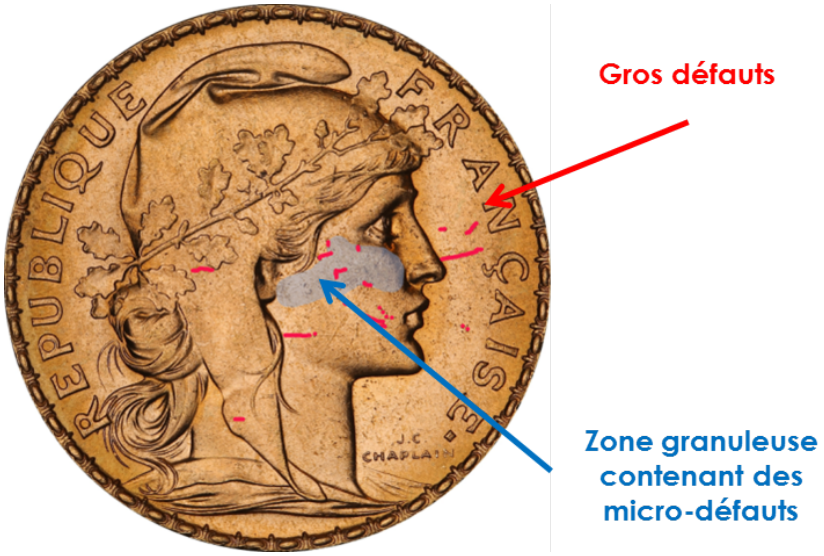


FIGURE 3.11 – Annotation des « éléments inattendus » sous des calques superposés à l'aide de Photoshop

diamètre d'une monnaie segmentée dans *MUSCLE CIS* est souvent d'une centaine de pixels. Évidemment, l'utilisation de petites images favorise le temps de calcul et l'espace de stockage. Pourtant, une résolution élevée est indispensable pour analyser les détails des monnaies.

- Segmentation : certaines bases de données sont composées d'images pré-segmentées. Pour la validation des méthodes de segmentation, il est bien évidemment nécessaire de travailler sur des images non-segmentées. En outre, pour une base de données comportant des petites images non-segmentées, les monnaies segmentées seront encore plus petites.

- Orientation : dans le cas des photos professionnelles, les monnaies sont intentionnellement bien orientées lors de la prise de photos, même si une petite fluctuation peut être introduite. Nous appelons cette situation « quasi-fixe ». Quant aux photos amateurs, l'orientation de la monnaie peut être complètement aléatoire. Concernant les données de la base de données *20FrancsCOQ_Photos*, une étape de recalage a été réalisée.

- Qualité : il s'agit une observation subjective. Nous utilisons le terme de qualité pour décrire si les éléments intrinsèques de la monnaie sont bien présentés et les bruits introduits sont minimisés. Évidemment, les photos professionnelles ont une meilleure qualité que les photos amateurs. Concernant la base *USA_Grading*, bien qu'elle provienne des photos professionnelles de la société PCGS, du fait de la faible résolution des images téléchargées et l'éclairage non uniforme, nous la notons comme étant de moins bonne qualité que les bases *GENI_coinPhotos* et *20FrancsCOQ_Photos*.

Table 3.1 – Comparaison des bases de données numismatiques

	N.img	**Résolu.(pixels)**	**Seg.**	**Orientation**	**Qualité**
MUSCLE CIS	60000	640×576	Non	Aléatoire	+
Romaines	180	$> 1000 \times 1000$	Oui	Aléatoire	+
USA_Grading	2598	500×500	Non	Quasi-fixe	++
GENI_CoinPhotos	217	$> 2000 \times 2000$	Non	Quasi-fixe	+++
LIRIS_CoinPhotos	250	1400×1400	Non	Quasi-fixe	+
20FrancsCOQ_Coins	286	2200×2200	Oui	Fixe	+++

3.2 Base de données des chiffres extraits des monnaies

Aujourd'hui la reconnaissance de caractères est un domaine bien exploré. Les algorithmes de reconnaissance de caractères sont partout dans la vie quotidienne comme dans les systèmes de lecture des plaques d'immatriculation, les machines de tri de courrier, *etc.* Cependant, les bases de données que nous trouvons dans la littérature sont essentiellement faites pour tester ces types d'algorithmes.

La reconnaissance des caractères extraits des monnaies est utilisée pour aider l'identification du type monétaire, notamment pour les monnaies anciennes. Comme nous l'avons mentionné dans l'état de l'art, les caractères extraits de monnaies sont difficiles à reconnaître par des approches classiques d'OCR. La plupart de ces techniques s'appuient sur une segmentation correcte du premier plan. Cependant dans le cas des monnaies, le premier plan est difficile à extraire car sa couleur ou sa texture est similaire à l'arrière-plan. Pour différencier ce type de police de caractères des autres polices, nous l'appelons police de « type creusé ». Dans ce cas, la séparation entre le premier plan et l'arrière-plan est donnée par les contours en relief. Cependant, ils peuvent être bruités par différents facteurs. Dans la littérature, certains chercheurs ont considéré l'ensemble de l'imagette de caractère comme un motif spécifique caractérisé par un descripteur local (Arandjelović 2012 ; Kavelar, Zambanini et Kampel 2012, 2014).

Dans les monnaies « modernes » considérées dans cette thèse, l'intérêt de la reconnaissance des caractères est la détection du millésime. Bien que certaines légendes soient textuellement identiques parmi les différents types monétaires d'un pays et d'une même période, le millésime est souvent discriminant pour différencier les types monétaires visuellement similaires. Il est ainsi important de reconnaître les chiffres extraits des monnaies.

Parmi les bases de données de caractères connues dans la communauté, par exemple MINST (LeCun et al. 1998), *Chars74K* (Campos, Babu et Varma 2009), les bases de données des compétitions ICDAR (Lucas et al. 2003 ; Shahab, Shafait et Dengel 2011 ; Karatzas et al. 2013 ; Karatzas et al. 2015), *etc.*, nous en avons cherché une pour tester nos algorithmes de reconnaissance des chiffres extraits des monnaies. Cependant, aucune

base de données de caractères existante ne présente les mêmes difficultés. Nous avons donc construit une base de données de chiffres extraits des monnaies, et dans l'objectif d'encourager les chercheurs à proposer des algorithmes de reconnaissance destinés aux caractères en police de « type creusé » ou similaires. Nous avons mis notre base de données de chiffres appelée *CoinNUMS* à disposition de la communauté sur un des serveurs du laboratoire LIRIS [1].

3.2.1 Méthodes d'extraction

Pour mettre en place *CoinNUMS*, nous avons utilisé les bases de photos professionnelles de monnaies *USA_Grading* et *GENI_CoinPhotos* comme point de départ. L'intérêt d'utiliser deux sources différentes est d'avoir une grande variété de polices de caractères et de niveaux de bruit. *USA_Grading* contient un nombre important de monnaies usées, alors que la base de données *GENI_CoinPhotos* n'est composée que des monnaies bien conservées.

CoinNUMS est construite avec tous les chiffres extraits des monnaies de la base *GENI_CoinPhotos*, et les chiffres des millésimes d'une sous-base de *USA_Grading*. Cette présélection contient 300 monnaies ayant des similarités interclasses élevées qui nécessitent la détection du millésime pour réaliser une identification plus précise du type monétaire. Les chiffres se trouvent principalement dans la date parce que la valeur faciale est souvent en lettre comme « *QUARTER DOLLAR* » ou « *TWENTY CENTS* » ; de plus, c'est la date d'émissions qui nous intéresse. Dans les monnaies de la base *GENI_CoinPhotos*, les chiffres se trouvent également dans la valeur faciale, ou d'autres légendes contenant des chiffres.

Deux approches d'extraction de chiffres ont été utilisées. Pour faire une extraction précise et efficace, nous avons développé un outil interactif *CoinLegendExtract* qui permet à l'utilisateur d'extraire des imagettes de caractères à partir des images des monnaies. Grâce au fait que la position du millésime est relativement prédictible dans les monnaies américaines présélectionnées, une extraction automatique de chiffres a également été proposée pour la base *USA_Grading*. Les petites erreurs d'extraction introduites par le processus automatique sont intéressantes pour tester la robustesse des algorithmes développés.

Dans les paragraphes suivants, nous allons détailler ces deux approches d'extraction et les caractéristiques des données de *CoinNUMS*.

Approche manuelle avec l'outil interactif

L'outil interactif *CoinLegendExtract* a été développé dans l'objectif d'extraire les légendes et les caractères des monnaies de manière précise et efficace. L'utilisateur doit tracer une boîte englobante autour d'un caractère sur l'image de la monnaie à l'aide de

1. http ://liris.univ-lyon2.fr/Coins

la souris, puis saisir la vérité terrain aperçue à l'aide du clavier. L'imagette du caractère extrait sera annotée et répertoriée dans la base de données.

De manière générale, cet outil permet d'extraire des imagettes de taille quelconque sur n'importe quelle image (*cf.* Figure 3.12). Par ailleurs, quelques fonctionnalités ont été implémentées pour l'adapter au contexte des images de monnaies. Les fonctionnalités principales de *CoinLegendExtract* sont :

- Lecture d'images : cela permet de lire les images de monnaies les unes après les autres dans un répertoire donné. Une fois les opérations terminées sur l'image courante, l'image suivante sera prise en compte par le programme. Si l'utilisateur quitte le programme avant de terminer toutes les images du répertoire, les images traitées sont mémorisées et le prochain lancement s'initialisera sur les images non-traitées. Il est également possible de lire une image spécifique en donnant son chemin. N'importe quelle taille d'image peut être traitée, l'image affichée s'adapte à l'écran de l'ordinateur.

- Traçage d'une boîte englobante : l'utilisateur doit tracer un rectangle qui englobe le caractère sur l'image affichée. Une fois cette opération terminée, la zone extraite apparaît dans une nouvelle fenêtre zoomée, où l'utilisateur peut vérifier si la boîte englobante est satisfaisante ou pas (*cf.* Figure 3.12). Il est éventuellement possible de retracer une nouvelle boîte englobante pour remplacer l'ancienne avant de sauvegarder.

- Rotation : il est possible de corriger la rotation de la monnaie pour que les caractères ciblés soient positionnés horizontalement (*cf.* Figure 3.13 (b)). Les configurations par défaut permettent à l'utilisateur de faire pivoter la monnaie vers la gauche ou vers la droite d'un petit angle, mais aussi de pouvoir renverser de 180 degrés dans des cas particuliers. Il est déconseillé de lancer l'opération de rotation plusieurs fois car l'interpolation perd des informations de l'image originale. Par contre, le pas de rotation est modifiable de manière à pourvoir réaliser un angle important.

- Changement de coordonnées : la transformation en coordonnées polaires sert à « horizontaliser » les légendes circulaires. Dans la pratique, deux images en sens opposés sont affichées en même temps après cette transformation pour extraire directement les caractères dans le bon sens (*cf.* Figure 3.13 (c)).

- Sauvegarde des imagettes annotées : les imagettes sont sauvegardées dans le répertoire indiqué par l'utilisateur. Par défaut, le nom de l'imagette sauvegardée est $N_M.jpg$, où N est le nom du caractère et M est un indice. N doit être saisi par l'utilisateur avant la sauvegarde, et M est affecté automatiquement selon les fichiers existants. Les imagettes appartenant à la même classe N sont stockées automatiquement dans un sous-répertoire nommé N. L'annotation des imagettes sauvegardées est configurable selon les besoins de l'utilisateur.

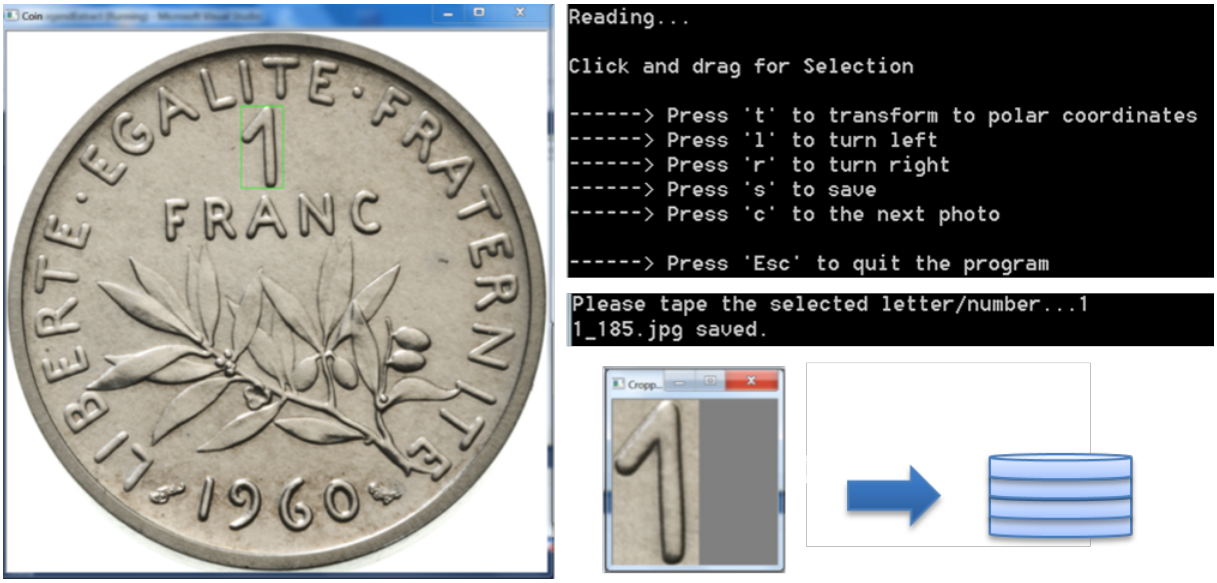


FIGURE 3.12 – Capture d'écran de l'utilisation de l'outil interactif *CoinLegendExtract*

(a) (b) (c)

FIGURE 3.13 – Extraction de légendes horizontales à l'aide de *CoinLegendExtract* : (a) image originale, (b) image corrigée par rotation, (c) image correspondante en coordonnées polaires.

Approche Automatique

En dépit des contraintes de la détection de légendes dans la monnaie, le millésime des monnaies dans *USA_Grading* peut être détectée assez facilement en se basant sur les hypothèses suivantes :

1) Le millésime est toujours composé de quatre chiffres. Au lieu de localiser des caractères individuels, la localisation de l'ensemble de la date est plus robuste et plus efficace.

2) Puisque les monnaies sont bien orientées, la position du millésime est relativement limitée en considérant une base de monnaies à une époque donnée et dans un pays donné. Ainsi, une liste de zones candidates susceptibles de contenir la date $\boldsymbol{Z}$ est utilisée pour éviter une recherche trop longue sur l'image entière.

Sur la base de données *USA_Grading*, la date se situe très souvent en bas de l'avers (Z_1) ou au milieu du revers (Z_2). La date étant parfois circulaire dans l'image originale, la

transformation en coordonnées polaires sert à l'« horizontaliser ». En conséquence, nous considérons cette zone suivant les deux types de coordonnées que nous notons respectivement Z_{1c} et Z_{1p}. La Figure 3.14 illustre les zones candidates $\boldsymbol{Z} = \{Z_{1c}, Z_{1p}, Z_2\}$. Ces zones candidates seraient à adapter en fonction des monnaies étudiées. La surface de ces zones candidates prédéfinies doit toutefois être suffisamment importante pour contenir de manière sûre la date recherchée.

Figure 3.14 – Zones candidates susceptibles de contenir la date dans la sous-base considérée de *USA_Grading*

3) La taille et la forme de la boîte englobante de la date est limitée pour des monnaies considérées. Nous notons $\Delta = \{l_{min}, l_{max}, h_{min}, h_{max}, r_{min}, r_{max}\}$ l'ensemble des extrêma de la largeur, la hauteur et le ratio d'aspect de la boîte englobante recherchée.

Pour détecter la position précise de la date, nous la cherchons dans toutes les zones candidates $\boldsymbol{Z}$ de manière exhaustive selon une fenêtre glissante de taille modifiable. Soit $W(\boldsymbol{Z}, \Delta)$ l'ensemble de fenêtres glissantes limitées par $\boldsymbol{Z}$ et Δ. L'idée est de chercher la fenêtre w_d qui englobe le mieux la date. Pour cela, pour chaque fenêtre $w_i \in W$ nous extrayons un histogramme normalisé $H(w_i)$ obtenu par scan vertical puis horizontal sur sa carte des gradients. $H(w_i)$ est ensuite comparé avec l'histogramme moyen $\widehat{H_M}$ obtenu à l'aide de modèles synthétiques (*cf.* Figure 3.15). $\widehat{H_M}$ est composé d'une large bosse et de quatre bosses étroites. Malgré les bruits, l'histogramme normalisé de la fenêtre qui englobe le mieux la date serait le plus corrélé à $\widehat{H_M}$, donnée par,

$$w_d = argmax\ Cor(H(w_i), \widehat{H_M}), w \in W(\boldsymbol{Z}, \Delta) \tag{3.1}$$

Les images présélectionnées pour construire la base de données de chiffres ont toutes des zones avec des dates. Dans le cas où la face de la monnaie ne contiendrait pas de date, la corrélation maximale serait faible (expérimentalement, $Cor_{max}(H(w_i), \widehat{H_M})$<0,65), sauf dans le cas où il y aurait d'autres motifs ou des textes alignés comme une date dans les

zones candidates. Les fausses détections sont causées principalement soit par une date trop usée, soit par des motifs ou des textes ressemblant à une date.

La Figure 3.16 illustre des exemples de zones dans lesquelles la date est détectée. Nous pouvons voir que la méthode proposée est assez robuste même si la distinction entre les chiffres et le fond n'est pas évidente à l'œil. La première ligne de la Figure 3.16 correspond aux exemples parfaits, et la deuxième ligne correspond aux erreurs typiques :

- Extraction non-horizontale (*cf.* Figure 3.16 (e))
- Extraction partielle avec une partie non majoritaire manquante (*cf.* Figure 3.16 (f))
- Extraction partielle avec une partie importante manquante (*cf.* Figure 3.16 (g))
- Extraction partielle avec un chiffre entier manquant (*cf.* Figure 3.16 (h))

À part les deux derniers cas qui font sans aucun doute échouer toute reconnaissance, les autres petites erreurs sont considérées comme acceptables et c'est la raison pour laquelle les boîtes englobantes sont vertes à la Figure 3.16. Le taux de détection avec une bonne boîte englobante (comme encadrée en vert à la Figure 3.16) est de 80% sur un test de 375 images.

À partir de la date extraite, il reste à le découper en quatre chiffres individuels. Le processus de découpage est réalisé automatiquement via une recherche de minima locaux de l'histogramme $H(w_d)$ les plus proches des séparations supposées : $0,25l_d, 0,5l_d, 0,75l_d$, où l_d est la largeur de w_d. Le pipeline de l'extraction automatique de chiffres à partir de la monnaie segmentée est illustré à la Figure 3.15.

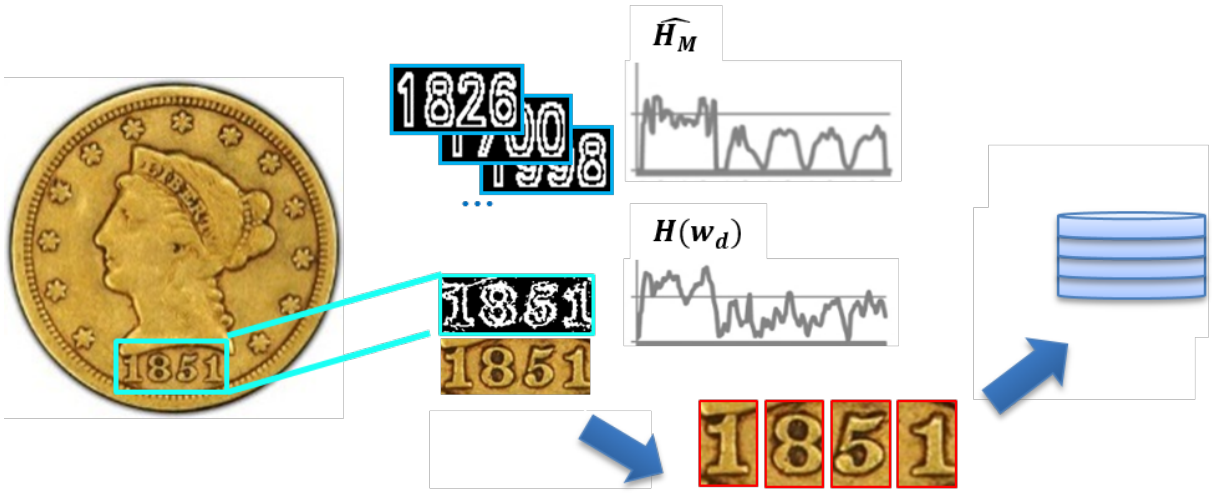


FIGURE 3.15 – Pipeline de l'extraction automatique des chiffres de la monnaie

(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 3.16 – Exemples de dates extraites : les bonnes ou quasi-bonnes encadrées en vert (a-f) ; les mauvaises encadrées en rouge (g-h)

3.2.2 Différentes sous-bases de CoinNUMS

CoinNUMS_geni

La sous-base *CoinNUMS_geni* consiste en 606 images de chiffres provenant de différentes légendes (millésime, valeur faciale, *etc.*) des monnaies de la base *GENI_CoinPhotos*. De ce fait, la taille des images de chiffres est très variable. L'extraction a été réalisée manuellement à l'aide de l'outil *CoinLegendExtract*. Dans ce cas-là, les erreurs de découpage sont minimisées.

Comme la base de données source ne contient que des images avec une haute résolution et les monnaies sont bien conservées, les caractères frappés sont clairement lisibles à l'œil nu. La vérité terrain est donnée par le chiffre extrait. Puisque les monnaies ne sont pas étiquetées dans la base de données *GENI_CoinPhotos*, l'annotation de l'image du chiffre est uniquement donnée par (Classe)_(Indice), où (Classe) est la vérité terrain du chiffre et (Indice) signifie que l'image actuelle est la $Indice^{ème}$ image sauvegardée pour cette classe dans la sous-base *CoinNUMS_geni*. Par exemple, 0_2.jpg signifie la deuxième image de « 0 » rajouté dans cette sous-base. Des exemples de la sous-base *CoinNUMS_geni* sont donnés à la Figure 3.17 (a).

CoinNUMS_pcgs_a

La sous-base *CoinNUMS_pcgs_a* consiste en 1200 images de chiffres provenant de la date d'émission de 300 images de *USA_Grading*. Pour ces 300 images de monnaies, la date a été correctement détectée (parfois avec de petites erreurs d'extraction) par l'approche automatique que nous avons proposée, et puis découpée en quatre chiffres.

Dans cette base de données, la vérité terrain a été manuellement vérifiée, en utilisant également les informations *a priori* sur le site de PCGS. Dans le cas de chiffres extrêmement dégradés ou coupés, nous avons regardé la monnaie entière et avons cherché

son millésime sur le site de PCGS pour que l'annotation soit correcte. Chaque image de chiffre est étiquetée par (Classes)_(Position dans date)_(Nom de monnaie). Par exemple, 1_1_GoldDollar2-55r.jpg signifie qu'il s'agit d'une image d'un chiffre « 1 », premier chiffre de la date, extraite de la monnaie *GoldDollar2-55r*. Des exemples de la sous-base *CoinNUMS_pcgs_a* sont illustrés à la Figure 3.17 (b).

CoinNUMS_pcgs_m

La sous-base *CoinNUMS_pcgs_m* consiste en 1200 images de chiffres provenant de la même source que *CoinNUMS_pcgs_a*. Les images de chiffres sont extraites manuellement à l'aide de l'outil *CoinLegendExtract*. L'annotation des données est exactement la même que celle de *CoinNUMS_pcgs_a*.

Le but de proposer *CoinNUMS_pcgs_a* et *CoinNUMS_pcgs_m*, est d'avoir deux sous-bases utilisant des données identiques mais obtenues par des approches d'extraction différentes. *CoinNUMS_pcgs_a* contient évidemment plus de bruit que *CoinNUMS_pcgs_m* dans la mesure où le processus d'extraction a été automatique. Des exemples de la sous-base *CoinNUMS_pcgs_m* sont illustrés dans la Figure 3.17 (c).

FIGURE 3.17 – Exemples d'images de la base de données *CoinNUMS* : (a) *CoinNUMS_geni* ; (b) *CoinNUMS_pcgs_m* ; (c) *CoinNUMS_pcgs_a*

3.2.3 Comparaison des propriétés visuelles des sous-bases de CoinNUMS

Ici, nous souhaitons comparer les propriétés visuelles des différentes sous-bases de la base *CoinNUMS* selon les critères suivants :

- Police de « type creusé » : une des caractéristiques remarquables que nous voulons mettre en avant dans la base de données *CoinNUMS* est le fait que le premier plan ne se discerne de l'arrière-plan que par des contours en relief. Des données idéales en police de « type creusé » sont illustrées par les images synthétiques à la Figure 3.18 (a). Dans le cas réel, les contours des caractères sont souvent soumis à des bruits. Notons toutefois que certains chiffres dans *CoinNUMS* possèdent quand même un premier plan distinct de son arrière-plan. Cela peut être causé par la patine, comme dans le cas des monnaies belles épreuves (*cf.* Chapitre 1) dont les légendes sont mattes sur un fond poli (*cf.* Figure 3.18 (b)).

- Propreté : nous utilisons le terme « propreté » pour décrire si les contours externes du caractère sont clairement présentés. Cette propriété est en fonction de la qualité de la photographie et de l'état de conservation de la monnaie. Normalement, les caractères extraits des photos professionnelles, avec une haute résolution, sont plus « propres » que ceux extraits des photos amateurs avec une faible résolution. Par ailleurs, la propreté du caractère extrait est évidemment moins bonne pour des monnaies usées. Par exemple, les images de chiffres provenant des photos professionnelles de la société GENI sont plus « propres » que celles provenant de la base *USA_Grading*, parce que les monnaies de la société GENI sont relativement bien conservées et les photos de meilleure qualité.

- Inclinaison : certains caractères sont intrinsèquement inclinés en fonction de la police. Mais le plus souvent, l'inclinaison est due à l'extraction automatique de la date dans des coordonnées non appropriées, par exemple, une date horizontale extraite en coordonnées polaires ou une date circulaire extraite en coordonnées cartésiennes. Une telle inclinaison est évidemment limitée à un petit degré dans le cas d'une extraction manuelle (*cf.* Figure 3.18 (c)).

Intégralité des caractères : il s'agit de caractériser le fait que le chiffre extrait est intégralement contenu dans l'image ou non. Dans le cas d'une extraction manuelle, le chiffre est toujours complet. Dans le cas d'une extraction automatique, une partie du caractère peut être découpée (*cf.* Figure 3.18 (d)).

- Bruit externe : il s'agit d'éléments qui n'appartiennent pas au caractère mais qui sont inclus dans l'image. Ces bruits peuvent provenir d'autres motifs en relief de la monnaie, ou être des « éléments inattendus » comme des rayures et des taches. Parfois ces éléments sont touchés ou superposés au caractère (*cf.* Figure 3.18 (e)).

- Bruit interne : il s'agit de motifs décoratifs ou de texture de la surface à l'intérieur du caractère. Ce type de bruit ne modifie pas la forme globale du caractère définie par les contours externes en relief mais peut poser problème pour la reconnaissance (*cf.* Figure 3.18 (f)).

La Table 3.2 récapitule la comparaison de toutes les sous-bases de la base *CoinNUMS*. En résumé, les données dans la sous-base *CoinNUMS_geni* sont proches des caractères idéaux en police de « type creusé » grâce à l'extraction manuelle et à la bonne qualité des monnaies sources. Il faut mentionner que la distribution des données n'est pas égale. Certaines classes possèdent très peu de données par rapport à d'autres (*cf.* Table 3.3). En particulier, le chiffre « 1 » est très représenté dans de nombreuses dates.

(a) (b) (c) (d) (e) (f)

FIGURE 3.18 – Exemples d'imagettes de chiffres : (a) images synthétiques « idéales » des chiffres en police de « type creusé » et exemples réels : (b) chiffres avec un premier plan distinct ; (c) chiffres inclinés ; (d) chiffres légèrement découpés ; (e) chiffres avec du bruit externe ; (f) chiffre avec du bruit interne

TABLE 3.2 – Propriétés visuelles des différentes sous-bases de CoinNUMS

Sous-base	**Police**	**Prop.**	**Intég.**	**Incl.**	**Bruit ext.**	**Bruit int.**
« Type creusé » idéal	Oui	Oui	Oui	Non	Non	Non
CoinNUMS_geni	Oui*	Oui	Oui	Oui*	Oui*	Oui*
CoinNUMS_pcgs_a	Oui*	Non*	Non*	Non	Oui*	Oui*
CoinNUMS_pcgs_m	Oui*	Non*	Oui	Oui*	Oui*	Oui*

Le symbole * signifie qu'une petite partie des données ne correspond pas au critère.

La base de données *CoinNUMS* actuelle ne contient pas une quantité importante de données par rapport à d'autres bases de données pour la reconnaissance de caractères. Cette quantité n'est pas du tout suffisante pour faire de l'apprentissage profond, comme dans le travail de LeCun et al. (1998). D'autre part, certaines classes de chiffres possèdent très peu de données pour faire un apprentissage pertinent. La taille de la base *CoinNUMS* sera augmentée petit à petit au fur et à mesure des sessions de gradation de la société GENI. Autrement, nous envisageons d'étendre *CoinNUMS* à des lettres, des caractères étrangers et des légendes.

Table 3.3 – Distribution de données dans les sous-bases de *CoinNUMS*

Sous-bases	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Total**
CoinNUMS_geni	46	185	43	4	2	42	43	14	84	143	606
CoinNUMS_pcgs_a	80	331	57	81	24	66	70	149	213	139	1200
CoinNUMS_pcgs_m	80	331	57	81	24	66	70	149	213	139	1200

Cette base de données est disponible sous *Creative Commons License* sur le site associé : liris.univ-lyon2.fr/Coins (*cf.* Figure 3.19). Nous espérons que cette base de données attirera l'attention de chercheurs spécialisés en reconnaissance de caractères et les incitera à proposer des algorithmes performants. Les chercheurs intéressés peuvent télécharger gratuitement cette base de données sur une simple inscription sur le site.

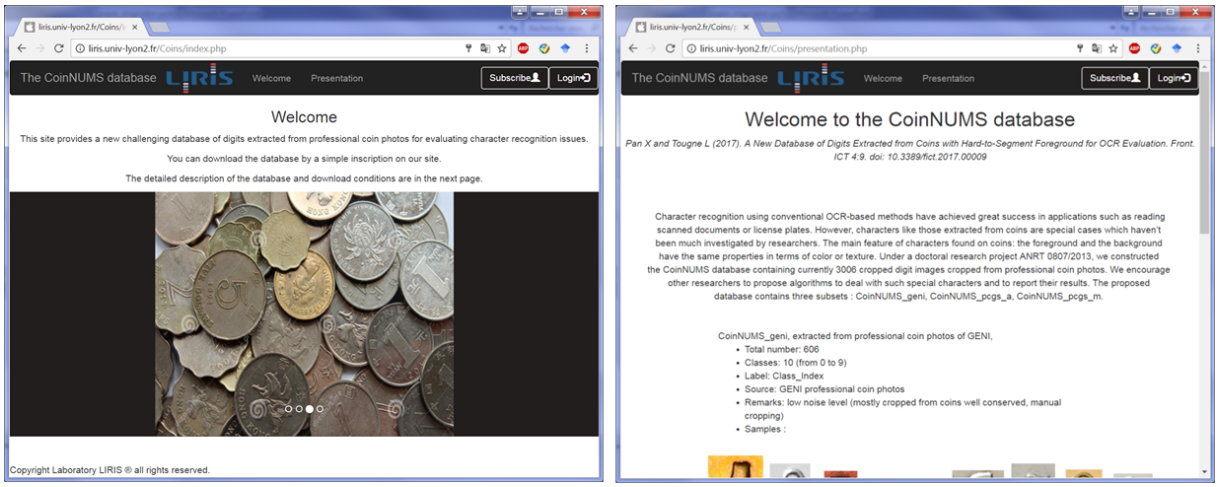


Figure 3.19 – Capture d'écran du site où *CoinNUMS* est disponible

3.3 Conclusion

Dans ce chapitre, nous avons d'abord présenté les bases de données pour l'identification du type monétaire utilisées dans la littérature : *MUSCLE CIS* pour les monnaies « modernes », et les monnaies de la République romaine pour les monnaies anciennes. Cependant, notre travail se focalise sur les monnaies « modernes » de collection. Malgré un grand nombre de données, *MUSCLE CIS* ne comporte que des images d'intensité de faible résolution qui ne correspondent pas à notre contexte d'étude.

Pour valider les différentes phases concernées dans le cadre de cette thèse, plusieurs bases de données ont été construites. La base de données *LIRIS_CoinPhotos* contenant des images semi-synthétiques de formes variées sera utilisée pour valider la segmentation précise des monnaies. La base *USA_Grading* qui contient des monnaies de différents états de conservation par type monétaire sera utilisée pour valider l'identification du type monétaire. Comme les images dans la base *USA_Grading* sont de faible résolution, elles ne

pourront être utilisées pour une gradation fine. La base de données *20FrancsCOQ_Photos* composée de photos professionnelles de bonne résolution sera utilisée pour la gradation. Pour évaluer les algorithmes de reconnaissance de caractères extraits de monnaies, à partir des photos professionnelles de monnaies, nous avons proposé une base de données de chiffres nommée *CoinNUMS*. Afin d'avoir une grande variabilité de polices et de niveaux de bruit, cette base de données est construite à partir de deux sources différentes (*GENI_coinPhotos* et *USA_Grading*), et par deux approches différentes d'extraction. L'extraction des chiffres dans la monnaie est soit réalisée manuellement à l'aide de l'outil interactif *CoinLegendExtract* que nous avons développé, soit automatiquement à l'aide de l'algorithme de détection et de décomposition de la date d'émission que nous avons proposé (Pan et Tougne 2016).

Chapitre 4
Segmentation des monnaies

DANS ce chapitre, nous détaillons notre méthode de segmentation des monnaies dédiée aux photos professionnelles. Cette méthode s'appuie sur une modélisation des différentes formes régulières de monnaies « modernes ». Cette segmentation réalisée à l'aide d'un modèle paramétrique a deux objectifs dans le contexte de cette thèse : premièrement, les monnaies parfaitement segmentées seront utilisées comme des illustrations dans les catalogues de l'entreprise. Deuxièmement, les monnaies segmentées seront des données d'entrée pour l'identification du type monétaire et la gradation.

La méthodologie proposée est détaillée dans la section 4.1. Dans la section 4.2, les expériences réalisées en utilisant notre algorithme et les résultats sont présentés. Enfin, la section 4.3 récapitule les avantages et les contraintes de cette méthode.

4.1 Méthodologie proposée

De manière générale, les chercheurs qui développent des algorithmes de segmentation de monnaies dans la littérature se satisfont d'une segmentation « grossière » que la forme de la monnaie soit ronde ou irrégulière. Cependant, dans le contexte de cette thèse, une demande de la société GENI était d'obtenir un détourage très précis pour des formes particulières. D'après les numismates de la société GENI, les monnaies considérées ont trois types de formes fréquentes qu'ils souhaitent traiter automatiquement : ronde, polygonale (à pans) et sinusoïdale (à dents). Par ailleurs, certaines monnaies peuvent posséder un trou au milieu.

La méthode proposée se base sur une modélisation des formes de monnaies concernées. Ainsi, la segmentation consiste à chercher un modèle qui correspond bien à la forme effective de la monnaie à segmenter.

- Modèle polygonal : la forme d'une monnaie polygonale est caractérisée par son centre O, le rayon R, le nombre de coins N et la rotation θ (l'angle entre le coin le plus proche de l'axe X et l'axe des abscisses lui-même). Un modèle polygonal est ainsi défini par $\Omega_{poly}(O, R, \theta, N)$ (*cf.* Figure 4.1 (a)).

- Modèle sinusoïdal : La forme d'une monnaie polygonale est caractérisée par son centre O, le rayon R, l'amplitude du sinus A et le nombre de crêtes N et la rotation θ (l'angle entre le segment long de taille $(R + A)$ le plus proche de l'axe X et l'axe des abscisses lui-même). Un modèle sinusoïdal est ainsi défini par $\Omega_{sin}(O, R, A, \theta, N)$ (*cf.* Figure 4.1 (b)).

Ces modèles représentent également les monnaies rondes parce qu'elles peuvent être considérées comme les monnaies polygonales ayant un nombre de coins infini ($N = \infty$) (*cf.* Figure 4.1 (c)) où les monnaies sinusoïdales ayant une amplitude nulle ($A = 0$).

- Modèle général : la plupart des paramètres dans le modèle polygonal et dans le modèle sinusoïdal sont similaires, seule l'amplitude A est unique pour le modèle sinusoïdal. Pour simplifier, quelle que soit la forme de la monnaie, nous la modélisons donc par un modèle général $\Omega(O, R, A, \theta, N, S)$ où $S = \{0, 1\}$ est une valeur binaire qui précise si le modèle est polygonal ou sinusoïdal.

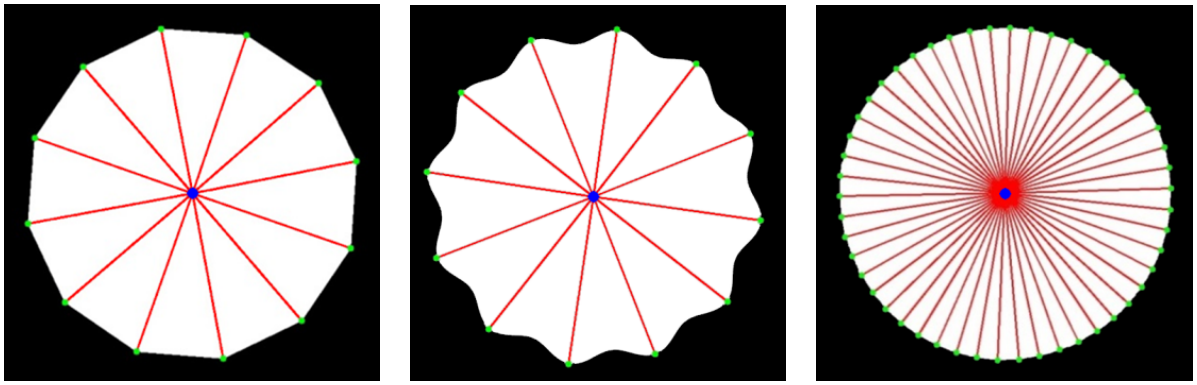
Le processus de segmentation proposé est composé de trois étapes :

- Segmentation « grossière » : il s'agit d'une extraction approximative de la région binaire de la monnaie à l'aide d'algorithmes basiques de traitement d'images sur l'image brute.

- Estimation du modèle : il s'agit d'estimer les paramètres qui constituent le modèle sur la région « grossièrement » segmentée. Les paramètres estimés sont considérés comme appartenant au voisinage des valeurs précises.

- Évolution du modèle : dans la dernière étape, les paramètres du modèle seront optimisés par une minimisation d'énergie. La segmentation finale est obtenue par le modèle optimal.

La Figure 4.2 illustre le pipeline composé de ces trois étapes que nous détaillerons dans les paragraphes suivants.



(a) (b) (c)

Figure 4.1 – Modèles de formes différentes : (a) forme polygonale ; (b) forme sinusoïdale ; (c) forme ronde

4.1.1 Segmentation « grossière »

L'objectif de cette étape est d'obtenir une région qui représente approximativement la forme de la monnaie. Dans cette étape, les petits bruits sont tolérés (*cf.* Figure 4.3 (c)) et, seront traités dans les étapes suivantes pour obtenir une segmentation précise (*cf.* Figure 4.3 (b)). Pour éviter des défauts trop importants qui pourraient modifier

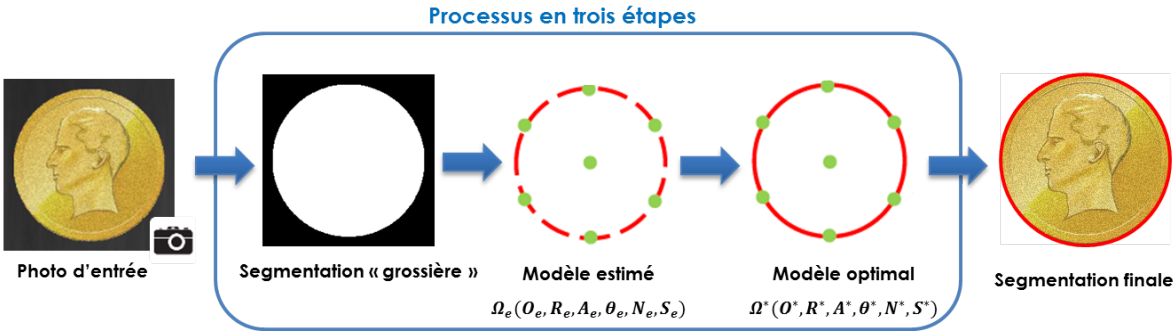
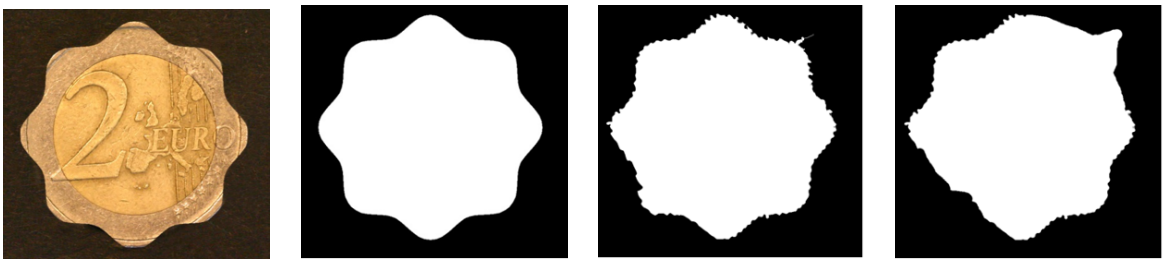


FIGURE 4.2 – Processus de la méthode de segmentation composée de trois étapes principales : segmentation « grossière », estimation du modèle et calcul du modèle optimal

.

complètement la segmentation finale (*cf.* Figure 4.3 (d)), nous combinons des techniques de base de traitement d'images à savoir, le seuillage et la détection de contours.



(a) (b) (c) (d)

FIGURE 4.3 – Qualité de la segmentation « grossière » : (a) monnaie originale (synthétique) de forme sinusoïdale ; (b) segmentation exacte ; (c) segmentation approximative ; (d) segmentation trop bruitée

- Seuillage basé sur la méthode d'Otsu

Dans les photos professionnelles, l'arrière-plan quasi-uniforme est souvent blanc ou noir. Les pixels appartenant à la monnaie sont ainsi supposés plus sombres ou plus lumineux que ceux appartenant au fond dans l'image d'intensité. Dans ce cas-là, un seul seuil suffit à séparer la région de la monnaie du reste de l'image. Nous utilisons la méthode d'Ostu (Otsu 1975) pour trouver ce seuil optimal. Pour l'image entrée I, l'image seuillée est notée I^T (*cf.* Figure 4.4 (b)). Même si l'image seuillée I^T comporte parfois des pixels mal affectés, normalement ces bruits sont considérés faibles et donc faciles à supprimer.

Cependant, à cause du contraste varié entre l'arrière-plan et la région de la monnaie, le seuil obtenu par la méthode d'Ostu n'est pas toujours « optimal ». Dans la Figure 4.4 (b), par exemple, nous pouvons voir que les parties foncées de la monnaie sont mal affectées à l'arrière-plan. C'est l'un des inconvénients de la segmentation par seuillage global que nous avons mentionné dans l'état de l'art. Par ailleurs, pour certaines monnaies multicolores, il est possible que l'intensité de l'arrière-plan soit inférieure à une partie de la monnaie mais supérieure à l'autre partie. Autrement dit, il n'existe pas un seul seuil optimal qui

permet d'extraire la région complète de la monnaie. Même si de tels cas sont très rares, il est important de développer une méthode robuste.

- Détection de contours par filtre de Canny

La forme de la monnaie est également définie par ses contours extérieurs indépendant du contraste entre les régions.

Sachant que nous avons des photos de très bonne qualité, la détection des contours extérieurs de la monnaie devrait être évidente. Après avoir testé plusieurs détecteurs de la littérature, nous avons décidé d'utiliser le filtre de Canny qui permet d'obtenir une carte des contours binaire de largeur d'un pixel. L'idée est de détecter les contours aux endroits où la méthode d'Otsu a mal binarisé, les faux négatifs étant plus graves que les faux positifs. Les deux seuils de Canny permettent de réaliser une détection sensible correspondant aux besoins. La carte des contours binaire est notée I^C (*cf.* Figure 4.4 (c)).

- Région approximative de la monnaie

La carte des probabilités de la monnaie I^P est obtenue par la superposition de I^T et I^C, traitée ensuite par l'opération morphologique fermeture (*closing*) (*cf.* Figure 4.4 (d)) :

$$I^P = (I^T + I^C) \bullet S_{elem}, \tag{4.1}$$

où S_{elem} est l'élément structurant et $\bullet$ est l'opérateur de fermeture.

En général, I^P et I^T sont proches. Cependant, dans le cas où I^T ne présente pas la totalité de la monnaie, I^C permet de « compléter » le bord extérieur. La région approximative de la monnaie correspond à la plus grosse composante connexe dans I^P. Les autres composantes connexes sont considérées comme les bruits de l'arrière-plan et nous les supprimons.

Grâce à une opération morphologique (*closing*) et un remplissage de l'enveloppe convexe de la plus grosse composante connexe dans I^P, nous obtenons la segmentation « grossière » I^S (*cf.* Figure 4.4 (e)). Il faut noter que I^S ne constitue qu'une segmentation « grossière » et en aucun cas une segmentation parfaite. En effet, nous pouvons noter des petits défauts sur l'image zoomée : le bord n'est pas parfaitement lisse (*cf.* Figure 4.5).

4.1.2 Estimation des paramètres du modèle

Cette étape a pour objectif d'estimer rapidement les paramètres initiaux du modèle géométrique Ω_0. L'estimation se fait sur l'image « grossièrement » segmentée I^S obtenue à l'étape précédente.

(a) (b) (c) (d) (e)

FIGURE 4.4 – Différentes étapes de la segmentation « grossière » : (a) images originales ; (b) images seuillées I^T ; (c) cartes des contours I^C ; (d) cartes des probabilités I^P ; (e) résultats de la segmentation « grossière » I^S

.

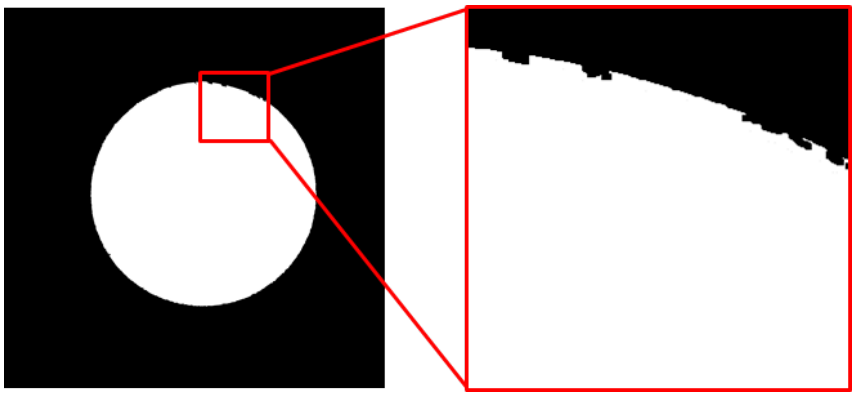


FIGURE 4.5 – Bord bruité sur I^S zoomée

.

- Estimation du centre $O_e(x_e, y_e)$

Les moments géométriques sont utilisés en image pour caractériser des formes. Ils correspondent à des sommes pondérées calculées comme suit :

$$m_{ij} = \sum_{(x;y)\in I^S} x^i y^j \tag{4.2}$$

où i et j définissent l'ordre des moments.

Dans le cas d'une image segmentée binaire, seuls les pixels de la région segmentée sont pris en compte. En particulier, le centre de la région est donné par

$$O_e(x_e, y_e) = (\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}}) \tag{4.3}$$

À partir du centre estimé, nous estimons les autres paramètres par des relations géo-

métriques.

- Estimation du rayon R_e et de l'amplitude A_e

En supposant que le modèle estimé est sinusoïdal, le rayon R_e et l'amplitude A_e sont estimés en même temps. Pour cela, nous cherchons d'abord le plus long segment d_{max} et le plus court segment d_{min} entre le centre estimé O_e et le bord de la région « grossièrement » segmentée I^S. Les points intersectés par ces segments respectifs sont notés P_{max} et P_{min} (*cf.* Figure 4.6).

Dans le cas d'une forme sinusoïdale, d_{max} est approximativement la longueur d'un segment entre le centre O_e et une crête et, d_{min} est approximativement la distance entre le centre O_e et une vallée. Dans le cas d'une forme polygonale, d_{max} est approximativement la longueur d'un segment entre le centre O_e et un sommet du polygone et, d_{min} la distance entre le centre O_e et le milieu d'un côté du polygone. Dans le cas d'une forme ronde, d_{max} et d_{min} devraient être égaux, mais dans la pratique, ils sont proches. En effet, d_{max} et d_{min} seuls sont sensibles aux bruits du bord. Cependant, dans le cadre de notre modèle les segments les plus longs (respectivement les plus courts) sont régulièrement repartis contrairement aux bruits. d_{max} (respectivement d_{min}) est donc donné par la médiane des n segments les plus longs (respectivement les plus courts).

En supposant que le modèle estimé soit sinusoïdal, le rayon R_e et l'amplitude A_e est données par

$$R_e = \frac{1}{2}(d_{max} + d_{min}) \tag{4.4}$$

$$A_e = \frac{1}{2}(d_{max} - d_{min}) \tag{4.5}$$

- Estimation du nombre de côtés ou de crêtes N_e

Le nombre possible de côtés ou de crêtes est limité. D'après les numismates de la société GENI, $\boldsymbol{N}_{poly} = \{8, 10, 11, 12, 18, 20, 22, 24\}$ est l'ensemble des nombres de coins possible pour les monnaies polygonales ; $\boldsymbol{N}_{sin} = \{12\}$ est le seul nombre de crêtes possible pour les monnaies sinusoïdales françaises ; mais pour des monnaies étrangères, ces nombres pourraient être différents et l'algorithme serait à adapter. Quelle que soit la forme attendue, nous utilisons $\boldsymbol{N}_{all} = \{8, 10, 11, 12, 18, 20, 22, 24, 50\}$ comme l'ensemble de tous les N possibles, où $N = 50$ correspond à la discrétisation d'une forme ronde.

Nous parcourons donc tous les $N \in \boldsymbol{N_{all}}$. Pour chaque N parcouru, nous simulons la répartition des N segments à partir du segment le plus long détecté. Le rapport des longueurs de ces différents segments simulés est d'autant plus proche de 1 que N_e correspond au nombre effectif de coins ou de crêtes de la région segmentée. L'Algorithme 1 donne le pseudocode permettant d'estimer N_e. La Figure 4.7 montre deux exemples de monnaies synthétiques comportant 12 côtés et 12 crêtes. Selon l'Algorithme 1, $r(N)$ est proche de

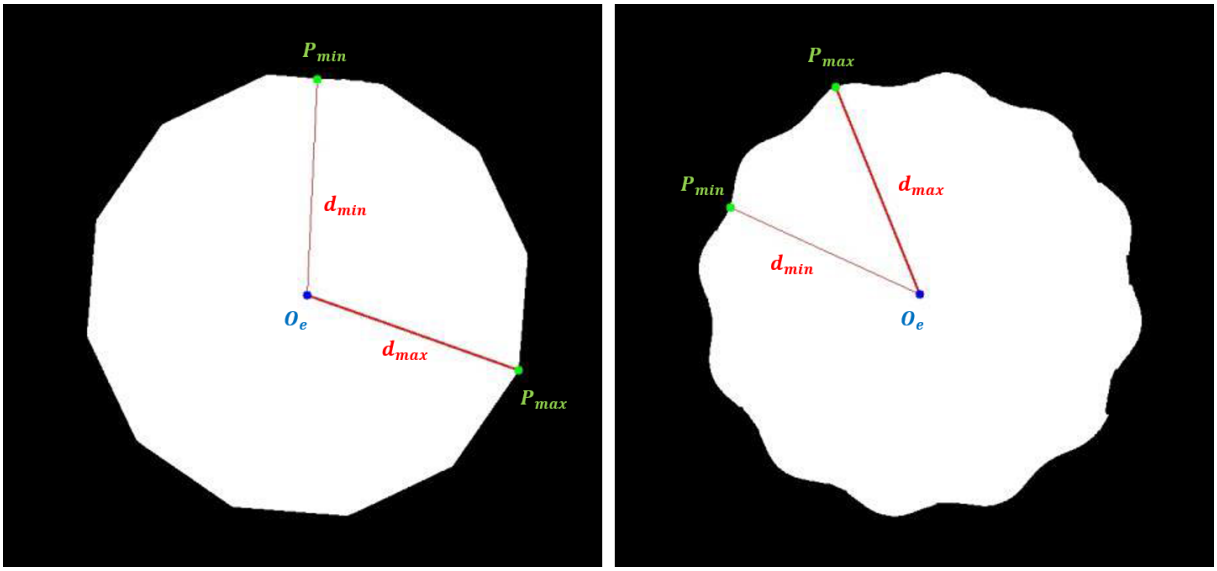


FIGURE 4.6 – Recherche du segment le plus long et du segment plus court sur une région polygonale et une région sinusoïdale

.

1 pour $N = 12$ et bien différent pour les autres valeurs.

La précision de l'estimation de N_e est importante car elle limitera la forme du modèle à l'étape suivante afin de guider la segmentation. Par contre, ce n'est pas catastrophique si d'autres paramètres comme le centre, le rayon, la rotation ou l'amplitude ne sont pas très bien estimés, car le modèle géométrique actif permettra d'obtenir les paramètres les plus adaptés. Mais plus ces paramètres seront bien estimés, moins de temps de calcul sera nécessaire à l'étape suivante, et cela évitera que le modèle ne tombe dans des minima locaux.

- Estimation de l'angle θ_e

L'angle de rotation estimé θ_e est obtenu en considérant l'angle entre le segment le plus proche de l'axe X et l'axe des abscisses lui-même (*cf.* Figure 4.8). Dans le cas où $N_e = 50$, nous n'avons pas besoin d'estimer la rotation car il s'agit d'une forme ronde.

- Estimation de la forme S_e

Trois formes sont considérées : ronde, polygonale et sinusoïdale. Notons que la forme ronde est un cas particulier des deux autres : soit une forme polygonale avec $n = \infty$, soit une forme sinusoïdale avec $A = 0$. Pour sélectionner un modèle correct pour la suite, il s'agit de distinguer si la forme approximative de la région segmentée I^S est polygonale ou sinusoïdale. Pour cela, nous pourrions utiliser le fait qu'une forme polygonale est strictement convexe alors qu'une forme sinusoïdale comporte des parties concaves. Mais en réalité, les bruits sur le bord de la forme rendent cette distinction difficile à évaluer.

Algorithm 1 Pseudocode de l'estimation de N_e

Entrée : région « grossièrement » segmentée I^S, centre estimé O_e, segment le plus long d_{max} et son point intersecté P_{max}
Sortie : nombre de coins ou de crêtes estimé N_e

$P_0 \leftarrow findIntersect(I^S, O_e, 0)$
$\phi_0 \leftarrow arccos \dfrac{\|O_e - P_{max}\|_2}{\|O_e - P_0\|_2}$
for all $N \in \boldsymbol{N}_{all}$ **do**
    $i \leftarrow 0$
    $\phi \leftarrow 0$
    $d\phi \leftarrow \dfrac{2\pi}{N}$
    while $i < N$ **do**
        $i \leftarrow i + 1$
        $P_i \leftarrow findIntersect(I^S, O_e, \phi_0 + \phi)$
        $\phi \leftarrow \phi + d\phi$
    end while
    $r(N) \leftarrow \sum_{i=1}^{N} \dfrac{\|O_e - P_i\|_2}{N \times d_{max}}$
end for
$N_e \leftarrow \underset{N \in \boldsymbol{N}_{all}}{argmin}(|r(N) - 1|)$

* La fonction $findIntersect(I^S, P, \phi)$ retourne le point P' appartenant au bord de la région I^S tel que l'angle entre $\overrightarrow{PP'}$ et l'axe des abscisses égal à ϕ.

Dans la suite, $S_e = 0$ signifie que la région estimée est polygonale convexe et $S_e = 1$ la forme est sinusoïdale.

L'approche adoptée est alors la suivante : nous simulons la répartition des segments les plus courts à l'aide de N_e, de la même manière que nous avons procédé pour les segments les plus longs pour estimer N_e. Nous avons donc obtenu N_e points intersectés par les segments longs, notés $P_i \in \boldsymbol{P}$, et N_e points intersectés par les segments cours, notés $Q_i \in \boldsymbol{Q}$, $i \in [1, N_e]$.

Notons un polygone $\Psi(V)$ par l'ensemble des sommets V, nous constituons deux formes polygonales : $\Psi_1(\boldsymbol{P})$ et $\Psi_2(\boldsymbol{P} \cap \boldsymbol{Q})$. N'importe quelle soit la forme effective, Ψ_1 est toujours convexe. Par contre, dans le cas d'une forme polygonale, Ψ_2 devrait être convexe et proche de Ψ_1, mais dans le cas d'une forme sinusoïdale Ψ_2 devrait être concave. Malgré les petits bruits du bord, la différence des aires entre Ψ_1 et Ψ_2 devrait être discriminante pour estimer la forme (*cf.* Figure 4.9). Nous utilisons

$$r_S = \alpha \frac{|Area(\Psi_1) - Area(\Psi_2)|}{d_{max}^2} \tag{4.6}$$

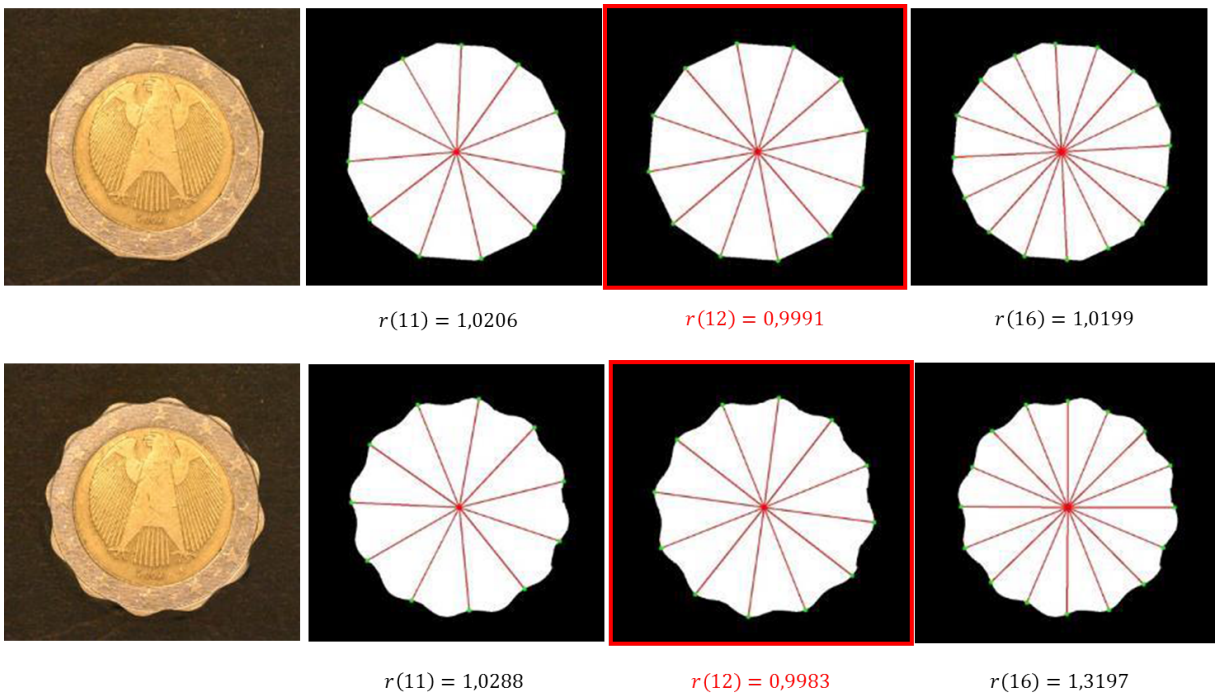


FIGURE 4.7 – Estimation de N_e à l'aide du rapport $r(N)$ des segments simulés

.

pour mesurer cette différence, où α est une constante qui permet à r_S de ne pas être trop petit. Le fait de diviser par d_{max}^2 permet d'être indépendant de la taille de la monnaie. r_S est d'autant plus proche de 0 que la forme polygonale, et beaucoup plus élevé si la forme est sinusoïdale. Nous obtenons finalement :

$$S_e = \begin{cases} 1; & \text{si } r_S > r_{th} \\ 0; & \text{sinon} \end{cases} \tag{4.7}$$

où r_{th} est une constante empirique pour distinguer les formes. Nous avons conduit un test sur 35 paires de photos dont chacune contient un exemplaire polygonal et un exemplaire sinusoïdal fabriqués à partir de la même photo. Le nombre de côtés et le nombre de crêtes sont identiques dans chaque paire d'images. Dans la Figure 4.10 qui présente les résultats, l'axe des abscisses correspond au numéro de test et l'axe des ordonnées à r_S. Nous observons que, pour la forme sinusoïdale, r_S dépend principalement du rapport amplitude-rayon A/R ; mais pour la forme polygonale, r_S ne dépend que des bruits du bord. Selon ce graphique, la distinction entre les deux formes est réalisée par r_S lorsque d_{max} et d_{min} sont bien estimés.

Enfin, nous obtenons un modèle $\Omega_e(O_e, R_e, A_e, \theta_e, N_e, S_e)$ définit par les paramètres estimés. Il s'agit en fait d'un modèle initial $\Omega^{(0)}$ qui bougera vers les paramètres optimaux dans l'étape suivante.

Il faut cependant mentionner que l'estimation des paramètres initiaux est sensible

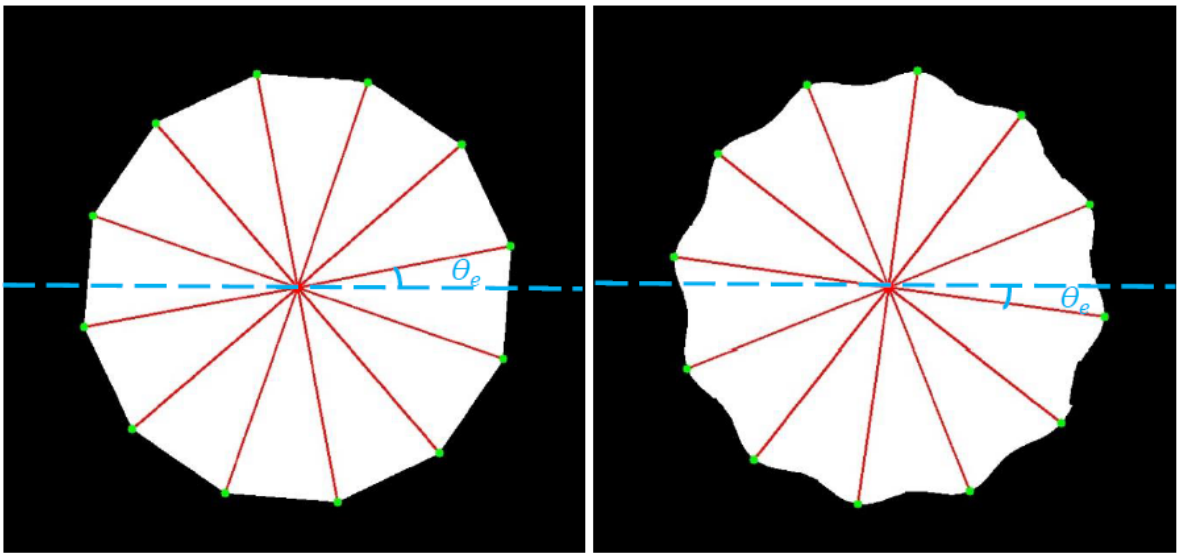


Figure 4.8 – Estimation de la rotation de la monnaie

.

à la binarisation de I^S. Mais les photos professionnelles nous permettent d'obtenir une binarisation suffisamment précise. Une estimation précise de N_e et S_e est importante car elle limitera la variation du modèle en évitant que le modèle ne tombe dans des minima locaux. En ce qui concerne les autres paramètres O_e, A_e, R_e et θ_e, ils sont supposés être dans le voisinage des valeurs précises.

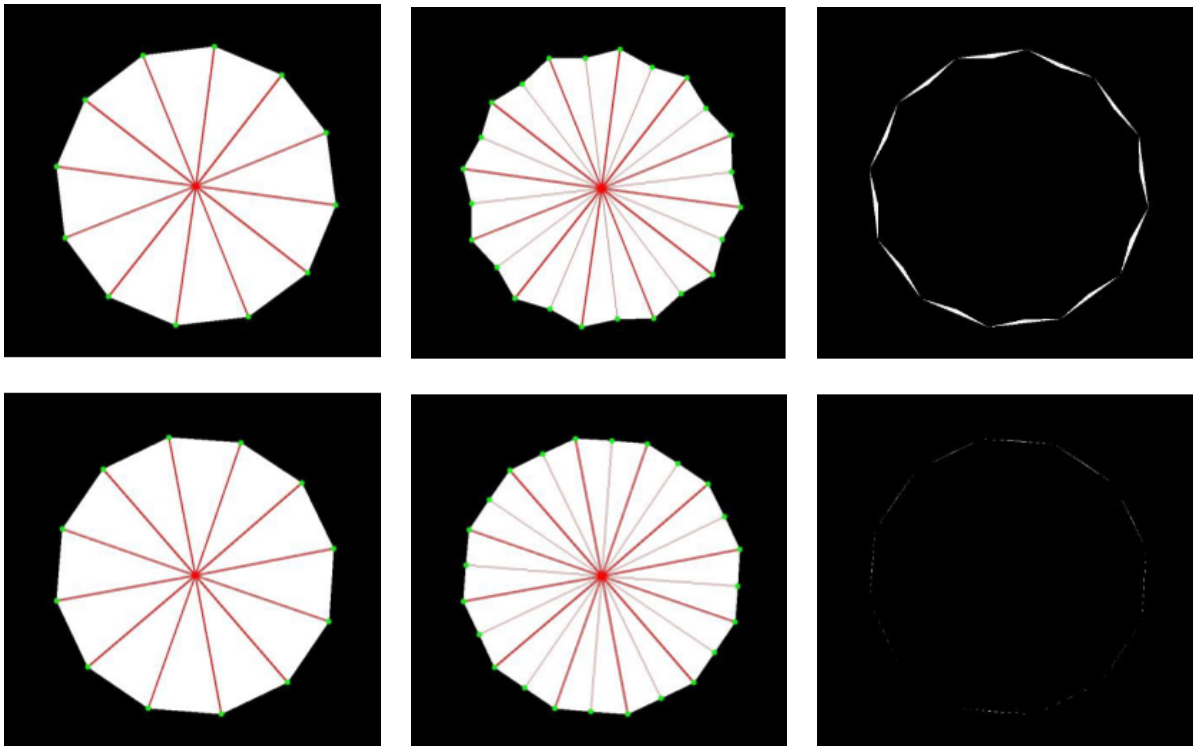
4.1.3 Segmentation précise par évolution du modèle actif

Le modèle initial $\Omega^{(0)}$ est obtenu à l'aide des paramètres estimés à l'étape précédente :

$$\Omega^{(0)}(O^{(0)} = O_e, R^{(0)} = R_e, A^{(0)} = A_e, \theta^{(0)} = \theta_e, N^{(0)} = N_e, S^{(0)} = S_e) \tag{4.8}$$

Notons $\Omega^{(t)}$ le modèle à l'instant t. L'énergie du modèle $E(\Omega^{(t)})$ est mesurée en considérant l'homogénéité intérieure et extérieure. La mesure de l'homogénéité est similaire à celle des contours actifs de Chan et Vese (2001). Soient c_{int} et c_{ext} les intensités moyennes à l'intérieur et à l'extérieur de la région définie par le modèle Ω,

$$\begin{aligned} c_{int}^{(t)} &= \frac{\int_{\Omega^{(t)}} I^S(x,y)dxdy}{\int_{\Omega^{(t)}} dxdy} \\ c_{ext}^{(t)} &= \frac{\int_{\overline{\Omega^{(t)}}} I^S(x,y)dxdy}{\int_{\overline{\Omega^{(t)}}} dxdy} \end{aligned} \tag{4.9}$$



(a) (b) (c)

FIGURE 4.9 – Estimation de la forme de la monnaie : (a) Ψ_1 ; (b) Ψ_2 ; (c) $|Area(\Psi_1) - Area(\Psi_2)|$

$E(\Omega^{(t)})$ est ensuite donnée par la somme de l'énergie extérieure et l'énergie intérieure :

$$E(\Omega^{(t)}) = E_{int}(\Omega^{(t)}) + E_{ext}(\Omega^{(t)}), \tag{4.10}$$

avec

$$\begin{aligned} E_{int}(\Omega^{(t)}) &= \int_{\Omega^{(t)}} |I^S(x,y) - c_{int}^{(t)}| dxdy \\ E_{ext}(\Omega^{(t)}) &= \int_{\overline{\Omega^{(t)}}} |I^S(x,y) - c_{ext}^{(t)}| dxdy \end{aligned} \tag{4.11}$$

Ensuite, nous faisons évoluer le modèle initial $E(\Omega^{(0)})$ de manière itérative pour qu'il s'adapte au mieux à la région segmentée. $\boldsymbol{\Omega}^{(t+1)}(\boldsymbol{O}^{(t+1)}, \boldsymbol{R}^{(t+1)}, \boldsymbol{A}^{(t+1)}, \boldsymbol{\theta}^{(t+1)}, \boldsymbol{N}^{(t+1)}, \boldsymbol{S}^{(t+1)})$ est l'ensemble des modèles possibles, à l'instant $(t+1)$, avec toutes les combinaisons des paramètres voisins. $\Omega' \in \boldsymbol{\Omega}^{(t+1)}$ est le modèle avec les paramètres retenus qui minimisent l'énergie à $(t+1)$, donné par,

$$\Omega' = \underset{\Omega \in \boldsymbol{\Omega}^{(t+1)}}{argmin}\ E(\Omega) \tag{4.12}$$

Le modèle optimal Ω^* sera obtenu lorsque l'énergie ne diminuera plus. Alors, les contours du modèle coïncideront avec le bord de la région segmentée. La région intérieure et la région extérieure définies par Ω^* seront quasi-homogènes.

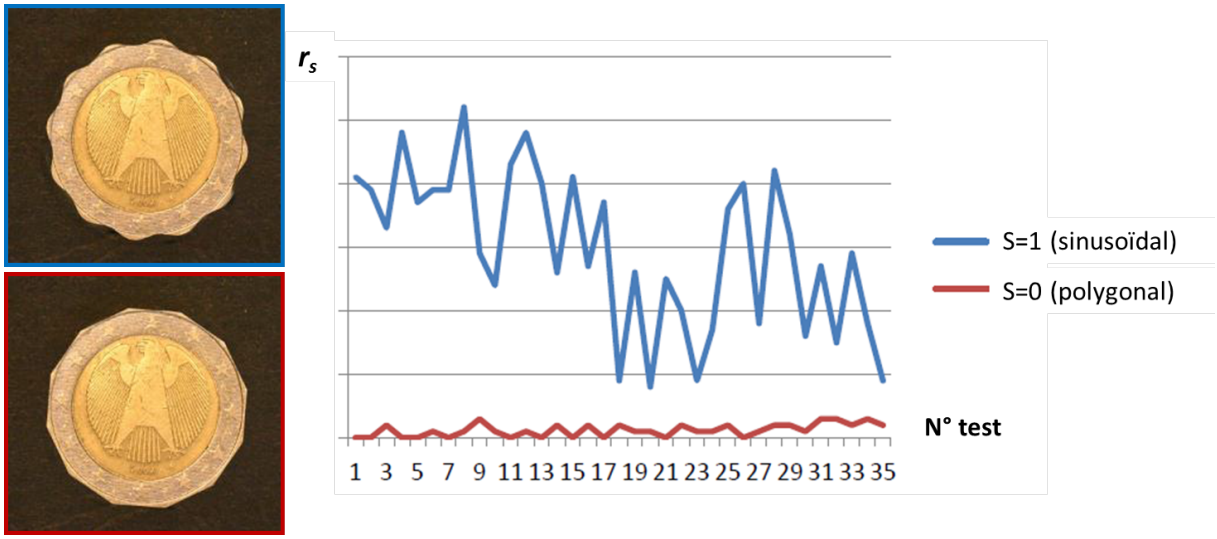
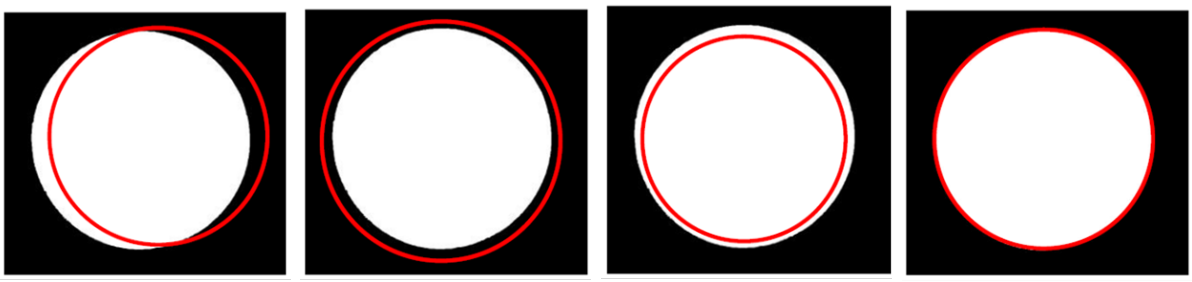


Figure 4.10 – Graphique des expériences sur r_S

.



(a) (b) (c) (d)

Figure 4.11 – Énergie du modèle avec différentes configurations : (a) $E_{ext} > 0, E_{int} > 0, E > 0$;(b)$E_{ext} = 0, E_{int} > 0, E > 0$; (c) $E_{ext} > 0, E_{int} = 0, E > 0$; (d)$E_{ext} = 0, E_{int} = 0, E = 0$

.

Théoriquement, il faut parcourir toutes les combinaisons des paramètres dans $\mathbf{\Omega}^{(t+1)}$ et sélectionner ceux qui minimisent l'énergie. Mais ce calcul sera trop lourd. Dans la pratique, il est possible d'optimiser les paramètres l'un après l'autre pour chaque itération.

$$\begin{aligned} O' &= \underset{O \in \mathbf{O}^{(t+1)}}{argmin}\ E_\Omega(O) \\ R' &= \underset{R \in \mathbf{R}^{(t+1)}}{argmin}\ E_\Omega(R) \\ A' &= \underset{A \in \mathbf{A}^{(t+1)}}{argmin}\ E_\Omega(A) \\ \theta' &= \underset{\theta^* \in \boldsymbol{\theta}^{(t+1)}}{argmin}\ E_\Omega(\theta) \end{aligned} \tag{4.13}$$

où,

$$
\begin{aligned}
\mathbf{O}^{(t+1)} &= \left\{ \begin{array}{lll} (x^{(t)}-dx, y^{(t)}-dy), & (x^{(t)}-dx, y^{(t)}), & (x^{(t)}-dx, y^{(t)}+dy) \\ (x^{(t)}, y^{(t)}-dy), & (x^{(t)}, y^{(t)}), & (x^{(t)}, y^{(t)}+dy) \\ (x^{(t)}+dx, y^{(t)}-dy), & (x^{(t)}+dx, y^{(t)}), & (x^{(t)}+dx, y^{(t)}+dy) \end{array} \right\} \\
\mathbf{R}^{(t+1)} &= \{R^{(t)}+dR, \quad R^{(t)}, \quad R^{(t)}-dR\} \\
\mathbf{A}^{(t+1)} &= \{A^{(t)}+dA, \quad A^{(t)}, \quad A^{(t)}-dA\} \\
\boldsymbol{\theta}^{(t+1)} &= \{\theta^{(t)}+d\theta, \quad \theta^{(t)}, \quad \theta^{(t)}-d\theta\}
\end{aligned}
\tag{4.14}
$$

Pour un maximum de précision, nous mettons dx, dy, dR et dA égaux à un pixel et $d\theta$ égal à un degré.

Nous aurions pu essayer de déterminer N par la même méthode en prenant N qui minimise l'énergie. Selon nos tests sur les monnaies polygonales, le bon N correspond toujours à celui dont l'énergie est minimale, mais des minima locaux dans la courbe $N - E_{min}$ impliquent des conditions d'arrêt compliquées. Dans la Figure 4.12, chaque courbe correspond à une monnaie polygonale, l'axe X correspond au nombre de côtés ou de crêtes estimé, l'axe Y donne l'énergie minimale. Par exemple, la courbe orange montre que l'énergie d'une monnaie polygonale comportant 16 côtés est minimale lorsque $N = 16$.

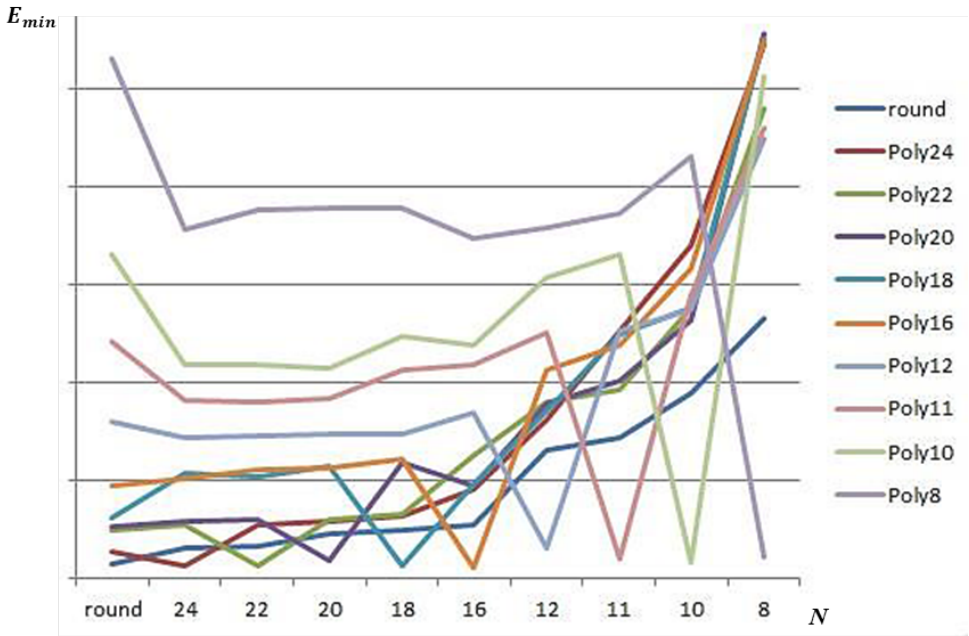


FIGURE 4.12 – Courbe $N - E_{min}$

.

Dans la pratique, nous supposons que N_e et S_e sont bien estimés. Nous imposons donc $N^{(t+1)} = N^{(t)} = N_e$ et $S^{(t+1)} = S^{(t)} = S_e$. Cela permet de réduire le nombre de paramètres qui doivent évoluer. Nous imposons également $A = dA = 0$ si $S_e = 0$ (la forme estimée

est polygonale), et $\theta = d\theta = 0$ si $N_e = 50$.

La segmentation finale I_{Ω^*} est obtenue via le modèle optimal Ω^*. Comme le modèle est une forme parfaite, tous les petits défauts sur le bord de la segmentation « grossière » sont « nettoyés ».

Un sous-ensemble des paramètres optimaux du modèle fournit une description globale de la forme de la monnaie segmentée :

$$F_{shape} = \{N^*, R^*, A^*, S^*\} \tag{4.15}$$

Cette description sera utilisée pour présélectionner les monnaies pour l'identification du type monétaire.

4.2 Optimisation et adaptation

4.2.1 k-estimation

Si les paramètres estimés du modèle initial sont trop loin des valeurs optimales, cela augmentera le nombre d'itérations et la probabilité que le modèle tombe dans des minima locaux. Par ailleurs, S_e et N_e sont deux paramètres essentiels dont nous souhaitons que leurs valeurs estimées soient précises. Dans la pratique, une seule estimation est parfois insuffisante. Au lieu de parcourir toutes les possibilités dans un jeu de paramètres pour obtenir des valeurs optimales, une solution souvent employée est de faire un recuit simulé pour sortir des minima locaux, mais cela risque également prendre du temps.

La solution proposée est d'estimer k modèles initiaux différents, en supposant qu'au moins un d'entre eux contient des paramètres estimés pertinents (notamment N_e et S_e) pour atteindre un bon résultat. Parmi tous les paramètres, seul O_e ne dépend que de la région segmentée I^S. Les autres dépendent fortement du segment le plus long (respectivement le plus court) initialisé. Nous estimons ainsi k fois le segment le plus long et le segment le plus court sur les k différentes sections du bord sur I^S afin d'obtenir k modèles estimés. Ces k modèles estimés devraient être quasi-identiques et leurs modèles optimaux respectifs devraient être les mêmes. Mais si les modèles optimaux obtenus à partir de ces k modèles initiaux sont différents, nous sélectionnons celui dont l'énergie est minimale.

$$\Omega^* = \underset{\Omega \in \{\Omega_1^*, \Omega_2^*, \ldots, \Omega_k^*\}}{argmin} E(\Omega) \tag{4.16}$$

Cette k-estimation diminue la sensibilité de notre algorithme à l'estimation. Avec un petit k, cette optimisation n'augmente pas trop le temps de calcul par rapport au parcours de tous les N et S possibles, et est plus efficace qu'un recuit simulé. Selon nos tests, $k = 2$ est suffisant pour obtenir des résultats robustes. Si les k modèles contiennent les mêmes N_e et S_e, et les autres paramètres sont proches, nous n'effectuons qu'une passe de l'évolution

du modèle. Par ailleurs, nous pouvons utiliser la k-estimation pour vérifier les cas qui comportent des symétries semblables, par exemple $N_e = 8$ et $N_e = 16$.

4.2.2 Adaptation aux monnaies trouées

Certaines monnaies peuvent posséder un trou au milieu qui permet de réunir plusieurs monnaies sur une barre ou par une corde afin de faciliter le stockage ou le transport. Pour les monnaies françaises traitées fréquemment par la société GENI, le trou central contient deux formes (carré et rond) (*cf.* Figure 4.13). À noter qu'il existe des monnaies avec un trou décentré ou de forme irrégulière, mais ces cas rares ne sont pas pris en compte dans le cadre de cette thèse.

(a) (b)

FIGURE 4.13 – Exemples de monnaies trouées françaises : (a) monnaie avec un trou rond (Essai Michelin Théodore 5 centimes) ; (b) monnaie avec un trou carré (Essai Michelin Théodore 4 centimes)

La méthode que nous avons détaillée précédemment ne considère pas un trou éventuel au milieu. Les pixels appartenant au trou sont considérés comme des pixels de la région de la monnaie lors de la segmentation « grossière ». Pour traiter les monnaies trouées, l'idée est de considérer la région du trou comme un objet d'intérêt en utilisant la même méthode que nous avons proposée pour la monnaie entière. Il faut ainsi détecter approximativement la région du trou, notée I^S_{hole}. Le processus de la segmentation des monnaies trouées est résumé à la Figure 4.14

Il s'agit de détecter un trou éventuel de la monnaie dans la carte des probabilités I^P, dans laquelle les pixels appartenant au trou sont de même couleur que l'arrière-plan. Notons $\overline{I^P}$ le complémentaire de I^P, trois hypothèses basées sur des connaissances *a priori* d'un trou central de la monnaie peuvent être faites :

- Un trou est toujours au centre de la monnaie.

- Un trou a une forme soit ronde soit carrée, c'est-à-dire que le ratio d'aspect de la boîte englobant du trou est proche de 1.

- Un trou a une taille relative à celle de la monnaie.

Notons respectivement l_i, h_i et o_i la largeur, la hauteur et le centre de chaque composante connexe $C_i \in \overline{I^P}$, nous filtrons C_i selon les trois critères suivants :

- position : $\|o_i - O_e\|_2 < d_{th}$,
- ratio d'aspect : $r_{th1} < \dfrac{l_i}{h_i} < r_{th2}$;
- aire : $a_{th1} < Area(C_i) < a_{th2}$;

où $d_{th}, r_{th1}, r_{th2}, a_{th1}$ et a_{th2} sont des seuils fixés expérimentalement.

La seule composante connexe retenue représente grossièrement la région du trou I^S_{hole}. Si aucune composante connexe ne correspond aux critères, nous pourrons dire que c'est une monnaie normale sans trou. Ensuite, en utilisant la même modélisation mais avec $\boldsymbol{N}_{hole} = \{4, 50\}$ et $\boldsymbol{S}_{hole} = \{0\}$, nous cherchons le modèle optimal Ω^*_h pour la région du trou. La segmentation précise du trou est notée $I_{\Omega_{hole}{}^*}$.

La segmentation finale de la monnaie trouée est donnée par,

$$I_{coin} = I_{\Omega^*} \oplus I_{\Omega_{hole}{}^*} \tag{4.17}$$

Les caractéristiques de forme de la monnaie trouée sont données par la concaténation des paramètres optimaux de la monnaie et ceux du trou.

$$F_{shape} = \{N^*, R^*, A^*, S^*, N_{hole}{}^*, R_{hole}{}^*\} \tag{4.18}$$

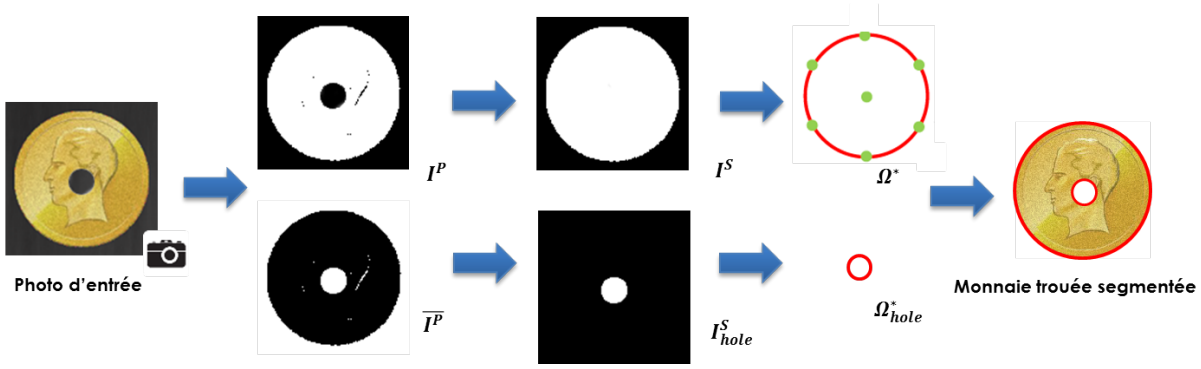


Figure 4.14 – Processus de segmentation des monnaies trouées

.

4.3 Expériences et résultats

Dans ce paragraphe nous présentons les expériences conduites en utilisant la méthode présentée dans la section 4.1 avec les optimisations proposées dans la section 4.2. La validation de notre méthode est effectuée sur la base de données *LIRIS_CoinPhotos* car elle comporte des monnaies synthétiques de formes particulières prises en compte par notre

méthode. Nous avons également testé la méthode sur la base de données professionnelle *GENI_coinPhotos* mais elle ne contient quasiment que des monnaies rondes ou quasi-rondes.

4.3.1 Expériences sur *LIRIS_CoinPhotos*

Comme nous l'avons présenté dans le Chapitre 3, la base de données *LIRIS_CoinPhotos* comporte 250 images de monnaies. Ce sont des monnaies réelles multinationales de forme ronde, et des monnaies synthétiques de forme polygonale, sinusoïdale et des monnaies trouées.

Rappelons que l'objectif est de détourer les monnaies afin de trouver une forme géométrique parfaite qui englobe au mieux la région de la monnaie. Ainsi, il ne s'agit pas de comparer la segmentation résultante avec la vérité terrain de la monnaie mais de s'assurer que la bonne forme est trouvée. En effet, grâce à la qualité des photos de la base *LIRIS_CoinPhotos*, la segmentation « grossière » est déjà très proche de la vérité terrain en termes de mesures classiques comme le coefficient de Dice utilisé dans le travail de Zambanini et Kampel (2009). Nous validons donc la segmentation en comparant le modèle optimal et la forme effective.

Dans cette base de données, nous avons 53 monnaies sinusoïdales et 197 monnaies polygonales (comportant les monnaies rondes). Par ailleurs, 34 monnaies comportent un trou synthétique rond ou carré au milieu. Parmi celles-ci, 35 monnaies sinusoïdales et 35 monnaies polygonales sont à l'origine de la même monnaie réelle et comportent le même nombre de côtés ou crêtes synthétiques. Nous avons tout d'abord fait un test sur ces données pour valider l'estimation de forme S_e (*cf.* Figure 4.10).

Les résultats obtenus présentés dans la Table 4.1 sont très encourageants d'autant que les conditions d'images dans *LIRIS_CoinPhotos* ne sont pas aussi optimisées comme des photos professionnelles. La Figure 4.15 montre des exemples de segmentations précises obtenues : le contour rouge est le modèle de forme extrait. Notons que les mauvaises segmentations sont généralement causées par une mauvaise estimation pour des monnaies polygonales avec N élevé (proche d'une monnaie ronde). Ceci est dû au fait l'influence des petits bruits du bord est plus importante que pour les monnaies avec un N petit.

Pour évaluer le temps de calcul, nous avons testé l'algorithme sur deux tailles d'images : le premier test effectué sur l'image originale de taille 1400 $\times$ 1400 pixels environ, et le deuxième sur une version réduite de taille 700 $\times$ 700 pixels afin d'obtenir des images affichables sur l'écran de l'ordinateur. Il faut mentionner que la taille des photos professionnelles est normalement encore plus importante. Selon nos tests, la taille de l'image ne change pas la qualité de la reconnaissance de la forme mais influence beaucoup sur le temps de calcul (*cf.* Table 4.2). Le calcul des deux premières étapes est relativement rapide (moins d'une seconde) mais la recherche du modèle optimal (troisième étape) dé-

Table 4.1 – Performance concernant l'estimation de forme des monnaies sur des images de taille 700 × 700 pixels environ de la base *LIRIS_CoinPhotos*

Tests	**N. photos**	**N. succès**	**Taux de réussite (%)**
Estimation S	70	70	100
Estimation N (sin)	53	53	100
Estimation N (poly)	197	195	98,9
Détection du trou	34	34	100
Total	250	248	99,2

Table 4.2 – Temps de calcul sur *LIRIS_CoinPhotos* (AMD Turion II 2,4GHz, 2G RAM, Win7 32bits)

Taille (pxiels)	**Temps moyen (s)**	**Temps max (s)**
700 × 700	0,8	3,1
1400 × 1400	2,9	7,5

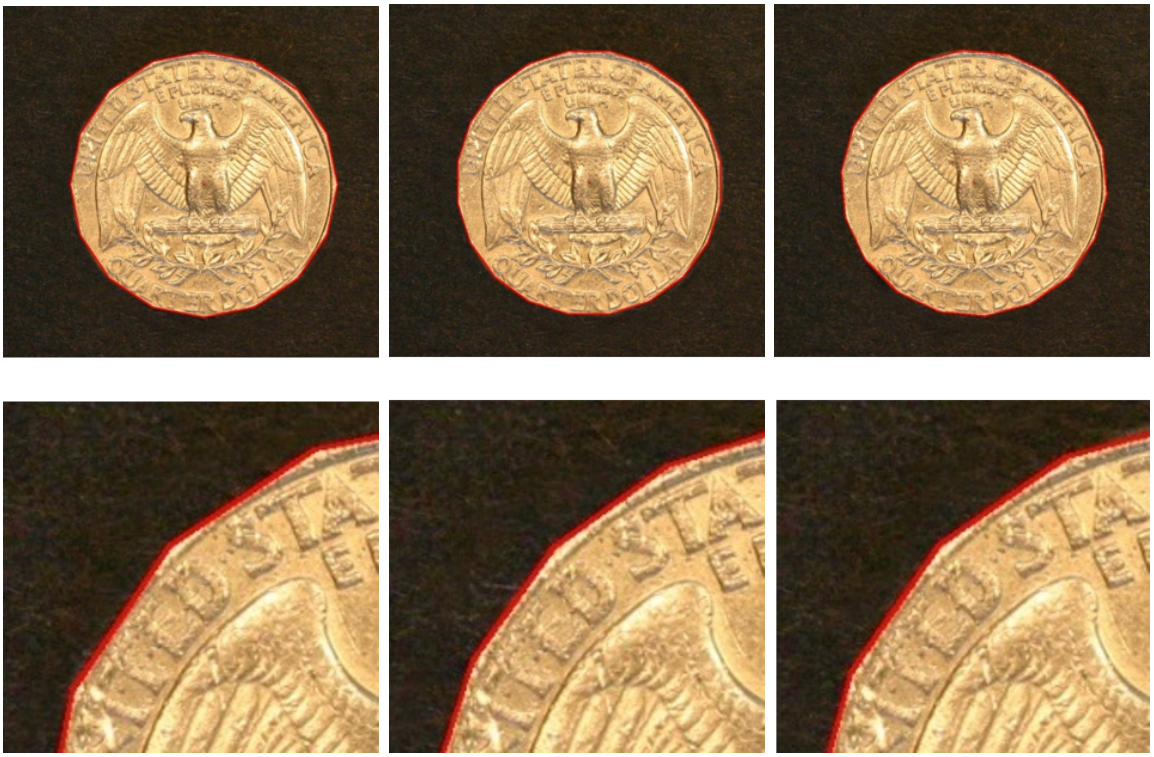
pend du nombre d'itérations. Évidemment, plus l'image a une taille importante, plus le nombre d'itération augmente et donc plus le calcul est coûteux.

4.3.2 Expériences sur *GENI_CoinPhotos*

Nous avons également testé notre méthode sur la base de données *GENI_CoinPhotos*. Ce sont des photos professionnelles mais ne comportant actuellement pas beaucoup de monnaies de formes particulières. Dans certaines photos, le photographe de GENI utilise un fond noir pour minimiser l'effet de l'ombre mais, en même temps, augmente le risque de mélanger l'arrière-plan noir avec le bord foncé de certaines monnaies. Grâce à la robustesse de notre méthode, nous avons réussi à segmenter toutes les monnaies rondes et celles de forme particulière par la méthode proposée. Certains exemples sont illustrés à la Figure 4.16. Nous pouvons voir qu'il est parfois difficile de reconnaître la forme exacte à l'œil nu sur la segmentation « grossière » bruitée. Cependant, le modèle optimal est robuste et permet de trouver la forme exacte de la monnaie à segmenter.

Il faut mentionner que cette base de données comporte également des monnaies particulières que nous ne pouvons pas exactement décrire par la modélisation proposée :

- Monnaies striées ou cannelées : il s'agit de monnaies rondes mais ayant le bord strié (*cf.* Figure 4.17 (a)) ou cannelé (*cf.* Figure 4.17 (b)). La cannelure est crénelée, alors que la striure a des pointes saillantes en forme de triangle. Parfois la cannelure ou la striure n'est présentée que sur une partie du bord. Dans le cadre de notre méthode, elles sont considérées comme des monnaies rondes avec un bord régulièrement bruité. En conséquence, une partie du bord sera légèrement coupée dans la segmentation finale (*cf.* Figure 4.17 (a)(b)).



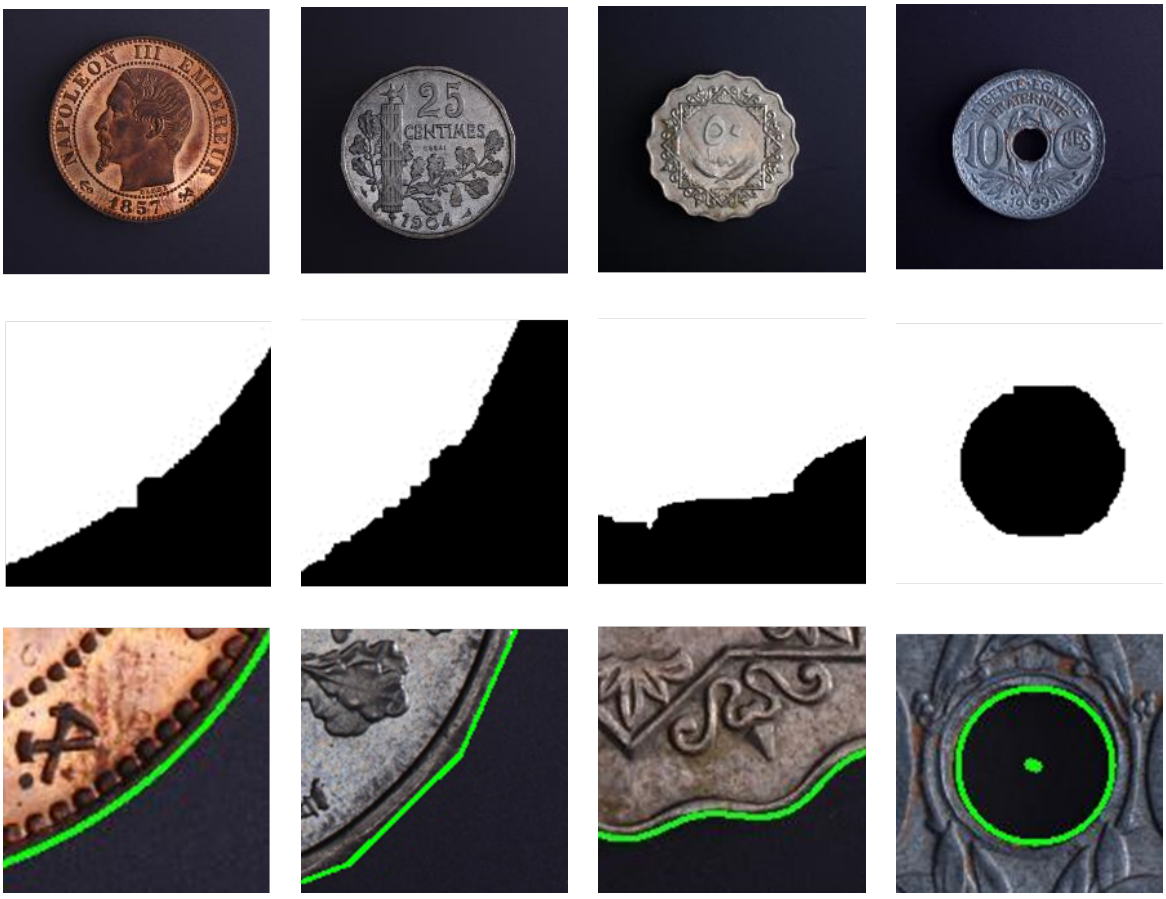
(a) (b) (c)

FIGURE 4.15 – Exemples de monnaies segmentées (repérées par un contour rouge) de la base *LIRIS_coinPhotos* à l'aide du modèle optimal : (a) monnaie à 16 pans, (b) à 22 pans, et (c) à 24 pans

.

- Monnaies de forme irrégulière quasi-ronde : il s'agit de monnaies de forme irrégulière mais quasi-ronde (*cf.* Figure 4.17 (c)). La segmentation résultante est une région ronde qui correspond le mieux à la région de la monnaie. Le résultat obtenu ressemble à la segmentation « grossière » de monnaies anciennes à l'aide de la transformée de Hough dans le travail de Arandjelovic (2010), mais la région ronde approximative est obtenue basée sur la région au lieu des contours.

- Monnaies distordues : Il s'agit de monnaies légèrement distordues à cause de la prise de photo. En effet, lorsque l'objectif n'est pas parfaitement parallèle au plan qui contient la monnaie, dans l'image la monnaie ronde apparaît comme une ellipse. Puisque le modèle appliqué est parfaitement rond, la segmentation résultante peut comporter quelques pixels de l'arrière-plan. Un exemple est illustré à la Figure 4.18. Sur la monnaie segmentée zoomée, nous pouvons trouver des bords noirs (appartenant à l'arrière-plan) de largueur d'un pixel en haut et en bas, mais les bords noirs n'existent pas à gauche et à droite. C'est la preuve que la monnaie dans l'image n'est pas exactement ronde. Nous observons également que les images segmentées de la même monnaie originelle, par exemple l'avers et le revers, n'ont pas toujours la même taille. Selon les tests sur *GENI_coinPhotos*, nous avons une erreur moyenne de 0,15% en diamètre sur les mêmes monnaies dans les



(a) (b) (c) (d)

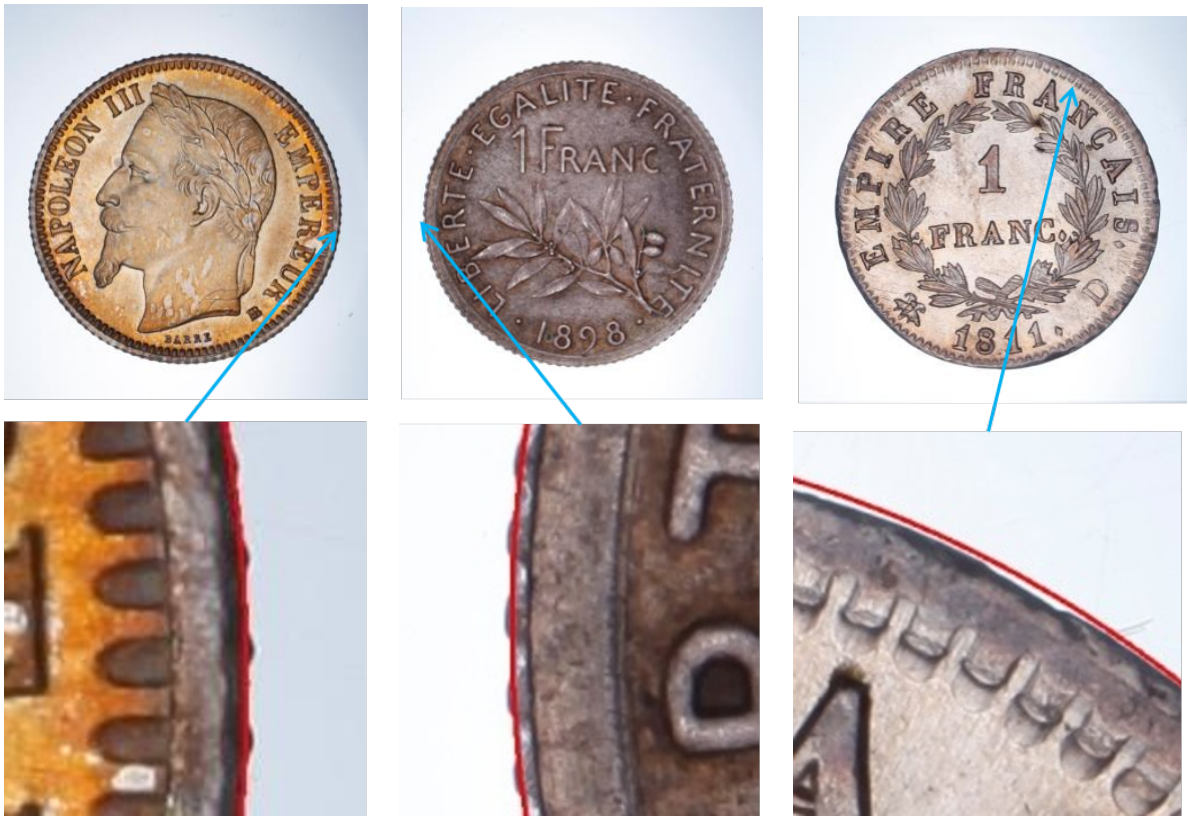
Figure 4.16 – Exemples de monnaies segmentées de la base *GENI_CoinPhotos* (première ligne : images originales ; deuxième ligne : segmentations « grossières » zoomées ; troisième ligne : segmentations précises repérées par un contour vert) : (a) monnaie ronde, (b) monnaie polygonale, (c) monnaie sinusoïdale, (d) monnaie trouée

.

différentes images. D'une part, les segmentations « grossières » ne sont pas exactement pareilles ; d'autre part, la position relative de la monnaie par rapport à l'objectif de la caméra est différente pour chaque prise de photo. Néanmoins, une telle imprécision due à la distorsion est très subtile, et est uniquement détectable sur des images de résolution importante. Elle est imperceptible à l'œil nu. La segmentation d'une monnaie légèrement distordue est donc considérée comme acceptable. Il n'est pas nécessaire d'envisager un modèle elliptique. Dans l'exemple montré à la Figure 4.18, il suffit d'utiliser un arrière-plan quasi-blanc pour avoir une illustration « parfaite » sans bord noir.

4.4 Conclusion

Dans ce chapitre, une nouvelle méthode de segmentation des monnaies est présentée. À l'aide d'un modèle paramétrique actif, nous avons réussi à segmenter de manière précise les monnaies dans les photos professionnelles. Par rapport à la segmentation des monnaies



(a) (b) (c)

FIGURE 4.17 – Exemples de monnaies particulières dans la base *GENI_CoinPhotos* (en haut : images originales ; en bas : segmentations zoomées repérées par un contour rouge) : (a) monnaie striée, (b) monnaie cannelée (c) monnaie quasi-ronde

.

dans la littérature, la méthode proposée possède les avantages suivants :

- Segmentation de monnaies de formes différentes : c'est la première méthode qui sert à segmenter des monnaies pas uniquement rondes, mais également des monnaies à pans (polygonales), des monnaies à dents (sinusoïdales) et des monnaies trouées.

- Précision au pixel près : cette précision fournit des illustrations parfaites pour le catalogue de l'entreprise. Par ailleurs, elle sert également à mesurer le diamètre de la monnaie en pixels. Le diamètre réel de la monnaie pourrait être ensuite obtenu en connaissant la focale de la lentille et la distance focale-objet, ou à l'aide d'une petite référence telle qu'un échiquier de taille connue dans un coin fixe de la photo originale. Cette information pourrait servir à l'identification du type monétaire.

- Extraction des caractéristiques de la forme de monnaie : les caractéristiques de forme de la monnaie sont extraites en même temps que la segmentation. Pour identifier le type monétaire, nous pourrions utiliser ces caractéristiques de forme pour présélectionner des types monétaires.

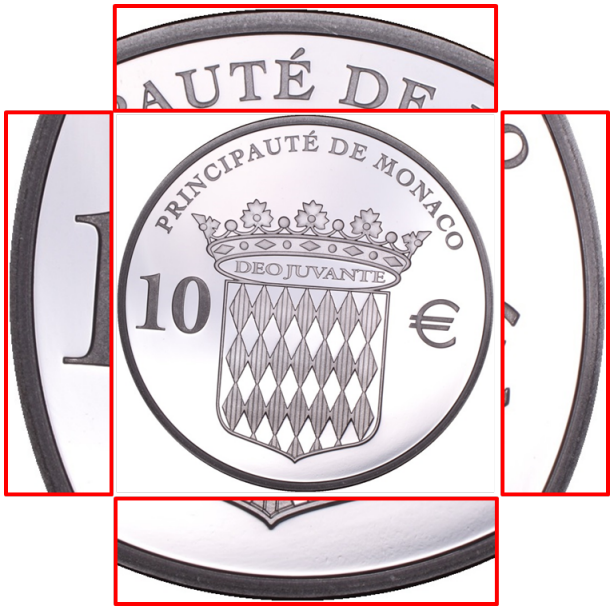


Figure 4.18 – Monnaie segmentée d'un arrière-plan noir avec des bords zoomés encadrés en rouge

.

En ce qui concerne les inconvénients, nous ne pouvons pas segmenter précisément des monnaies de forme irrégulière où avec des formes « exotiques » non comprises dans la modélisation actuelle. Par exemple, une monnaie de Dubaï à sept pans courbés est reconnue comme ronde (*cf.* Figure 4.19), car nous n'avons pas $N = 7 \in \boldsymbol{N_{all}}$. Nous pourrions éventuellement le rajouter mais cela n'a pas d'intérêt pour la société GENI pour l'instant. Pour certaines monnaies quasi-polygonales à des coins arrondis, la segmentation aux endroits des coins n'est pas parfaite (*cf.* Figure 4.20). La modélisation actuelle se base principalement sur des monnaies françaises que la société GENI souhaite traiter. Pour des monnaies étrangères ayant un nombre de coins ou de crêtes particuliers, il suffit d'adapter $\boldsymbol{N_{all}}$ sans modifier le nombre de paramètres du modèle. Pour adapter à des formes régulières « exotiques », y compris des pans courbés ou des coins arrondis, il est possible d'ajouter des paramètres supplémentaires comme le rayon du coin et le rayon du pan pour que les formes considérées puissent être modélisées de manière précise. La segmentation sera suivie la même procédure. Finalement, pour des monnaies irrégulières quasi-rondes, ou des monnaies striées ou cannelées, un travail supplémentaire sur le bord serait à rajouter sur la forme générale de la monnaie comme ce que Arandjelovic (2010) a proposé pour des monnaies anciennes.

Par rapport à des méthodes naïves, notre méthode permet d'obtenir de la précision mais avec un temps de calcul plus important. Cependant, grâce à la qualité des photos professionnelles, les paramètres sont initialisés dans le voisinage des bonnes valeurs, ce qui minimise quand même le temps de calcul. La combinaison de l'estimation et l'évolution du modèle nous offre un bon compromis entre temps de calcul et précision. Par ailleurs,

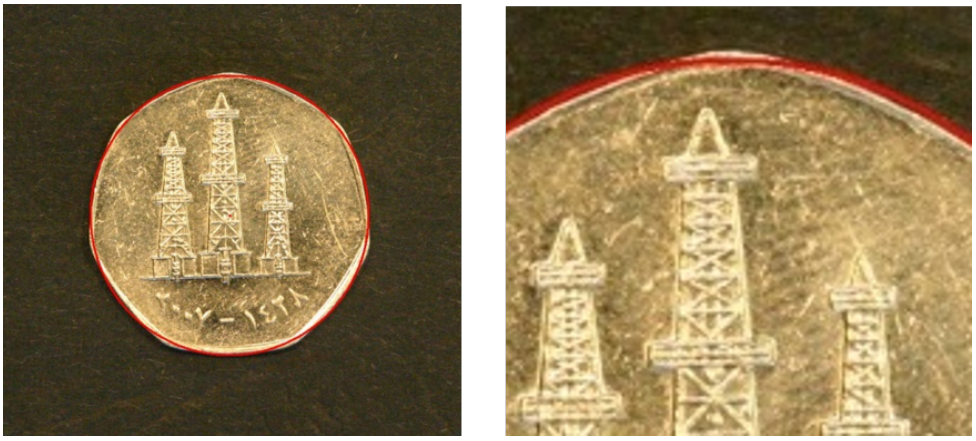


FIGURE 4.19 – Segmentation d'une monnaie polygonale ayant des pans courbés (repérée par un contour rouge)

.

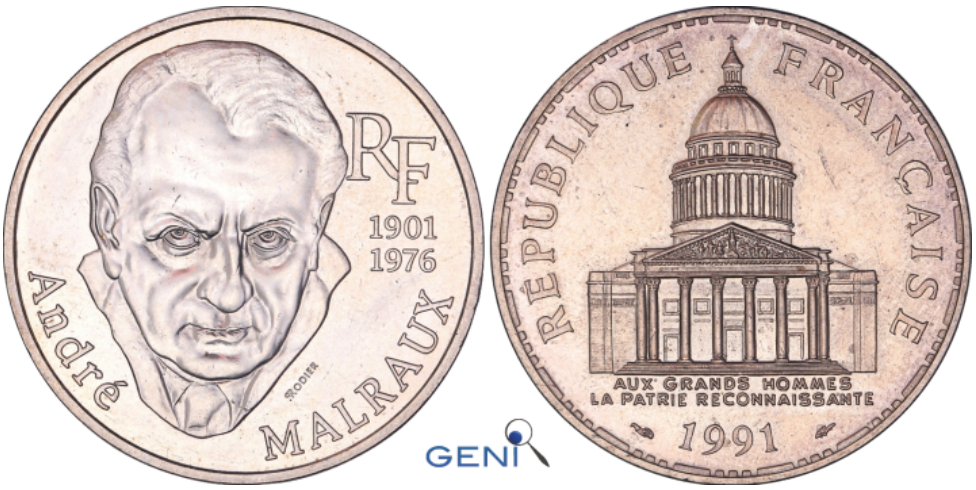
FIGURE 4.20 – Segmentation d'une monnaie polygonale ayant des coins arrondis (repérée par un contour rouge)

.

nous avons également optimisé l'algorithme afin de ne pas parcourir tous les paramètres. Si l'image d'entrée est trop grande, inspiré par le schéma *corse-to-fine* proposé par Zambanini et Kampel (2012), nous pouvons d'abord travailler sur une version réduite de l'image et puis, sur l'image de taille originale. Nous assurons ainsi peu itérations sur l'image de taille originale. Au vu des expériences réalisées, la société GENI est satisfaite du temps de calcul actuel de la méthode proposée, même si elle peut encore être optimisée.

Chapitre 5

Identification du type monétaire



DANS ce chapitre, nous détaillons notre méthode d'identification du type monétaire. Comme nous l'avons mentionné précédemment, le nombre de données est souvent insuffisant pour les approches basées apprentissage dans le contexte numismatique. Dans ce cas-là, l'identification du type monétaire se base sur la recherche du type monétaire de référence correspondant à la monnaie requête. Deux scores de similarité ont été proposés pour mesurer la similarité entre une monnaie requête et une référence.

D'un point de vue numismatique, un type monétaire est déterminé par les motifs des deux faces (avers et revers) et par la tranche. Cependant, notre algorithme développé se concentre sur la comparaison de deux images. Par simplicité, dans ce chapitre et dans la suite de cette thèse, le terme « type monétaire » est souvent utilisé pour décrire la classe d'une seule face d'une monnaie. Dans ce cas-là, deux images de deux faces en provenance d'une même monnaie sont considérées comme deux classes différentes à distinguer.

Dans la section 5.1, nous décrivons l'algorithme d'identification du type monétaire à l'aide de deux scores de similarité : le premier score se base sur la correspondance entre caractéristiques locales ; le deuxième score se base sur les caractéristiques semi-globales. La section 5.3 présente les expériences conduites sur la base de données *USA_Grading*. Enfin la section 5.4 est la conclusion.

5.1 Méthodologie proposée

En supposant que le type monétaire de la monnaie à identifier existe déjà dans la base de références, le type monétaire est celui de la référence qui ressemble le plus à la monnaie requête. Ainsi, nous cherchons une mesure de similarité qui permet de déterminer si la monnaie requête et une monnaie de référence appartiennent au même type monétaire.

Pour mesurer la similarité entre deux images dans notre contexte, nous avons proposé deux scores de similarité. Le premier score de similarité est le nombre de descripteurs locaux en correspondance. Ce descripteur mesure ainsi la similarité des motifs locaux. Pour réaliser cela, nous détectons d'abord des points d'intérêt sur les deux images à comparer. Ensuite, nous mettons les points d'intérêt en correspondance en fonction de leurs descripteurs et de certaines contraintes spatiales. Le second score de similarité est la combinaison de la similarité locale et d'une similarité semi-globale menée à l'aide de graphes. Les scores de similarité sont calculés sur deux monnaies de forme identique et de taille normalisée, pour cela nous utilisons les résultats du Chapitre 4.

La Figure 5.1 présente le schéma global utilisé pour calculer ces deux scores de similarité que nous allons détailler dans les paragraphes suivants. Par ailleurs, dans le traitement global, l'identification du type monétaire est l'étape suivante de la segmentation de monnaies. Les caractéristiques de forme que nous avons obtenues lors de la segmentation nous permettent de présélectionner les références pour la monnaie requête (*cf.* Figure 5.2).

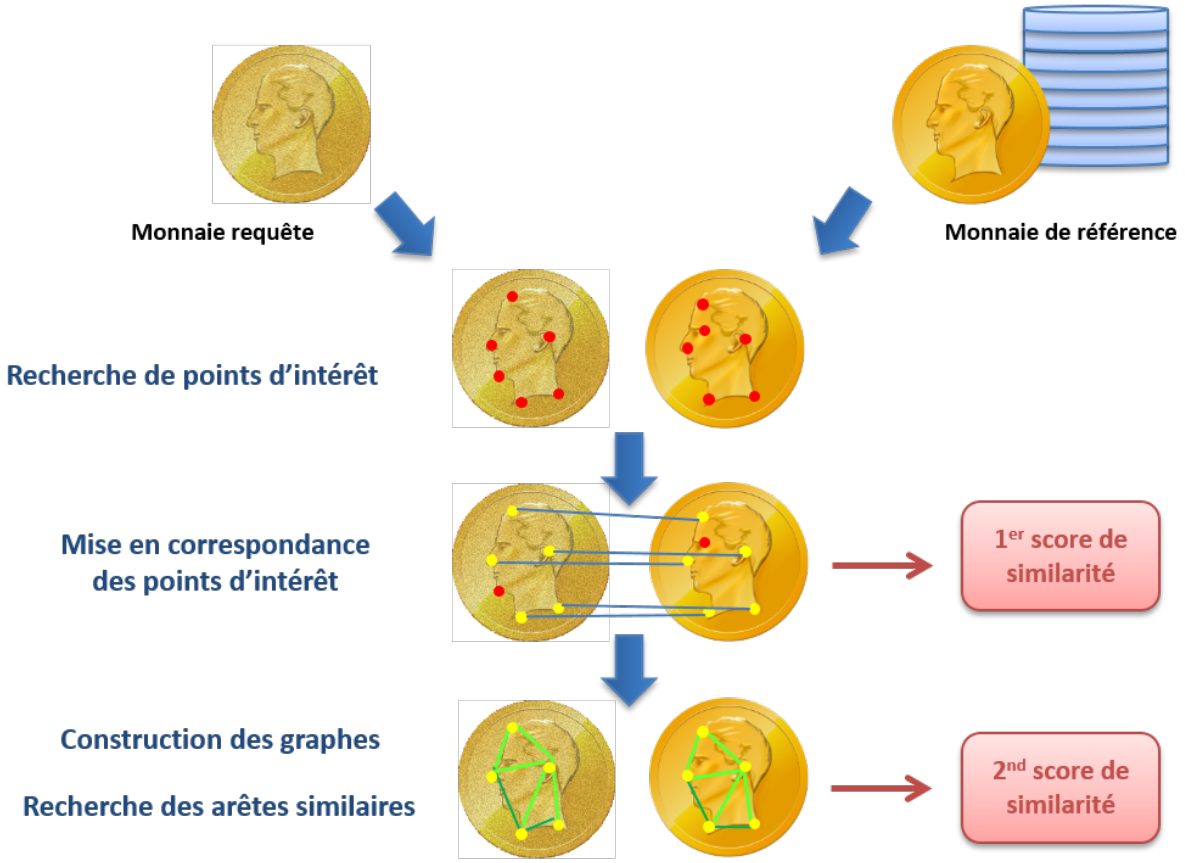


Figure 5.1 – Schéma global du calcul des scores de similarité

5.1.1 Score de similarité à l'aide de correspondances locales

Dans l'état de l'art, l'appariement de descripteurs locaux est l'une des méthodes les plus populaires pour traiter les monnaies anciennes (Kampel et Zaharieva 2008). Dans le cadre de cette thèse, les monnaies « modernes » que nous traitons présentent une difficulté similaire aux monnaies anciennes : la description globale du type monétaire n'est pas toujours discriminante. Concernant les monnaies anciennes, ceci provient de la dégradation élevée et de la déformation intraclasse. Alors que dans notre contexte, nous avons une similarité intraclasse élevée pour certains types monétaires qui rend les approches globales peu performantes. Pour être robuste aux détails des motifs de la monnaie, nous caractérisons les monnaies à l'aide des caractéristiques locales.

Détection de points d'intérêt locaux

Dans le domaine de la vision, des caractéristiques locales invariantes à l'échelle ou à la rotation sont souvent utilisées pour résoudre le problème général de la détection et de la recherche de correspondances entre images. Les caractéristiques les plus discriminantes dans l'image, par exemple les angles du contour ou les changements brusques d'intensité, sont encodés par des points d'intérêt (*keypoints*). De nombreux algorithmes de détection de points de caractéristique sont étudiés depuis une vingtaine d'années. Tous les algo-

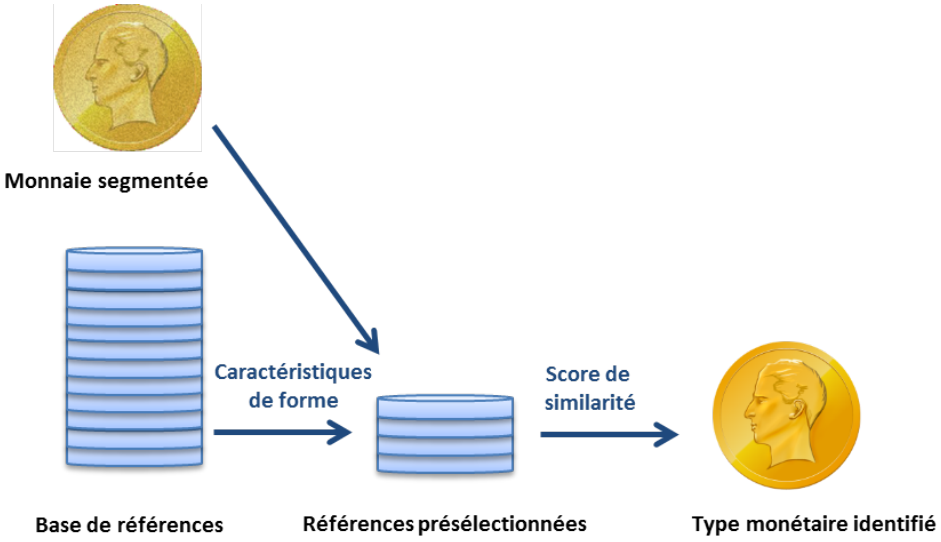


FIGURE 5.2 – Pipeline de l'identification du type monétaire

rithmes considérés contiennent souvent deux parties principales : le détecteur qui localise les points d'intérêt sur l'image, et le descripteur qui encode les caractéristiques locales. La plupart des descripteurs locaux mentionnés dans la littérature sont des descripteurs basés imagettes car ils encodent les imagettes centrées sur des points d'intérêt. L'ensemble des descripteurs de points d'intérêt établit ainsi une véritable signature numérique du contenu de l'image.

Dans le contexte de l'identification du type monétaire, nous souhaitons que le détecteur localise des points d'intérêt autour des motifs en relief de la monnaie. Quant au descripteur, idéalement il devrait être sensible aux motifs des monnaies mais robuste au bruit « non-motif » comme la texture du métal, l'usure ou l'éclairage. Plus précisément, nous souhaitons que les imagettes contenant des motifs similaires puissent avoir des descripteurs proches.

L'invariance du descripteur à certains changements joue également un rôle important. Puisque nous travaillons sur des monnaies segmentées de photos professionnelles, l'invariance au changement d'échelle et à une transformation affine semble peu importante. Par contre, l'invariance à la rotation est importante puisque des monnaies n'ont pas exactement la même orientation.

Parmi tous les algorithmes de détection de points d'intérêt, SIFT est le plus populaire pour la description des monnaies anciennes (Kampel et Zaharieva 2008 ; Zambanini et Kampel 2012). Cependant, il s'agit d'un algorithme gourmand en temps de calcul. Sachant que certaines invariances du descripteur ne sont pas obligatoires dans notre contexte, il est intéressant d'utiliser un algorithme performant mais plus léger. Après avoir comparé de nombreux algorithmes de la littérature (Lowe 2004 ; Bay, Tuytelaars et Van Gool 2006 ; Agrawal, Konolige et Blas 2008 ; Calonder et al. 2010 ; Leutenegger, Chli et Siegwart 2011 ; Rublee et al. 2011), nous avons sélectionné ORB (*Oriented FAST and Rotated BRIEF*)

(Rublee et al. 2011) pour les raisons suivantes :

- ORB est beaucoup plus rapide que les autres algorithmes étudiés (100 fois plus rapide que SIFT) ;

- la distribution des points d'intérêt localisés par son détecteur oFAST (*Oriented Features from Accelerated Segment Test*) semble plus appropriée à notre contexte. Plus précisément, nous avons comparé les points d'intérêt détectés par différents algorithmes sur la même image, et ORB extrait des points d'intérêt plus concentrés sur les motifs en relief que les autres (*cf.* Figure 5.3).

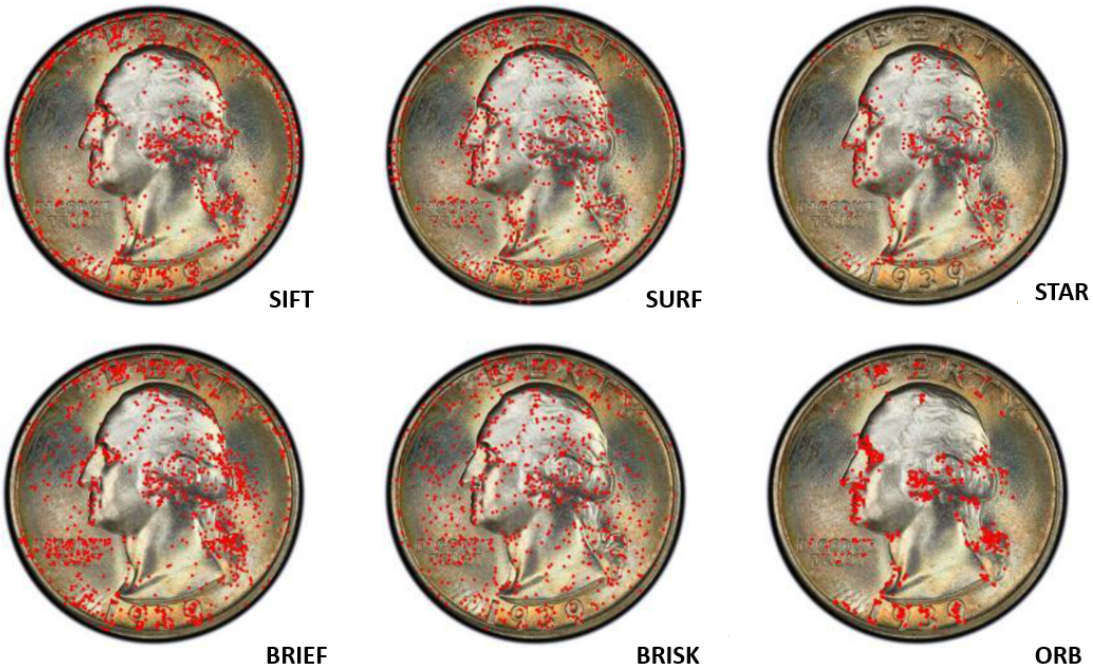


Figure 5.3 – Points d'intérêts détectés par différents algorithmes sur la même image

5.1.2 Mise en correspondance des points d'intérêt

Un point d'intérêt détecté est noté $v_i(p_i, d_i)$, où p_i est sa position et d_i est son descripteur qui caractérise l'imagette centrée sur p_i. Après l'extraction des descripteurs $d_i^Q \in D^Q$ de l'image de la monnaie requête I^Q, et $d_j^R \in D^R$ de l'image de la monnaie de référence I^R, l'objectif est de les mettre en correspondance. Cette étape est également appelée *matching* ou appariement des descripteurs locaux.

Normalement, la distance entre descripteurs devrait être petite si les motifs locaux qu'ils encodent sont similaires. Par exemple, sur les deux monnaies à la Figure 5.4 (a), le descripteur du point d'intérêt sur le nez du portrait dans l'image de gauche devrait être proche de celui sur le nez dans l'image de droite. Ainsi, deux points d'intérêt en

correspondance, notés $v_i^Q \leftrightarrow v_j^R$, devraient correspondre à des descripteurs dont la distance L2 est par,

$$v_i^Q \leftrightarrow v_j^R : \|d_i^Q - d_j^R\|_2 < d_{th} \tag{5.1}$$

où d_{th} est un seuil empirique. S'il n'y a aucun $d_j^R \in D^R$ ayant une distance L2 inférieure à d_{th} pour d_i^Q, v_i^Q n'a pas de point apparié dans l'image I^R.

Dans la Figure 5.4 (b), tous les $d_i^Q \in D^Q$ et $d_j^R \in D^R$ sont reliés en fonction de leurs distances. Cependant, il est à noter que les points d'intérêt sont souvent mis en correspondance à tort en utilisant une simple distance. Par ailleurs, un descripteur obtient parfois plusieurs descripteurs correspondants, et nous souhaitons établir des correspondances « un-à-un » entre les points d'intérêts.

Les mauvaises correspondances ont plusieurs raisons. D'une part, les descripteurs ORB ne sont pas toujours bien discriminants ; d'autre part, d'éventuels motifs locaux répétitifs se trouvent dans différents endroits et peuvent augmenter la confusion (par exemple, les étoiles qui entourent le portrait dans la Figure 5.4).

Pour cela, l'objectif suivant est d'augmenter les « vraies » correspondances et de diminuer les « fausses » correspondances.

(a) (b)

FIGURE 5.4 – Détection et appariement des points d'intérêt ORB sur deux monnaies *US Barber Quarter obv.* (points rouges : points d'intérêt détectés non appariés ; points jaunes : points d'intérêt appariés) : (a) une paire des points d'intérêt appariés sur les nez du portrait ; (b) appariement naïf de tous les points d'intérêt en utilisant un seuil de distance.

.

- Nombre total de points d'intérêt : le nombre total de points d'intérêt N_{kpt} que nous souhaitons détecter est un paramètre sur lequel nous pouvons jouer pour contrôler le nombre de correspondances. Plus nous détecterons de points d'intérêt, plus nous obtiendrons des « vraies » correspondances mais aussi des « fausses » correspondances.

- Méthode des k-plus proches voisins : une « vraie » correspondance de points d'intérêt ne correspond pas toujours à la plus petite distance entre descripteurs, mais sou-

vent à l'une des plus petites. La recherche des k plus proches descripteurs (*knn-matching*) peut ainsi augmenter le nombre de « vraies » correspondances.

Si N_{kpt} et k sont trop petits, le nombre de « vraies » correspondances ne sera pas significatif ; et inversement, si N_{kpt} et k sont trop grands, le nombre de fausses correspondances et le temps de calcul augmenteront. Selon nos tests, $N_{kpt} \approx 1000$ et $k \approx 10$ s'adaptent bien à nos données (base de données *USA_Grading*, image de taille 500×500 pixels). La Figure 5.5 montre une comparaison de différents paramètres.

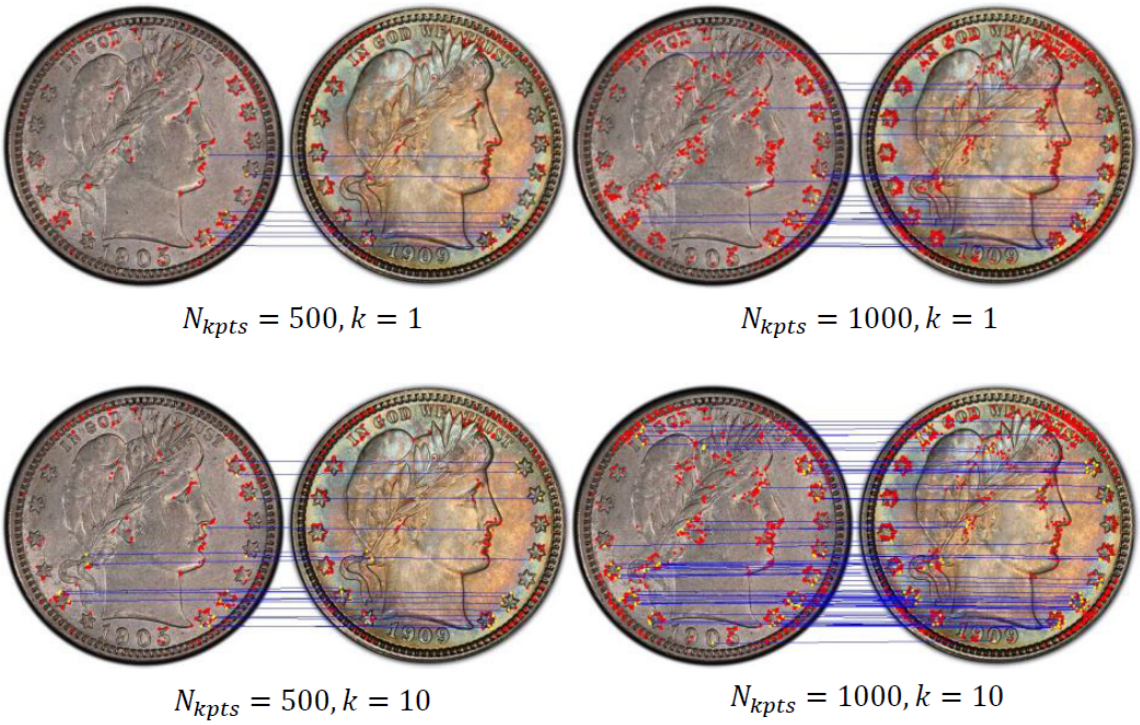


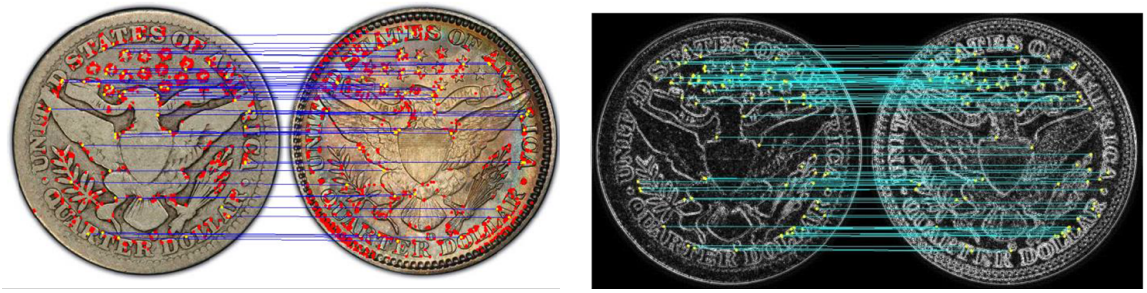
Figure 5.5 – Comparaison entre le nombre de « vraies » correspondances selon différents paramètres

- Carte des gradients

La carte des gradients est obtenue en utilisant l'opérateur Laplacien qui mesure la rapidité des changements d'intensité dans un voisinage et renforce les contours des motifs en réduisant les caractéristiques locales qui ne nous intéressent pas. Bien que les points d'intérêt ORB détectés se concentrent sur les contours des motifs dans l'image originale, d'après nos tests, les « vraies » correspondances retenues sont généralement plus nombreuses sur la carte des gradients que sur l'image originale (*cf.* Figure 5.6).

Pour minimiser le nombre de « fausses » correspondances, nous proposons ensuite deux contraintes spatiales.

- Contrainte de voisinage



(a) (b)

FIGURE 5.6 – « Vraies » correspondances obtenues entre deux monnaies en utilisant les mêmes paramètres (a) sur les images originales et (b) sur les cartes des gradients associées

Il s'agit d'une contrainte en fonction des positions relatives de deux points d'intérêt mis en correspondance. Nous appliquons cette contrainte selon les hypothèses suivantes :

- pas de variation d'échelle des caractéristiques locales appariées puisque les images d'entrée sont des monnaies segmentées de taille normalisée ;

- pas de translation ou déformation des caractéristiques locales appariées puisque les motifs du même type monétaire sont censés être rigides ;

- les orientations des monnaies sont quasi similaires dans les photos professionnelles.

Ainsi, les deux points d'intérêt v_i^Q, v_j^R en correspondance devraient se trouver dans des coordonnées proches sur les deux images. La contrainte de voisinage est donc donnée par :

$$C_n : \|p_i^Q - p_j^R\|_2 < r_n \tag{5.2}$$

où r_n est un rayon de voisinage. Pour avoir des calculs peu coûteux, seuls des petits voisinages sont considérés. La contrainte de voisinage est illustrée à la Figure 5.7. À noter que la performance de détection des points en correspondance en appliquant cette contrainte chute si les orientations des monnaies sont trop différentes, car la troisième hypothèse n'est plus vraie.

- Contrainte de bordure

La bordure d'une monnaie, souvent riche en petits motifs répétitifs, n'est pas toujours discriminant. Ce sont plutôt les motifs de l'intérieur qui déterminent le type monétaire.

Nous imposons donc une contrainte sur la position absolue des points d'intérêt détectés. Cette contrainte est donnée par,

$$C_b : \|p_i - O\|_2 < 0,95R, \tag{5.3}$$

où O est le centre de la monnaie et R est son rayon. Dans le cas des monnaies segmentées de taille normalisée, O et R sont constantes. La contrainte de bordure est illustrée à la Figure 5.8.

Grâce à ces deux contraintes spatiales, nous pouvons éliminer la plupart des « fausses » correspondances ou des correspondances peu intéressantes. Par ailleurs, le temps de calcul est réduit de manière significative car nous n'avons pas besoin de calculer les distances entre toutes les paires de descripteurs.

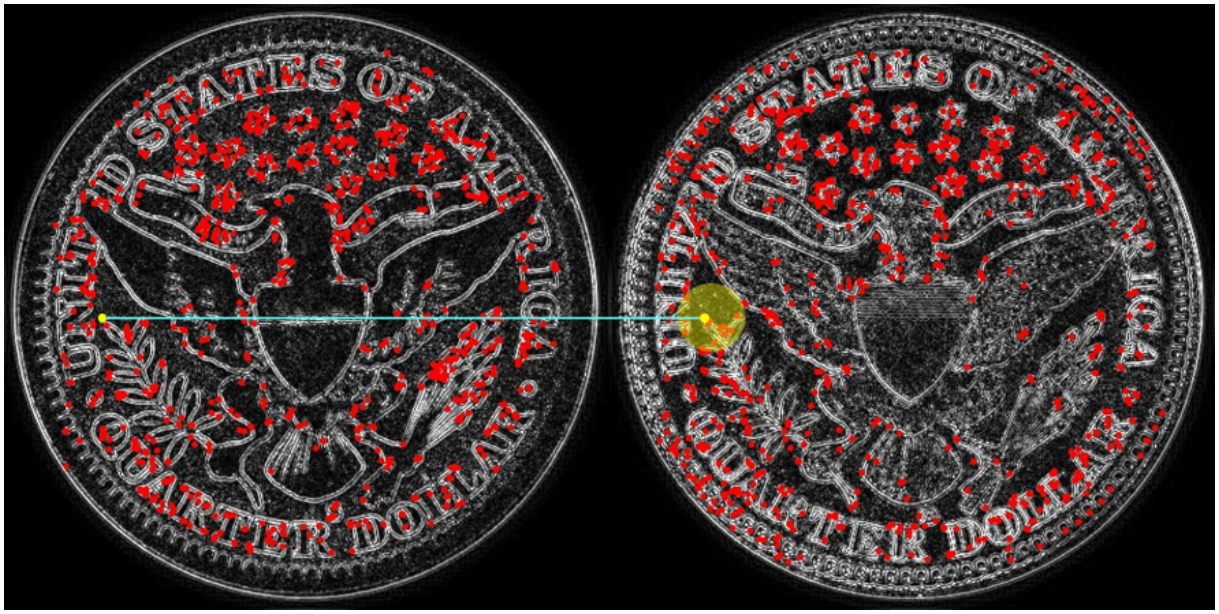


Figure 5.7 – Contrainte de voisinage

Figure 5.8 – Contrainte de bordure

Score de similarité basé sur la correspondance locale

Les correspondances entre points d'intérêt (*cf.* Figure 5.10) sont données par,

$$C_V^{G^Q \leftrightarrow G^R} = \{v_i^Q \leftrightarrow v_j^R | \|d_i^Q - d_j^R\|_2 < d_{th}, C_n, C_b\}, i, j \in [1, N_{kpt}]. \tag{5.4}$$

Le nombre de correspondances est le premier score de similarité que nous proposons,

$$|C_V^{G^Q \leftrightarrow G^R}|. \tag{5.5}$$

Ce score de similarité, basé les correspondances locales, est assez stable pour des photos professionnelles. Une monnaie requête du même type monétaire, quelle soit neuve ou usée, devrait avoir un score plus élevé qu'une monnaie d'un autre type monétaire. Nous observons que ce score peut être influencé par le grade des monnaies. Évidemment, une monnaie requête de même grade que la référence a probablement un score plus élevé qu'une référence usée. Ainsi pour avoir une bonne performance avec ce score de similarité, la base de références devrait comporter des monnaies de même grade. Par ailleurs, certaines monnaies pourraient donner un score élevé sans pour autant être similaires. Par exemple, des monnaies de différents types monétaires peuvent avoir de nombreuses caractéristiques locales similaires, et engendrer un *SSN* plus grand que nous pourrions le penser (*cf.* Figure 5.9).

(a) (b)

FIGURE 5.9 – Score de similarité non discriminant dans certains cas : (a) deux monnaies appartenant à des types monétaires différents ; (b) deux monnaies appartenant au même type monétaire

5.1.3 Score de similarité semi-global basé graphe

Pour répondre aux contraintes du premier score de similarité, nous proposons un deuxième score de similarité, que nous allons détailler, qui vise à traiter des monnaies globalement différentes mais avec beaucoup de motifs locaux correspondants. Ce nouveau score de similarité semi-global est calculé à l'aide de graphes construits à partir des points d'intérêt en correspondance.

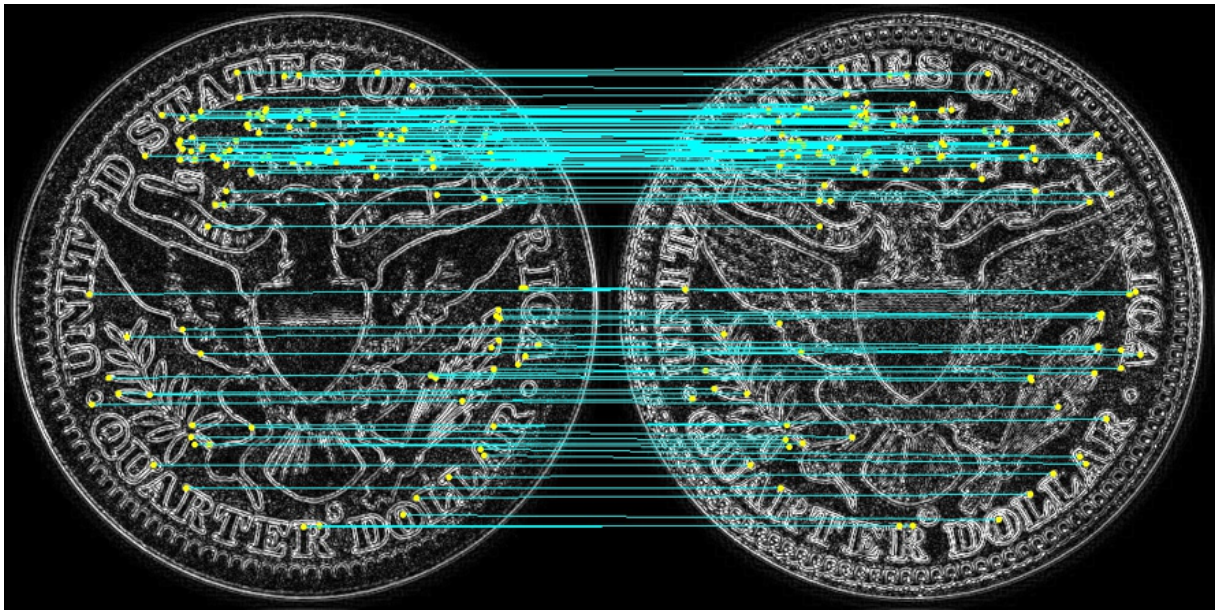


Figure 5.10 – Points d'intérêts ORB retenus

.

Construction des graphes

Sur chaque image, nous construisons un graphe $G(V, E)$ en utilisant les points d'intérêt retenus comme sommets, $v_i \in V$ et les segments entre v_i et v_j comme arêtes, $e_{ij} = (v_i, v_j), e_{ij} \in E$, de longueur $l_{ij} = \|v_i - v_j\|_2$. Le nombre de sommets n_V est toujours égal au score de similarité que nous avons proposé, mais le nombre d'arêtes n_E dépend de n_V et de la façon de construction du graphe. Sachant que le graphe complet avec $n_E = \dfrac{n_V(n_V - 1)}{2}$ arêtes serait trop lourd à gérer en temps de calcul, nous avons choisi de construire le graphe en utilisant les triangulations de Delaunay à partir des points d'intérêt retenus,afin de n'avoir que $n_E = 2n_V - n_B - 2$ arêtes, où n_B est le nombre de sommets dans l'enveloppe convexe du graphe.

Nous construisons d'abord le graphe G^R sur la monnaie de référence I^R à l'aide de la triangulations de Delaunay. Sachant que dans une paire de points d'intérêt en correspondance les points ne sont pas exactement aux mêmes endroits, si le graphe G^Q était construit de la même manière sur la monnaie requête I^Q, G^Q sera probablement bien différent de G^R. Afin d'obtenir deux graphes quasi-identiques dans ce cas, nous construisons toutes les arêtes $e^Q_{ij} = (v^Q_i, v^Q_j)$ tel que $v^Q_i \leftrightarrow v^R_n$ et $v^Q_j \leftrightarrow v^R_m$. Cette astuce permet de construire G^Q avec toutes ses arêtes intrinsèquement appariées à G^R, notées $e^Q_{ij} \leftrightarrow e^R_{nm}$.

Ainsi, G^Q et G^R sont deux graphes avec le même nombre de sommets, le même nombre d'arêtes et des structures quasi-identiques (*cf.* Figure 5.11). Les sommets et les arêtes de deux graphes sont intrinsèquement en correspondance un à un. La seule différence entre les deux graphes est le fait que les sommets correspondants ne sont pas aux mêmes coordonnées. Nous n'avons donc pas besoin de faire un appariement de graphes *graph matching* classique dont la complexité est généralement élevée.

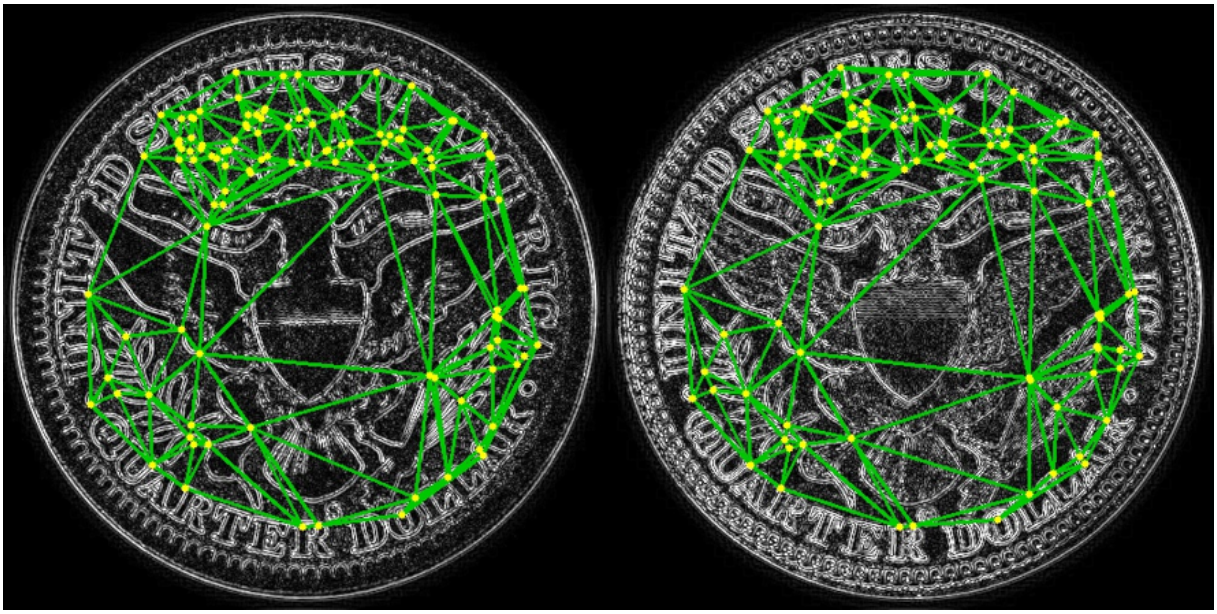


FIGURE 5.11 – Construction de deux graphes quasi-identiques à l'aide des triangulations de Delaunay

.

Recherche d'arêtes similaires correspondants

Notre hypothèse est que, sur des monnaies de même type monétaire, les graphes quasi-identiques devraient couvrir des motifs similaires ; par contre, pour des types monétaires différents, les graphes couvriront des motifs différents. Sachant que les sommets des deux graphes encodent déjà des motifs locaux censés être similaires, l'objectif est ainsi de quantifier la similarité des motifs traversés par les arêtes correspondantes.

Inspiré du *strip-based descriptor* de *D-nets* proposé dans le cadre d'une approche *graph matching* (Hundelshausen et Sukthankar 2012), nous définissons un nouveau descripteur basé sur la séquence d'intensité le long d'une arête. Pour une arête e_{ij} de longueur l_{ij}, nous la divisions en s sections et calculons l'intensité moyenne de chaque section. Puis, nous normalisons ces s moyennes entre 0 et 1. Si la valeur maximale et la valeur minimale sont très proches, nous considérons que la séquence d'intensité est quasi-homogène et donc remplaçons toutes les moyennes par 0,5. Ce descripteur d'arête f_{ij}, sous forme d'un vecteur composé de s éléments normalisés, encode la transition d'intensité le long de l'arête $e_{ij} = (v_i, v_j)$ dans le sens de v_i à v_j. La Figure 5.12 illustre deux descripteurs d'arête avec $s = 3$. L'invariance à la luminosité globale est l'un des avantages de ce descripteur. Par ailleurs, ce descripteur est également invariant à la rotation et à la longueur d'arête.

Trois optimisations ont ensuite été proposées pour faire que ce descripteur d'arête s'adapte au mieux à notre contexte.

- **Épaisseur d'une arête** : au lieu de prendre une arête de largeur un pixel, nous considérons une bande de largeur w_f(w_f=3 ou 5) pixels pour qu'il capture plus d'informations et qu'il soit plus robuste aux bruits (*cf.* Figure 5.13).

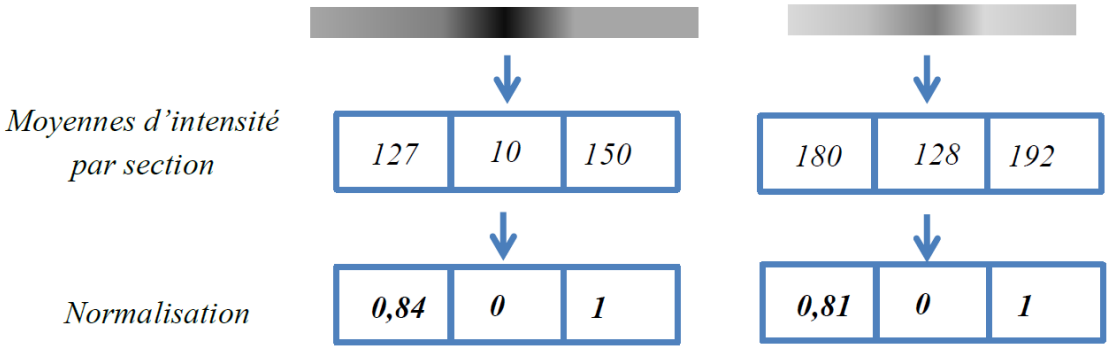


Figure 5.12 – Construction du descripteur d'arête avec $s = 3$

.

- **Arête centrale** : lors du calcul de descripteur f_{ij} de l'arête $e_{ij} = (v_i, v_j)$, nous excluons les deux extrémités et ne considérons que la partie centrale (80%) entre v_i et v_j, de $p_i + 0,1\|p_j - p_i\|_2$ à $p_i + 0,9\|p_j - p_i\|_2$. Cette optimisation permet à f_{ij} qui encode la transition entre les deux sommets v_i et v_j d'être moins influencée par leur position absolue. Par ailleurs, les caractéristiques locales autour des sommets sont déjà prises en compte dans les points d'intérêt retenus.

- **Suppression d'arêtes trop courtes** : nous ne prenons pas une arête si elle est trop courte car il s'agit d'une arête dont les deux extrémités se trouvent dans le même voisinage. La transition entre ces deux sommets est trop locale et n'apporte rien de plus que les caractéristiques locales. Ainsi, les descripteurs d'arête ne sont calculés que sur les arêtes assez longues tel que

$$C_l : l_{ij} > l_{th} \tag{5.6}$$

où l_{th} est un seuil empirique que nous avons fixé dans la pratique.

Deux arêtes correspondantes $e_{ij}^Q \leftrightarrow e_{ij}^R$ sont considérées comme « similaires » si la distance entre leurs descripteurs respectifs est inférieure à un seuil, soit,

$$\|f_{ij}^Q - f_{nm}^R\|_2 < f_{th} \tag{5.7}$$

où f_{th} est un seuil empirique que nous avons fixé dans la pratique.

Score de similarité basé sur les transitions semi-globales

Ainsi, l'ensemble des arêtes similaires correspondantes fournit une description semi-globale qui encode la transition d'intensité entre les motifs locaux en correspondance,

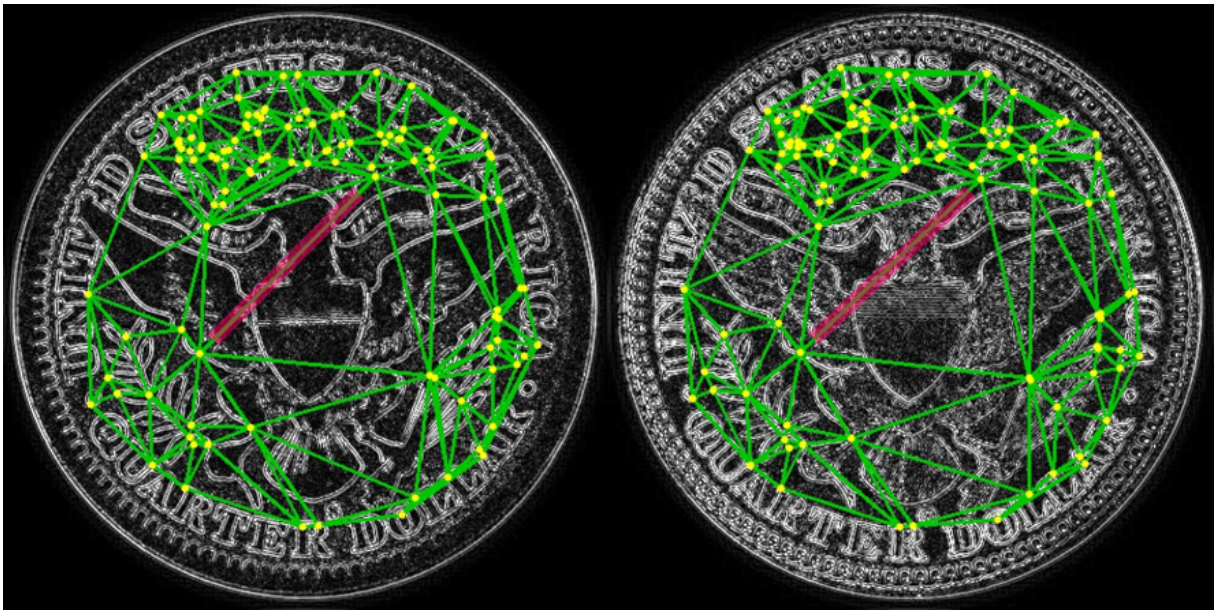


FIGURE 5.13 – Comparaison des arêtes correspondantes : bandes centrales (en rouge) prises en compte

donné par,

$$S_E^{G^Q \leftrightarrow G^R} = \{(e_{ij}^Q, e_{ij}^R)\} | \|f_{ij}^Q - f_{ij}^R\|_2 < f_{th}, C_l\}. \tag{5.8}$$

Le nombre d'arêtes similaires correspondantes est ainsi le second score de similarité que nous proposons, *Similarity Score by graph Edges* (SSE), donné par,

$$SSE = |S_E^{G^Q \leftrightarrow G^R}|. \tag{5.9}$$

Deux monnaies du même type monétaire devraient avoir plus d'arêtes similaires que deux monnaies de types monétaires différents. Sur la Figure 5.14, malgré des grades différents (la monnaie de droite a plus de détails que celle de gauche), la plupart des arêtes correspondantes suffisamment longues sont considérées comme similaires par notre méthode.

Le premier score de similarité quant à lui nous le nommons *Similarity Score by graph Nodes* (SSN), puisqu'il ne prend en compte que les sommets des graphes. En revanche, des monnaies de types monétaires différents n'ont souvent pas beaucoup de points d'intérêt en correspondance et donc peu d'arêtes dans leurs graphes, donc SSE renforce le résultat obtenu par SSN. Par contre, si deux monnaies ont un SSN élevé mais un SSE petit, nous pouvons dire que ces deux monnaies partagent certaines caractéristiques locales mais les motifs intérieurs sont globalement différents, donc elles n'appartiennent pas au même type monétaire.

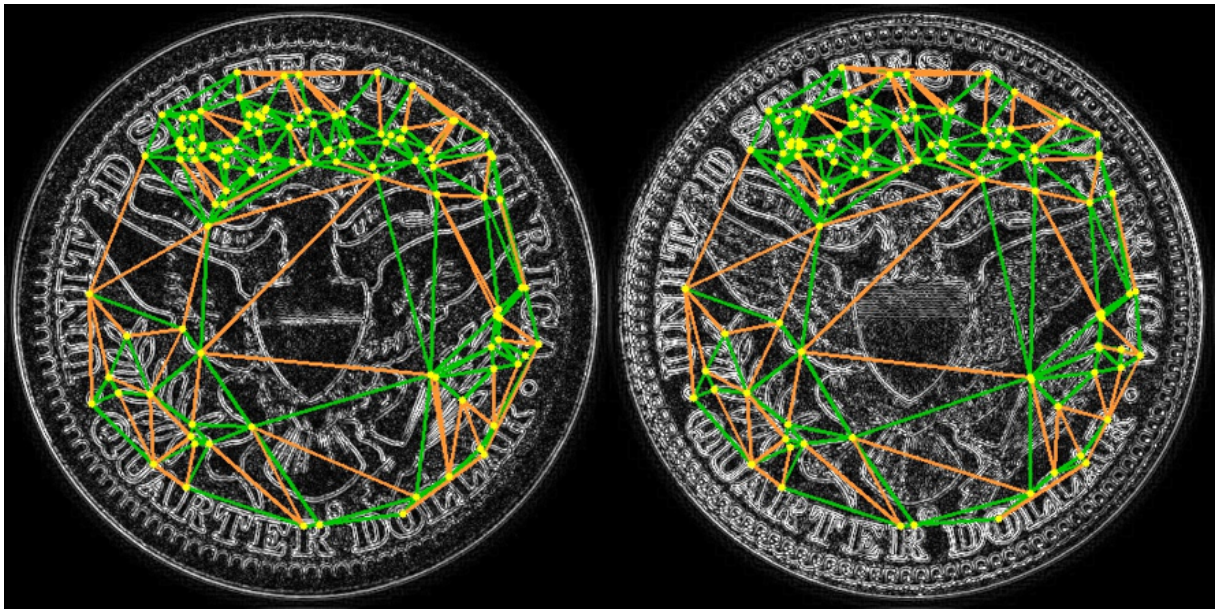


Figure 5.14 – Arêtes similaires marquées en orange (arêtes courtes non considérées)

.

5.2 Expériences et discussion

Pour tester notre méthode, nous avons utilisé la base de données *USA_grading* que nous avons présentée dans le chapitre 3. Pour rappel, les principales caractéristiques de cette base sont :

- un nombre d'images et un nombre de classes importants,
- un nombre de monnaies par type monétaire quasi-égal (une dizaine d'images),
- une variabilité intraclasse de l'apparence des monnaies en fonction de leur grade,
- une similarité interclasse élevée pour certains types monétaires.

Il faut noter toutefois que l'éclairage et donc les reflets sont parfois variables sur certains types monétaires. Ce sont des cas qui ne vérifient pas l'hypothèse que nous avons faite pour notre méthode.

Toutes les images de la base *USA_Grading* sont de taille normalisée 500 × 500 pixels quelle que soit la taille réelle de la monnaie, sachant que la taille de la monnaie segmentée est proportionnelle à sa taille réelle. Le diamètre de la monnaie est l'une des caractéristiques de forme extraites dans l'étape de segmentation pour présélectionner les références. Dans les expériences suivantes, nous utilisons donc les diamètres réels des monnaies concernées comme connaissance *a priori*.

5.2.1 Expériences sur une sous-base de *USA_Grading*

Nous avons d'abord construit une base de test pour fixer les paramètres. Cette base nommée *USA_Grading2* est une sous-base de la base *USA_Grading* qui contient 550 images réparties en 30 classes différentes (une classe est une face d'un type monétaire en

terme numismatique). Les paramètres testés sont comme suit :

- Images d'intensité vs. cartes des gradients

Dans les études de la littérature concernant les monnaies anciennes, les descripteurs locaux sont en général calculés sur l'image originale. Cependant, nous pensons que la carte des gradients aide à mieux cibler les points d'intérêt sur la monnaie que l'image originale. Nous comparons donc les scores de similarité calculés respectivement sur les images d'intensité et les cartes des gradients.

- Une référence vs. deux références

Nous utilisons les monnaies de grade 35 comme références pour qu'elles soient aussi proche des monnaies usées que les monnaies bien conservées. Par ailleurs, dans cette base de données, les monnaies de grade « moyen » contiennent tous les motifs principaux mais moins de reflets que les monnaies de grade élevé. Elles semblent donc appropriées comme références.

Cependant, nous souhaitons tester également notre méthode en considérant deux références par type monétaire (la monnaie de grade 35 et celle de grade 60) pour diminuer le biais d'une seule référence. Dans ce cas-là, les scores de similarité sont calculés deux fois en comparant avec ces deux références, et nous comparons les moyennes des deux calculs.

Les résultats sur la base *USA_Grading2* sont donnés dans la Table 5.1. Parmi tous les tests, le Test 4 (carte des gradients, deux références par classe) a obtenu le meilleur résultat, 100% pour SSN et 99,5% pour SSE. Selon les résultats, la carte des gradients est plus robuste que l'image originale, et que l'utilisation de deux références diminue le biais d'une seule référence. Il est dommage que SSE conçu pour améliorer la performance de SSN paraisse moins performant sur cette base. Bien que SSE qui se base sur SSN soit théoriquement plus global, les images de *USA_Grading* ne correspondant pas exactement à notre hypothèse : l'éclairage inhomogène et les reflets variables dans certaines monnaies ne sont pas des conditions favorables pour cette mesure.

Table 5.1 – Récapitulatif des divers tests faits sur la base *USA_Grading2*

Test	Img. d'entrée	N.img	N.ref par classe	SSE(%)	SSN(%)
1	Img. org.	550	1	94,9	92,9
2	Gradient	550	1	99,3	98,9
3	Img.org	520	2	97,6	97,1
4	Gradient	520	2	100	99,5

5.2.2 Expériences sur USA_Grading

La base de monnaies requêtes contient 2450 images appartenant à 148 classes différentes. La base de références contient 148 monnaies de grade 35 correspondant à chaque

classe. Selon les résultats obtenus sur la sous-base, les scores de similarité ont été calculés sur les cartes des gradients. Notre méthode a été développée dans l'optique d'obtenir des résultats corrects sans avoir beaucoup de données pour chaque type monétaire. Bien que l'utilisation plus de références soit évidemment plus robuste, dans cette expérience nous utilisons une seule référence de grade 35 par classe, ce qui est plus réaliste.

Nous avons également évalué la performance de notre méthode sur trois sous-bases de *USA_Grading* en fonction du matériau des monnaies : monnaies en nickel (gris), monnaies en or (jaunes), monnaies en cuivre (brunes). En général, la couleur dominante de la monnaie peut être une caractéristique pour présélectionner les références. Mais la couleur en provenance du matériau peut parfois être modifiée par la dégradation ou la patine. Ainsi, l'objectif d'utiliser trois sous-bases est de valider la stabilité des algorithmes.

Les résultats sont donnés dans la Table 5.2. Le taux d'identification le plus élevé est de 94,4% sur la base totale, et les résultats sur trois sous-bases sont stables (*cf.* Table 5.2). Les deux scores de similarité permettent des taux d'identification élevés, et SSE est légèrement meilleur que SSN sur la base totale mais pas forcément sur toutes les sous-bases. En considérant les k-plus proches voisins sur la base totale, le bon type monétaire est parmi les trois premiers à plus de 99%, et dans les cinq premiers avec SSE à 100% (*cf.* Figure 5.15). SSE est donc plus discriminant que SSN.

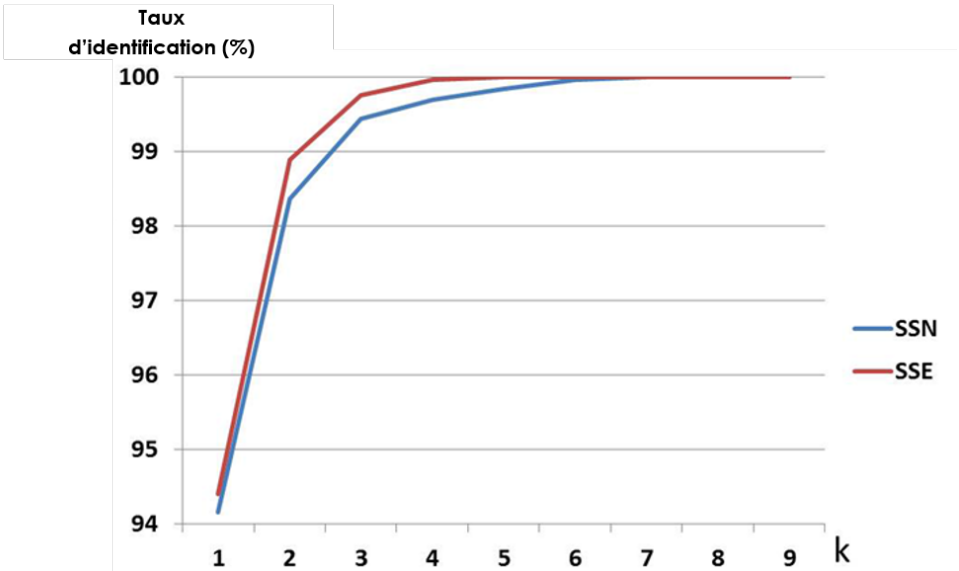


Figure 5.15 – Résultats de classification des k-plus proches voisins sur la base *USA_Grading*

- Rotation des monnaies

La Figure 5.16 présente les courbes de tolérance à la rotation entre la monnaie requête et la monnaie de référence, en affichant les k-plus proches voisins. Les deux scores de similarité proposés ne sont pas invariants à la rotation à cause de la contrainte de voisinage imposée sur SSN. Cependant, les monnaies sont supposées être bien orientées dans les

TABLE 5.2 – Résultats sur différentes sous-bases de *USA_Grading*

Tests	N.img	N.classe	SSN(%)	SSE(%)
Nickel (grise)	1512	88	94,3	95
Or (jaune)	676	44	93,8	93,5
Cuivre (brune)	262	16	94,3	95,4
Total	1450	148	94,2	94,4

photos professionnelles. La petite différence de rotation introduite par l'opérateur est cependant acceptable au vu des résultats dans la Figure 5.16. Dans la pratique, ce n'est pas nécessaire d'augmenter la tolérance à la rotation avec le coût de taux d'identification ou du temps de calcul.

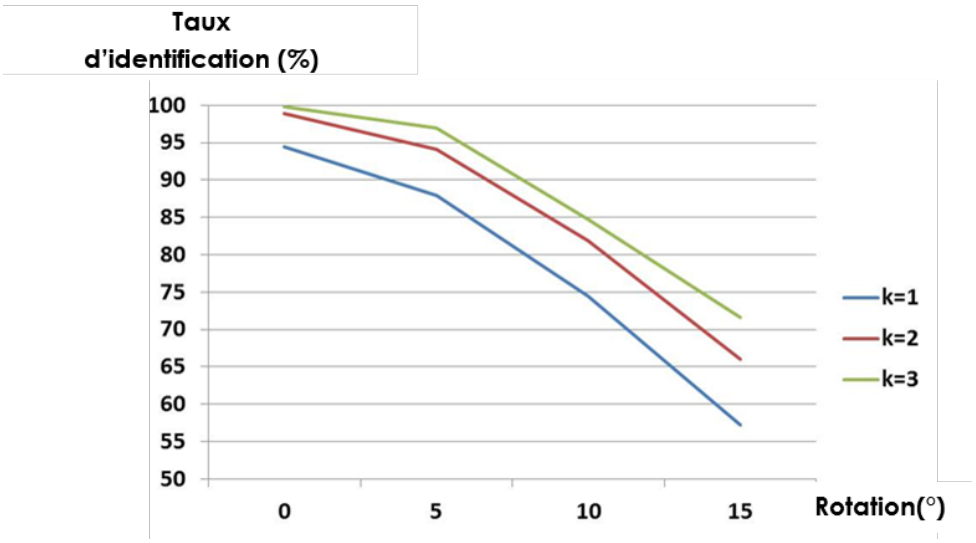


FIGURE 5.16 – Influence de la rotation des monnaies sur le taux d'identification en utilisant SSE

- Types monétaires avec une similarité interclasse élevée

Certaines monnaies de classes différentes sont très ressemblantes et contribuent donc souvent aux fausses identifications. Ce sont des monnaies appartenant à des types monétaires différents mais de même taille et avec des motifs similaires, voire identiques. Par simplicité, nous les appelons « classes ambiguës » (*cf.* Figure 5.17).

Il faut noter que nous utilisons les étiquettes de PCGS comme les vérités terrains. Les sous-types identifiés dont les motifs ne sont pas exactement identiques mais qui ont le même nom de type monétaire sont considérés comme appartenant à une seule classe. Autrement dit, nous ne distinguons que ceux qui sont étiquetées différemment.

Pour la table 5.3 il est très intéressant de noter que SSE a obtenu de bien meilleurs résultats que SSN sur certaines « classes ambiguës ». Cependant, les monnaies bien identifiées par SSN mais pour lesquelles SSE donne un résultat erroné ne se concentrent pas sur certaines classes spécifiques.

Figure 5.17 – Exemples de monnaies issues des classes ambiguës dont les nuances sont marquées en rouge : *Barber10.rev* et *Seated10.rev*

Table 5.3 – Extraits de résultats sur des « classes ambiguës »

Bonne classe	**Classe similaire**	**N.img**	**SSN(%)**	**SSE(%)**
Barber10.rev	Seated10.rev	18	77,8	94,4
Draped25SE.obv	Draped25LE.obv	15	33,3	66,7
Draped50SE.obv	Draped50LE.obv	12	8,3	50
DrapedHalfDimeLE.obv	DrapedHalfDimeSE.obv	12	50	91,7

Les monnaies appartenant aux « classes ambiguës » devraient avoir un SSE plus élevé que SSN, car l'appariement des caractéristiques locales dans un voisinage n'est pas sensible aux nuances éventuelles. Par contre, SSE étant semi-globale, il pourrait mettre en évidence ces nuances dans la mesure où certaines arêtes correspondantes traversent les motifs différents. Il faut noter toutefois que les « classes ambiguës » concernées dans la base de données *USA_Grading* n'ont pas tous SSE > SSN, puisque les scores de similarités dépendent aussi largement de la qualité des images individuelles et que certaines différences sont très localisées.

Concernant le temps de calcul, il dépend principalement de la taille d'image et du nombre de comparaisons. Pour une image de 500 × 500 pixels, le calcul des deux scores de similarité coûte 0,25s par comparaison (sous Mac OS X, 2,7 GHz Intel Core i7, 4G DDR3). Concernant les expériences sur *USA_Grading*, si nous utilisons uniquement le diamètre pour la présélection des références, le nombre de références à comparer sera de 2 à 14 (0,5 à 3,5 secondes pour identifier une monnaie requête). Dans la vraie application de la société GENI, des informations supplémentaires telles que la taille, la forme, le poids, l'avers ou le revers, *etc.*, peuvent largement réduire le nombre de références à comparer.

5.3 Conclusion

Dans l'objectif d'identifier le type monétaire via l'image, nous avons proposé une méthode qui consiste à chercher le type monétaire de référence le plus similaire en fonction de deux scores de similarité. Le premier score est le nombre de points d'intérêt en correspondance qui caractérise la similarité des motifs locaux. Le second score qui correspond au nombre d'arêtes correspondants similaires, prend en compte les similarités semi-globales. En résumé, SSE est une mesure dépendante de SSN, mais il est plus discriminant et plus robuste dans notre contexte, notamment pour certains types monétaires ayant une similarité interclasse élevée. En évitant des algorithmes lourds comme l'appariement de graphes à partir de nuages des points, la méthode proposée est logiquement et mathématiquement simple, avec un temps de calcul acceptable en vue d'une application réelle.

Cependant, SSE reste très sensible au bruit qui modifie la transition d'intensité. Malgré le fait que la base de test *USA_Grading* ne comportait pas que des photos professionnelles qui correspondent parfaitement à nos hypothèses, nous avons quand même obtenu des taux d'identification très encourageants qui devraient être meilleurs avec la future base de la société GENI.

Une limite de notre méthode est la dépendance de la base de références. Nous ne pouvons pas identifier une monnaie requête sans avoir sa référence dans la base de références. À la manière de certains systèmes de la littérature, par exemple *Dagobert* (Nölle et al. 2003), nous pourrons inclure un système de rejet si la monnaie requête n'obtient que des scores de similarité faibles. En effet, des scores de similarité faibles peuvent être non seulement causés par le manque de référence correspondant, mais aussi par une monnaie requête usée. Il faut également noter que nous précisions un biais dans notre cadre de validation : les référence du type monétaire ont été choisies de manière manuelle. La construction d'une base de références non biaisée reste à étudier.

Au vu des images de la base de données *GENI_CoinPhotos* qui correspond bien à nos critères souhaités, nous sommes confiants sur le fait que la méthode proposée fonctionnera très bien sur des photos professionnelles de la société GENI.

Chapitre 6

Reconnaissance des chiffres extraits des monnaies

Dans le cadre de notre thèse, c'est la similarité interclasse élevée qui rend l'identification du type monétaire difficile. En ce qui concerne les monnaies « modernes » de collection que nous étudions, la similarité interclasse élevée est principalement causée par le fait que les auteurs de monnaies créent souvent un nouveau type monétaire en héritant de l'ensemble des motifs de la version précédente, avec seulement des modifications subtiles ou non. Ainsi, l'identification automatique du type monétaire avec une telle similarité interclasse devient un véritable défi pour l'analyse d'images. Pour répondre à cette problématique, soit nous pouvons mesurer les petites nuances parmi les motifs globalement similaires, soit nous pouvons repérer une information évidente et discriminante. Concernant la première piste, nous avons proposé, dans le Chapitre 5, une méthode pour associer à une monnaie requête une monnaie de référence correspondante grâce à des scores de similarité. La performance est bonne de manière générale mais pour certains types monétaires très similaires, cette méthode n'est pas suffisante. Concernant la deuxième piste, nous constatons qu'une légende spécifique, le millésime (date d'émission), peut différencier les types monétaires similaires mais frappés dans des périodes différentes (*cf.* Figure 6.1).

Figure 6.1 – Périodes d'émission différentes de types monétaires similaires : *DrapDolSE.rev* (encadrés en bleu) : 1795-1798 ; *DrapDolLE.rev* (encadré en rouge) : 1798-1803

Un nouveau pipeline prenant en compte la détection du millésime est illustré à la Figure 6.2. Dans ce pipeline, la monnaie segmentée obtenue par la méthode présentée dans le Chapitre 4, est comparée avec les types monétaires de référence. Selon la méthode d'identification du type monétaire que nous avons présentée dans le Chapitre 5, la monnaie requête est parmi les cinq premières références les plus similaires dans 100% des cas. Dans le cas où les deux premières références, ou plus, ont une similarité interclasse élevée comme à la Figure 6.1, nous détectons et reconnaissons le millésime de la monnaie segmentée. La monnaie requête est finalement identifiée par un type monétaire similaire avec la même période de parution connue. La détection du millésime de la monnaie segmentée et le découpage en imagettes de chiffres sont expliqués dans le Chapitre 3.

Dans ce chapitre, nous nous concentrons sur la reconnaissance des imagettes de chiffres extraites des monnaies. Nous appelons de tels caractères, ayant la même couleur que le fond, des caractères en police de « type creusé ». Dans la section 6.1, nous détaillons une méthode de reconnaissance des chiffres qui ne nécessite pas d'apprentissage. En effet,

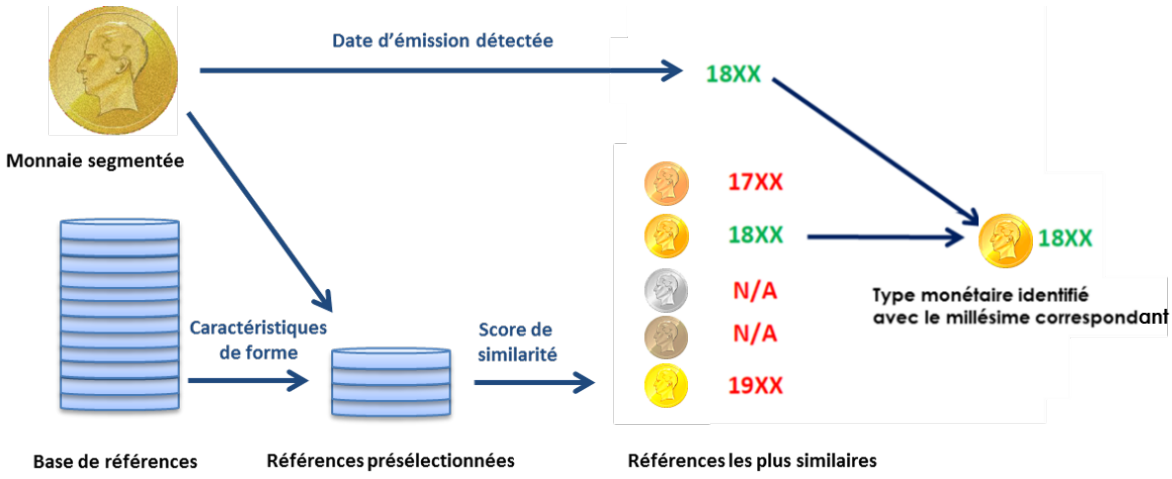


Figure 6.2 – Pipeline de l'identification du type monétaire prenant en compte la détection du millésime

n'ayant pas une base de données d'apprentissage suffisamment importante, nous proposons une solution sans apprentissage. Deux versions différentes de notre méthode sont en réalité proposées pour reconnaître le chiffre à l'aide de sa « topologie » extraite. La section 6.2 montre les expériences conduites sur la base de données *CoinNUMS* que nous avons présentée dans le Chapitre 3. Enfin, la section 6.3 est la conclusion.

6.1 Méthodologie proposée

Les légendes présentent sur les deux faces d'une monnaie sont des caractéristiques informatives pour des numismates. Comme nous l'avons dit dans le Chapitre 2, la plupart des légendes dans notre situation ne sont pas discriminantes. En effet, nous retrouvons très souvent les mêmes textes pour différents types monétaires issus d'un pays donné. Il n'est donc pas utile de tenter de reconnaître toutes les légendes. Par contre, le millésime fournit une information discriminante de la monnaie étudiée. Ainsi, notre objectif est de reconnaître uniquement les chiffres du millésime et non pas les autres caractères.

Contrairement aux travaux de reconnaissance de chiffres dans d'autres domaines, par exemple OCR sur des documents scannés, les chiffres extraits des monnaies présentent une difficulté particulière. En effet, les chiffres sont souvent de la même couleur que le fond de la monnaie, représentés par les contours des reliefs. Par ailleurs, ces contours sont parfois faibles voire quasi inexistants à cause de l'usure. Ainsi, toutes les méthodes de reconnaissance de caractères de la littérature basées sur une bonne segmentation du premier plan sont difficiles à appliquer à notre contexte.

Au vu de la grande variabilité d'apparence des chiffres (police, taille, inclinaison, *etc.*), il est techniquement difficile de les reconnaître par analyse d'images sans apprentissage adéquat. Cependant, ceci requiert une quantité importante de données pertinentes. Par

ailleurs, les données professionnelles fournies par la société GENI ne sont pas suffisantes pour un apprentissage machine performant. L'objectif dans le contexte considéré est de proposer une solution qui ne nécessite pas d'apprentissage machine afin que la société GENI puisse reconnaître directement le millésime de la monnaie traitée.

6.1.1 « Topologies » des chiffres

Nous présentons d'abord ce que nous appelons la « topologie » des chiffres bien segmentés. Pour une imagette de chiffre bien segmentée, nous appelons « trou » une composante connexe à l'intérieur dans son enveloppe convexe n'appartenant pas au chiffre et ne touchant pas l'enveloppe convexe. Nous nommons « ouverture » une composante connexe à l'intérieur de son enveloppe convexe n'appartenant pas au chiffre et touchant l'enveloppe convexe. Par exemple, un chiffre « 8 » contient deux trous en haut et en bas (*cf.* Figure 6.3 (a)) ; un chiffre « 5 » contient deux ouvertures en haut à droite et en bas à gauche (*cf.* Figure 6.3 (c)).

Nous pouvons caractériser un chiffre par le nombre et, la position des trous et des ouvertures. De manière générale, nous appelons par la suite les trous et les ouvertures « bulles ». Nous appelons carte de bulles l'image binaire contenant l'ensemble des bulles (*cf.* Figure 6.3 (b) et (d)). Il est intuitif de reconnaître certains chiffres bien segmentés par leurs cartes de bulles, par exemple « 8 » et « 5 ». Cependant, dans le cas où le chiffre est extrait d'une monnaie plus ou moins usée, la reconnaissance n'est pas facile. Nous proposons donc de nous appuyer sur la carte de gradients qui contient les contours bruités du chiffre.

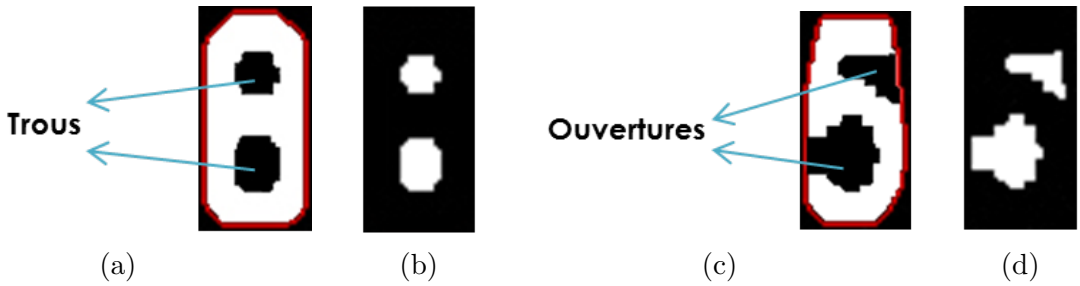


FIGURE 6.3 – Exemples de chiffres binaires et leur enveloppe convexe (rouge) (a) et (c) ; avec leur carte de bulle associée (b) et (d)

Construction de la carte de bulles du chiffre

Étant donnée une imagette de chiffre I à reconnaître, notons I^G sa carte des gradients. Cette carte est obtenue en utilisant l'opérateur Laplacien. Nous appliquons des opérations morphologiques sur I^G pour que les gradients appartenant aux contours du chiffre soient connectés. Puis, nous utilisons la méthode d'Ostu sur cette carte des gradients pour obtenir une image binaire I^{G_B}. Nous extrayons alors la plus grosse composante connexe

de I^{G_B} afin de construire une image binaire I^B ne contenant que cette composante connexe appelée C. La carte de bulles B est donc l'inverse de I^B dans l'enveloppe convexe Γ de C. Le processus pour obtenir la carte de bulle à partir de l'imagette de chiffre originale est illustré à la Figure 6.4

Si C correspond à un chiffre bien segmenté, la carte de bulles ne contient que des bulles représentatives du chiffre. Par opposition, lorsque le chiffre n'est pas bien segmenté, des bulles « parasites » peuvent survenir. Nous introduisons donc le vocabulaire suivant :

- **Bulles principales** : celles qui caractérisent le chiffre bien segmenté ;

- **Bulles parasites** : celles qui sont dues à des bruits ou une police particulière.

Nous supposons que les bulles principales d'un chiffre bruité se trouvent à des endroits similaires à ceux d'un chiffre bien segmenté de la même classe. Ainsi, l'objectif suivant est de filtrer les bulles parasites.

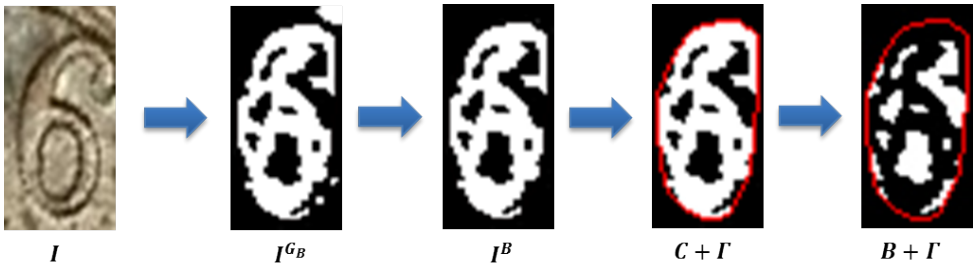


Figure 6.4 – Construction de la carte de bulles

Caractérisation des bulles

Pour vérifier si une bulle $b \in B$ est principale ou parasite, nous la caractérisons de manière suivante :

- τ_a : rapport entre la surface de b et celle dans l'enveloppe convexe Γ.
- τ_l : rapport entre la longueur du bord de b qui touche Γ et le périmètre de Γ.
- θ : angle du centre de b par rapport au centre de la région dans l'enveloppe convexe Γ.
- r_1 : ratio d'aspect de la boîte englobante ϕ de la bulle.
- r_2 : rapport entre la surface de b et celle de sa boîte englobante ϕ ;
- r_3 : rapport entre la surface de b et son enveloppe convexe γ ;
- r_4 : rapport entre la longueur entre le bord de b qui touche Γ et la racine carrée de sa surface.

Parmi ces caractéristiques, τ_a, τ_l et θ sont des caractéristiques décrivant la taille, la position et le type (trou ou ouverture) de la bulle, alors que $r1$, $r2$, $r3$ et $r4$ sont celles décrivant la forme de la bulle.

Filtrage des bulles parasites

Après avoir caractérisé toutes les bulles, nous les filtrons selon les trois hypothèses suivantes :

- le nombre de bulles principales est limité ($|B| \leq 4$) ;

- les bulles principales sont régulièrement convexes ($r_3 \approx 1$), sauf celle dans « 3 » ;

- dans la plupart des cas réels, les bulles principales sont plus grandes que les bulles parasites.

La dernière hypothèse provient des bruits « internes » aux chiffres (*cf.* Chapitre 3) et de la façon dont nous traitons la carte des gradients. Par contre, si la monnaie est neuve et polie (par exemple, les monnaies « belles épreuves »), la carte de bulles d'un chiffre extrait ne contiendra probablement qu'une bulle parasite. Il s'agit d'un trou particulier dû à la police de « type creusé ». Dans ce cas particulier, cette bulle parasite pourrait être considérée comme un chiffre segmenté « amaigri », et elle devrait avoir $\tau_l \approx 0, \theta \approx 0$, r_2 et r_3 assez petits. Par ailleurs, la surface de l'enveloppe convexe de cette bulle devrait être proche de celle dans l'enveloppe convexe Γ. Il faut mentionner cependant que ce cas est rare pour les vraies données bruitées. Dans la pratique, il suffit de vérifier les caractéristiques de la bulle dont la surface dans γ_i est la plus importante, mais nous ne le prenons pas en compte dans le reste de la méthode. La méthode abordée ici se concentre sur les chiffres bruités dont les premiers plans sont difficiles à extraire.

Basés sur les hypothèses précédentes, nous enlevons d'abord les bulles trop petites (empiriquement, τ_a<2%) et puis sélectionnons les quatre bulles les plus grandes (s'il en reste en assez). À l'aide de τ_l et θ, les bulles retenues sont étiquetées par l'une des sept étiquettes suivantes :

- *CH (Center Hole)*
- *TH (Top Hole)*
- *BH (Bottom Hole)*
- *TLO (Top Left Opening)*
- *TRO (Top Right Opening)*
- *BLO (Bottom Left Opening)*
- *BRO (Bottom Right Opening)*

Cependant, r_1, r_2, r_3 et r_4 caractérisant la forme de la bulle, ils nous permettent de faire un étiquetage plus robuste. Par exemple, une ouverture avec un r_4 très faible sera étiquetée comme un trou. Par ailleurs, si plusieurs bulles ont la même étiquette, nous ne garderons que la bulle dont τ_a est le plus élevé et enlèverons le reste. Enfin, les bulles retenues sont ainsi considérées comme les bulles principales. La carte de bulles retenues B_p est censée contenir les mêmes informations « topologiques » qu'un chiffre bien segmenté.

Pour reconnaître le chiffre à partir de B_p, deux stratégies différentes sont détaillées dans les paragraphes suivants.

6.1.2 Système de votes

Dans cette méthode, chaque bulle étiquetée $b_i \in B_p$ participe individuellement à prédire la classe à l'aide d'un système de votes. La somme pondérée des votes des éléments de B_p pour la classe c est donnée par,

$$V(c) = \sum_{i=1}^{|B_p|} v_c(l_{b_i}) \times w_c(b_i), \tag{6.1}$$

où $v_c(l_{b_i})$ est le vote de la bulle b_i en fonction de son étiquette l_{b_i} (*cf.* Table 6.1). Il s'agit d'une valeur binaire pour déterminer si une bulle principale est possible dans une classe c ou non. $w_c(b_i)$ est le poids associé. La prédiction de la classe est donnée par,

$$c_{predict} = argmax\ V(c), c \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\} \tag{6.2}$$

Le poids $w_c(b_i)$ est un paramètre expérimental dépendant des caractéristiques de forme de la bulle ($r1, r2, r3$ et $r4$). Bien qu'un réglage robuste dépende largement des données, il y a deux conseils généraux à suivre :

- le poids pour une classe, votée par une bulle, supposée contenir une seule bulle devrait être plus important que le poids pour une classe, votée par la même bulle mais supposée contenir plusieurs bulles, par exemple $w_7(b_i) > w_4(b_i)$;

- les bulles de forme irrégulière pèsent plus pour certaines classes concernées, par exemple, $w_1(b_i) > w_2(b_i)$ si b_i est trop mince (r_1 faible), $w_3(b_i) > w_2(b_i)$ si b_i est concave (r_3 faible). En général, nous mettons d'autres bulles de forme régulière avec les poids $w_c(b_i) = 1$.

- Nous ne calculerons pas de poids associé $w_c(b_i)$ si $v_c(l_{b_i}) = 0$.

Si B_p obtenue correspond à la « topologie » du chiffre bien segmenté, la classe sera correctement reconnue avec le plus de votes. Dans le cas où B_p est encore bruitée, par exemple avec une bulle de plus ou de moins, le chiffre sera quand même reconnu si les bulles principales retenues votent plus pour la bonne classe que pour les autres. L'utilisation du système de votes est illustrée à la Figure 6.5

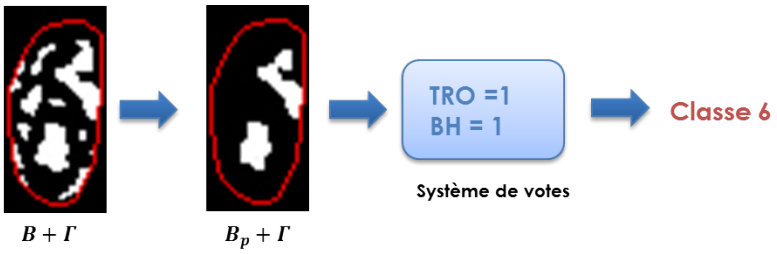


Figure 6.5 – Pipeline de l'utilisation du système de votes

TABLE 6.1 – Vote d'une bulle en fonction de son étiquette

l_b	v_0	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9
CH	1	0	0	0	1	0	0	0	0	0
TH	0	0	0	0	0	0	0	0	1	1
BH	0	0	0	0	0	0	1	0	1	0
TLO	0	1	1	1	0	0	0	0	0	0
TRO	0	1	0	0	1	1	1	0	0	0
BLO	0	1	1	1	1	1	0	1	0	1
BRO	0	1	1	0	1	0	0	0	0	0

6.1.3 Graphes de Reeb

Nous constatons que l'inverse de B_p dans l'enveloppe convexe Γ, notée $\overline{B_p}$, est censée être un chiffre binaire reconstruit. Inspiré du travail de Thome et al. (2011), nous avons essayé d'utiliser le graphe de Reeb pour représenter la « topologie » du chiffre de manière directe. La reconnaissance des chiffres est donc réalisée par la comparaison des graphes de Reeb construits sur les chiffres binaires reconstruits $\overline{B_p}$ et sur les chiffres de référence. Dans les paragraphes suivants, nous détaillons la construction des graphes de Reeb, la méthode proposée pour adapter les graphes aux bruits de nos données spécifiques, et la comparaison des graphes.

Construction des graphes de Reeb

Un graphe de Reeb $G(V, E)$ est généré par une approche incrémentale directionnelle. Chaque sommet $v_i \in V$ est étiqueté selon sa nature *begin*, *end*, *merge* et *split*, avec l'étiquette $l_{v_i} \in \{b, e, m, s\}$. Nous scannons l'image binaire qui contient le chiffre de gauche à droite colonne par colonne. Une composante connexe, ou aussi appelée cellule, à la colonne c est notée $cell_k^{(c)} (0 < k < n^{(c)} - 1)$, où $n^{(c)}$ est le nombre de cellules de cette colonne. Le nombre de cellules adjacentes sur les deux colonnes adjacentes à $cell_k^{(c)}$ est noté respectivement $n(cell_k^{(c)})^{(c-1)}$ et $n(cell_k^{(c)})^{(c+1)}$.

Les sommets du graphe seront générés et étiquetés au centre de $cell_k^{(c)}$ selon les critères suivants :

- $n(cell_k^{(c)})^{(c-1)} = 0$, créer un $v(b)$ (*begin :* sommet de début) ;
- $n(cell_k^{(c)})^{(c+1)} = 0$, créer un $v(e)$ (*end :* sommet de fin) ;
- $n(cell_k^{(c)})^{(c-1)} > 1$, créer un $v(m)$ (*merge :* sommet de fusion) ;
- $n(cell_k^{(c)})^{(c+1)} > 1$, créer un $v(s)$ (*split :* sommet de division).

Dans le cas où une cellule possède éventuellement deux étiquettes (*merge* et *split*, *begin* et *split*, *merge* et *end*), nous créons deux sommets superposés avec deux étiquettes respectives au même endroit. L'arête $e_{ij} = (v_i, v_j) \in E$ est construite entre deux sommets v_i et v_j appartenant à des cellules adjacentes.

Pour les chiffres bien segmentés avec une police classique, le graphe de Reeb représente bien leur « topologie ». Sur les graphes extraits de dix chiffres bien segmentés (*cf.* Figure 6.6), nous voyons bien une boucle dans les graphes pour « 0 », « 4 », « 6 » et « 9 », et deux boucles pour « 8 » ; par ailleurs, deux sommets adjacents *begin* ou *end* signifient une ouverture avec sa position correspondante.

La Figure 6.7 montre quelques exemples de graphes de Reeb extraits des $\overline{B_p}$ des chiffres originalement en police de « type creusé ». Nous pouvons voir que les graphes générés à partir de chiffres non bruités de la même classe sont similaires. Par contre, le graphe d'un chiffre bruité possède souvent plus de sommets que le nombre attendu.

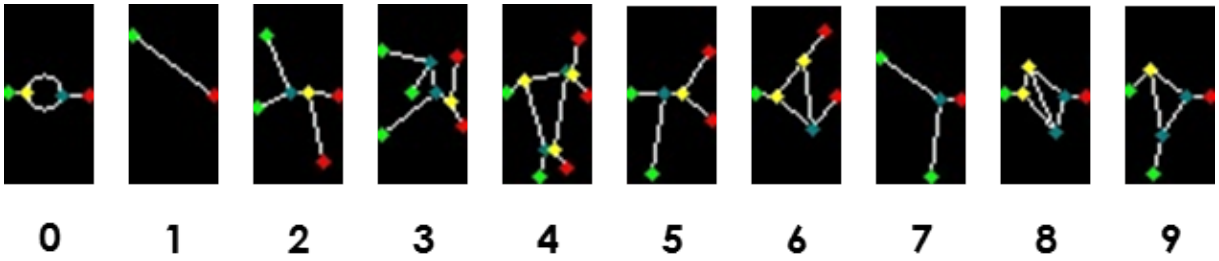
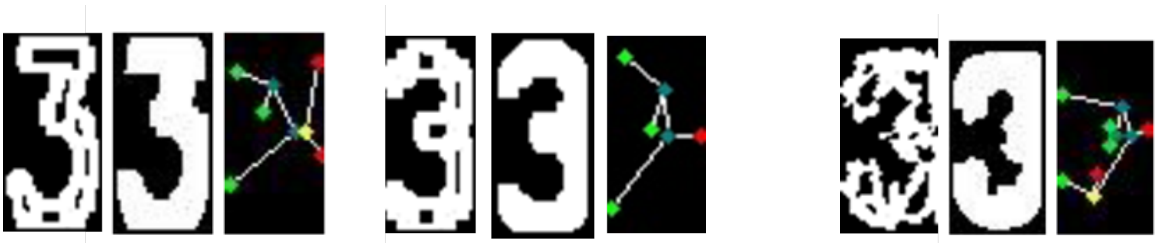


Figure 6.6 – Graphes de Reeb générés de dix chiffres bien segmentés en police courante



(a) (b)

Figure 6.7 – Exemples de graphes de Reeb générés de chiffres originalement en police de « type creusé » : (a) chiffres non bruités (b) chiffre bruité

Topologies simplifiées des chiffres

Nous souhaitons que le graphe de Reeb présente uniquement le nombre et la position relative des bulles. Cependant, les graphes générés à partir des données bruitées sont sensibles aux bruits du bord ainsi qu'à la forme des bulles. Même pour les chiffres bien segmentés, des polices différentes génèrent souvent des graphes différents. Puisqu'un graphe dans notre contexte ne contient pas beaucoup de « bons sommets » en total par rapport aux sommets associés au bruit, l'élagage d'un graphe « bruité » est difficile. L'idée est donc de « lisser » les bords bruités des chiffres binaires avant de construire les graphes.

Pour cela nous ne lissons pas directement l'image mais, extrayons le premier plan synthétique du chiffre avec les bords lisses de manière suivante. Nous construisons une

image binaire, à l'aide de B_p, qui ont les mêmes caractéristiques « topologiques » que $\overline{B_p}$ mais avec des bords lisses. Nous appelons cette image binaire la Topologie Simplifiée (TS) du chiffre à reconnaître, et nous la construisons de manière suivante. Dans l'enveloppe convexe Γ, nous remplaçons un trou de forme quelconque par un disque complet et une ouverture de forme quelconque par un disque partiel (pour l'ouverture concave de « 3 », nous la remplaçons par deux disques partiels). Des exemples sont montrés à la Figure 6.8. Les positions des disques remplacés sont en fonction des positions des bulles originales dans B_p, mais il faut faire attention de ne pas connecter des bulles originalement détachées.

Si B_p ne contient pas de bulle parasite, la TS obtenue à partir de B_p possède le même nombre de trous et d'ouvertures aux mêmes positions que le premier plan effectif du chiffre. Ainsi, nous pouvons générer des graphes de Reeb à partir des TS.

FIGURE 6.8 – Exemples de Topologies Simplifiées (TS) (encadrées en rouge) des chiffres bruités

Ensuite, le graphe de Reeb généré de la TS du chiffre à reconnaître est comparé avec ceux des chiffres de références (bien segmentés) des dix classes, et la classe reconnue sera celle ayant le graphe le plus similaire.

Cependant, d'une part à cause des polices variées, les chiffres bien segmentés de la même classe n'ont pas toujours la même « topologie » (*cf.* Figure 6.9 (a)). D'autre part, les « topologies » ne sont pas toujours discriminantes entre certains chiffres de classes différentes (*cf.* Figure 6.9 (b)).

Dans le cas où la police implique des « topologies » discriminantes, les graphes de Reeb générés à partir des TSs associées sont considérés comme graphes de référence, comme illustrés à la Figure 6.6. Dans la négative, il est plus intéressant de chercher les bulles représentatives dans B_p pour certaines classes, par exemple une ouverture mince dans un chiffre « 1 ». Cette étape ressemble à la pondération des bulles de forme particulière dans la méthode que nous avons proposée précédemment.

Par conséquent, la solution adoptée est de vérifier d'abord si B_p du chiffre à reconnaître contient des bulles particulières, sinon nous extrayons le graphe de Reeb de sa TS et comparons avec les dix graphes de référence. Pour cela, nous mettons expérimentalement une liste de contrôle L_C qui contient les bulles de formes symboliques appartenant aux chiffres considérés. Si $b_i \in B_p$ répond à la liste de contrôle L_C, le chiffre sera reconnu directement comme appartenant à la classe associée à cette bulle représentative ; sinon le chiffre sera reconnu comme appartenant à la classe avec le graphe de référence le plus similaire à l'aide d'une comparaison des graphes. Le pipeline de cette approche est illustré

à la Figure 6.10

(a) (b)

Figure 6.9 – Exemples de polices courantes de chiffres binaires et leur enveloppe convexe (en rouge) : (a) chiffres « 1 » avec des « topologies » différentes ; (b) chiffres appartenant aux différentes classes avec des « topologies » similaires

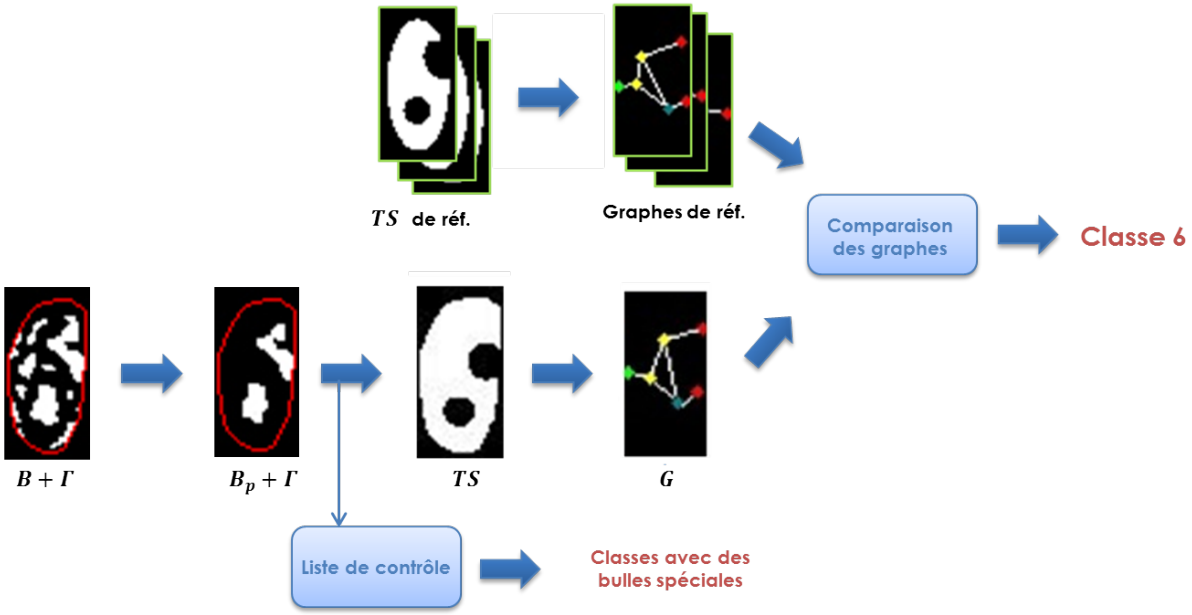


Figure 6.10 – Pipeline de la comparaison des graphes de Reeb

Comparaison des graphes

La similarité entre les deux graphes de Reeb est mesurée par une simple distance. Le calcul de cette distance se base sur l'appariement des sommets et des arêtes des graphes.

Étant donné un graphe requête $G^Q(V^Q, E^Q)$ et un graphe de référence $G^R(V^R, E^R)$, nous faisons d'abord l'appariement des sommets en fonction de leur position et de leur étiquette. La distance de deux sommets $v_i^Q \in V^Q$ et $v_j^R \in V^R$ est donnée par,

$$d_v(v_i^Q, v_j^R) = f_v(l_{v_i^Q}, l_{v_j^R})\sqrt{\alpha(x_i^R - x_j^Q)^2 + \beta(y_i^R - y_j^Q)^2} \tag{6.3}$$

$$f_v(l_{v_i^Q}, l_{v_j^R}) = \begin{cases} 1, si\ l(v_i^Q) = l(v_j^R) \\ \rho, si\ l(v_i^Q) \neq l(v_j^R) \end{cases} \tag{6.4}$$

où α et β sont des poids expérimentaux ; $\rho(\rho > 0)$ est une constante de pénalisation si les sommets ont des étiquettes différentes. Plus $d_v(v_i^Q, v_j^R)$ sera petite, plus les sommets auront de chance d'être appariés, notés $v_i^Q \leftrightarrow v_j^R$.

Selon notre hypothèse, deux graphes représentant des « topologies » similaires devraient avoir plus d'arêtes en correspondance que ceux qui représentent des « topologies » différentes. Pour chaque arête $e_{ij}^Q = (v_i^Q, v_j^Q)$, nous vérifions s'il existe une arête correspondante $e_{mn}^R = (v_m^R, v_n^R)$ définie par deux sommets appariés à v_i^Q, v_j^Q, avec $v_i^Q \leftrightarrow v_m^R$ et $v_j^Q \leftrightarrow v_n^R$. La distance entre graphes d_G est donnée par,

$$d_G = \sum_{1}^{min(n_E^Q, n_E^R)} \frac{f_E(e_{ij}^Q, e_{mn}^R)}{min(n_E^Q, n_E^R)} \tag{6.5}$$

où n_E^Q, n_E^R sont les nombres d'arêtes respectives sur les deux graphes.

$$f_E(e_{ij}^Q, e_{mn}^R) = \begin{cases} 0, si\ l_{v_i^Q} = l_{v_m^R},\ l_{v_j^Q} = l_{v_n^R} \\ 0,5, si\ l_{v_i^Q} = l_{v_m^R},\ l_{v_j^Q} \neq l_{v_n^R} ou\ l_{v_i^Q} \neq l_{v_m^R},\ l_{v_j^Q} = l_{v_n^R} \\ 1, si\ l_{v_i^R} \neq l_{v_m^Q},\ l_{v_j^R} \neq l_{v_n^Q} \end{cases} \tag{6.6}$$

La TS du chiffre à reconnaître ayant une « topologie » identique que celle du chiffre de référence, la distance à son graphe de référence sera $d_G = 0$; sinon il sera associé à celle avec la d_G la plus petite.

6.2 Expériences

Bien que notre méthode soit conçue pour les chiffres extraits des monnaies avec un premier plan difficile à segmenter, elle permet de reconnaître également des chiffres bien segmentés. Ainsi, pour la tester sur les caractères avec un premier plan facile à détecter, nous utilisons la base de données *Chars74K* (*cf.* Figure 6.11) qui contient deux sous-bases de chiffres. Pour tester notre méthode dans des conditions idéales, nous avons également fait une base de test en simulant la police de « type creusé » non bruitée.

L'apprentissage automatique à l'aide des réseaux de neurones (ANN) est normalement performant pour la reconnaissance de caractères. Même si l'objectif était de proposer une méthode qui ne nécessite pas d'apprentissage machine par manque de données dans notre contexte, nous avons également testé un réseau de neurones avec une structure basique de trois couches sur les bases de données considérées. Toutes les images d'entrée sont redimensionnées en 32×32 pixels. Nous avons ainsi 1024 neurones d'entrée correspondant à chaque pixel et 10 neurones de sortie correspondant aux classes. Le nombre de neurones de la couche cachée est fixé expérimentalement.

6.2.1 Bases de données

La base de données *CoinNUMS* a été déjà présentée dans le Chapitre 3. Cette base de données contient des chiffres plus ou moins bruités extraits des monnaies réelles. Plus précisément, CoinNUMS_geni contient les imagettes de chiffres moins bruitées manuellement extraites de monnaies bien conservées ; CoinNUMS_geni_m contient des imagettes de chiffres manuellement extraites de monnaies plus ou moins usées ; CoinNUMS_geni_a contient des imagettes de chiffres encore plus bruitées automatiquement extraites de monnaies plus ou moins usées, le découpage des chiffres n'étant pas parfait.

Concernant la base de données *Chars74K* (*cf.* Figure 6.11), les chiffres sont faciles à segmenter ou déjà segmentés. La sous-base *Chars74K_EnglishImg* contient 593 images de chiffres issus de scènes naturelles, avec une grande variabilité en couleur et en police. Cependant, la distinction entre le premier plan et l'arrière-plan est beaucoup plus simple que notre cas. La sous-base *Chars74K_EnglishHnd* contient 550 images binaires de chiffres manuscrits. La difficulté réside ici dans une grande variation de polices manuscrites, qui sont parfois trop « exotiques » pour notre méthode.

La base de test synthétique contient 100 chiffres synthétiques. Il s'agit de chiffres binaires avec des contours nets et connectés. Ce sont également des chiffres que nous souhaitons extraire sur des monnaies bien conservées dans des photos de qualité.

La différence est significative entre les chiffres synthétiques nets et les cartes des gradients binarisées obtenues sur les chiffres réels (*cf.* Figure 6.12). Par ailleurs, certains chiffres extraits de monnaies usées sont extrêmement bruités (*cf.* Figure 6.13 (a)), et leurs cartes des gradients sont souvent à peine reconnaissables à l'œil nu (*cf.* Figure 6.13 (b)).

Figure 6.11 – Exemples de la base de données *Chars74K* : *Chars74K_EnglishImg* (en haut) ; *Chars74K_EnglishHnd* (en bas)

La distribution des données de toutes les bases considérées est listée dans la Table 6.2.

Nous avons testé deux versions de notre méthode basée sur la « topologie » du chiffre sur toutes les bases de données mentionnées, et le réseau de neurones basique sur les bases de données *CoinNUMS* et *Chars74K*. À cause du nombre de données faible pour certaines bases, les bases d'apprentissage utilisées pour entraîner le réseau de neurones ne sont pas très grandes (100 et 200 images). Pour *CoinNUMS_geni*, nous rajoutons des

FIGURE 6.12 – Exemples de chiffres synthétiques testés (en haut) et de cartes des gradients binarisées obtenues sur la base *CoinNUMS_pcgs_a* (en bas)

(a) (b)

FIGURE 6.13 – (a) Chiffres extraits de monnaies usées (b) cartes des gradients binarisées « illisibles »

images synthétiques pour certaines classes car n'ayant pas assez d'exemples.

6.2.2 Résultats et discussions

Sur la Figure 6.14, nous pouvons voir que la méthode proposée à l'aide d'un système de votes a obtenu les meilleurs résultats sur la base de données *CoinNUMS*, et que le ANN a obtenu les meilleurs résultats sur Chars74K. Évidemment, plus nous utilisons de données d'apprentissage dans le réseau de neurones, meilleurs sont les résultats. Les taux de reconnaissance sur la base *CoinNUMS* sont globalement plus faibles que sur la base *Chars74K*. Cependant, notre méthode à l'aide d'un système de votes est légèrement meilleure que le réseau de neurones sur ces données bruitées. Quelle que soit la méthode utilisée, les résultats obtenus sur les données moins bruitées sont généralement meilleurs que sur les données bruitées.

Le taux de reconnaissance le plus élevé est de 92% obtenu sur les chiffres synthétiques

TABLE 6.2 – Distribution de données dans différentes bases de test

Base de données	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Total**
CoinNUMS_geni	46	185	43	4	2	42	43	14	84	143	606
CoinNUMS_pcgs_m	80	331	57	81	24	66	70	149	213	139	1200
CoinNUMS_pcgs_a	80	331	57	81	24	66	70	149	213	139	1200
Chars74K_EnglishImg	105	79	55	50	47	64	64	48	32	34	593
Chars74K_EnglishHnd	55	55	55	55	55	55	55	55	55	55	550
Chiffres synthétiques	10	10	10	10	10	10	10	10	10	10	100

par les deux versions de la méthode proposée basée sur la « topologie ». Cependant, l'approche par comparaison des graphes de Reeb n'est pas performante sur toutes les bases de données réelles, alors que celle utilisant le système de votes est robuste sur d'autres bases de test. En effet, le graphe de Reeb est très sensible aux bruits du bord et aux polices « exotiques », y compris celles manuscrites.

Concernant les réseaux de neurones, nous ne pouvons pas nier le fait qu'avec une quantité importante de données d'apprentissage, nous pourrions probablement avoir un taux de reconnaissance élevé. Cependant, avec le peu de données d'apprentissage que nous avons, le réseau de neurones n'est pas assez robuste à la variabilité interclasse élevée et aux bruits des chiffres extraits des monnaies.

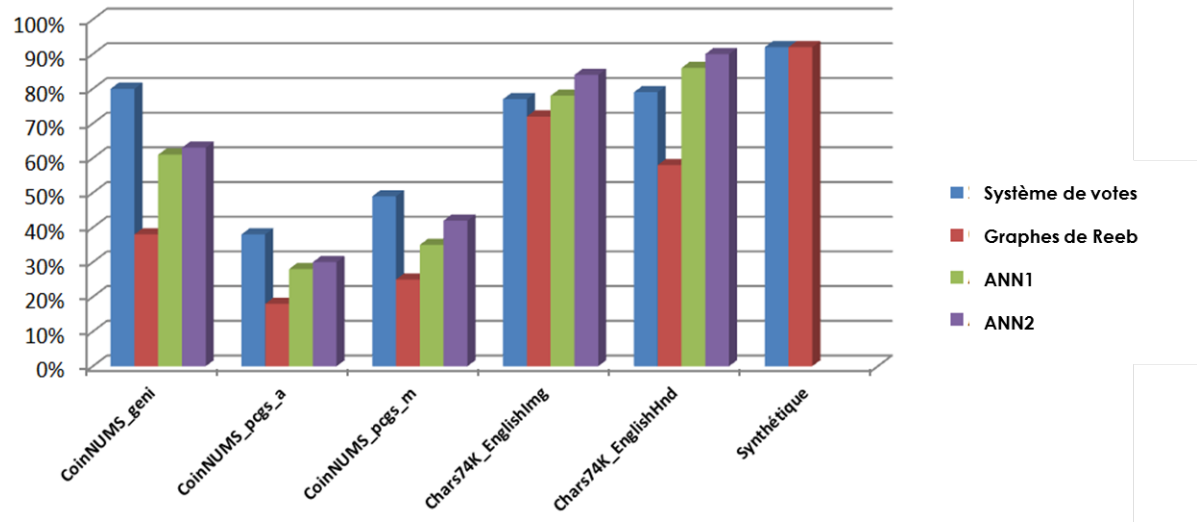


Figure 6.14 – Résultats obtenus sur les différents tests (ANN1 : ANN avec 100 images d'apprentissage ; ANN2 : ANN avec 200 images d'apprentissage)

Nous avons également évalué les résultats à l'aide de matrices de confusion M_c. Dans la matrice de confusion M_c, la valeur à chaque position $M_c(i, j)$ correspond au pourcentage d'imagettes appartenant au chiffre i et reconnu comme le chiffre j. La diagonale de M_c représente les classifications correctes (en vert), et ses valeurs devraient idéalement se rapprocher des 100%. Toutes les valeurs non nulles en dehors de la diagonale représentent donc des mauvaises classifications (en rouge). À l'aide de la matrice de confusion, nous pouvons donc savoir quelles classes sont plus sensibles aux bruits. La Figure 6.15 montre deux matrices de confusion obtenues sur la base synthétique (données non bruitées) et la base *CoinNUMS_pcgs_a* (données bien bruitées) avec la méthode de système de votes.

Pour les chiffres synthétiques non bruités, il est normal qu'un « 1 » ou un « 3 » soit facilement confondu avec un « 7 » s'il possède une ouverture en bas à gauche un peu grosse ou moins concave ; et qu'un « 4 » soit facilement confondu avec un « 9 » lorsque le trou en haut et l'ouverture en bas à gauche sont dominantes.

Pour les données bruitées de *CoinNUMS_pcgs_a*, les chiffres sont plus facilement

confondus. La classe « 3 » est très mal reconnue car la forme concave (r_3) de la bulle clef est sensible aux bruits. Par contre, les autres classes sont souvent reconnues en classe « 2 ». En effet, en considérant les polices courantes dans nos données, les ouvertures de trois positions différentes (*TLO, DLO, DRO*) votent pour la classe « 2 ». Si la carte des gradients est partiellement déconnectée à cause des bruits, il y aura probablement plus d'ouvertures qui votent pour « 2 ».

Système de votes : chiffes synthétiques

	0	1	2	3	4	5	6	7	8	9
0	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%
1	0%	70%	0%	0%	0%	0%	0%	30%	0%	0%
2	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%
3	0%	0%	0%	80%	0%	0%	0%	20%	0%	0%
4	0%	0%	10%	0%	70%	0%	0%	0%	0%	20%
5	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%
6	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%
7	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%
8	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%
9	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%

Système de votes : *CoinNUMS_pcgs_a*

	0	1	2	3	4	5	6	7	8	9
0	61%	3%	8%	0%	2%	0%	17%	0%	3%	5%
1	2%	54%	32%	1%	3%	1%	1%	5%	1%	1%
2	8%	0%	61%	0%	5%	0%	3%	3%	16%	5%
3	5%	3%	45%	3%	3%	13%	5%	3%	18%	3%
4	8%	0%	8%	0%	77%	0%	8%	0%	0%	0%
5	3%	3%	16%	0%	18%	29%	26%	0%	5%	0%
6	0%	0%	29%	0%	11%	0%	50%	0%	11%	0%
7	2%	2%	30%	11%	1%	7%	1%	35%	6%	4%
8	8%	3%	21%	2%	7%	1%	15%	2%	37%	3%
9	6%	3%	29%	2%	5%	8%	3%	2%	9%	31%

FIGURE 6.15 – Matrices de confusion

6.3 Conclusion

Dans l'objectif d'identifier les types monétaires notamment pour ceux qui ont une similarité interclasse importante mais avec des millésimes différents, nous avons proposé une méthode qui consiste à reconnaître les chiffres extraits des monnaies de manière automatique sans avoir besoin d'apprentissage spécifique. Pour traiter les chiffres en police

de « type creusé » avec un premier plan difficile à segmenter, nous les extrayons d'abord sur la carte des gradients binarisée. Ensuite, l'ensemble des trous et des ouvertures dans l'enveloppe convexe du chiffre est représenté dans la carte de bulles. Nous filtrons la carte de bulles en enlevant les petites bulles dues aux bruits et supposons que les bulles retenues représentent les informations « topologiques » du chiffre. Et puis, deux stratégies sont proposées pour reconnaître le chiffre à partir de la carte de bulles filtrée : la première approche se base sur un système de votes qui calcule la somme pondérée de toutes les bulles retenues ; la deuxième est de représenter la « topologie » du chiffre à l'aide d'un graphe de Reeb et de reconnaître le chiffre par la comparaison de graphes. Pour cette approche, le premier plan synthétique est introduit afin de lisser les bords du chiffre car le graphe de Reeb est très sensible aux bruits.

Selon les expériences effectuées, les deux versions de notre méthode ont obtenu des taux de reconnaissance élevées de 92% sur des imagettes de chiffre synthétiques en police de « type creusé » non bruitées. C'est un résultat encourageant puisque la société GENI a pour l'objectif de produire des photos professionnelles pour que nous puissions extraire nettement les contours des légendes. L'approche basée sur un système de votes est beaucoup plus robuste que celle basée sur la comparaison des graphes de Reeb, notamment sur des données bruitées. Par ailleurs, le système de votes permet plus facilement de gérer certaines polices courantes pour lesquelles la « topologie » ne permet pas une discrimination entre certaines classes. non « topologiquement » discriminantes interclasse. Cependant, cette méthode a quand même certains inconvénients suivants que nous allons améliorer dans le futur.

1) La forme effective de la bulle n'est pas considérée dans la méthode proposée. Par exemple, une ouverture triangulaire peut distinguer « 1 » et « 7 », et un trou triangulaire peut distinguer « 4 » et « 9 ». De manière similaire, les contours plats ou courbés du chiffre ne sont pas considérés non plus. Cependant, de telles caractéristiques qui permettent d'optimiser la reconnaissance sont souvent considérablement bruitées sur les données réelles.

2) L'obtention d'une carte des gradients binarisée la plus « idéale » possible est l'étape cruciale pour la suite de notre algorithme. Nous avons proposé des opérations morphologiques pour que la carte des gradients traitée soit exploitable. Si la carte des gradients est déconnectée ou trop bruitée, nous ne pouvons pas extraire la bonne carte de bulles pour analyser la « topologie » du chiffre contenu dans l'imagette.

3) Notre méthode ne peut pas reconnaître des chiffres avec des polices « exotiques » qui ne sont pas prises en compte par l'algorithme. Par exemple, un chiffre « 9 » avec son trou en haut ouvert (*cf.* Figure 6.16 (a)). Cependant, une ouverture trop fine risque d'être fermée dans la carte des gradients binarisée (*cf.* Figure 6.16 (b)).

Notre méthode basée sur la « topologie » du chiffre est robuste à certains types de bruits qui ne modifient pas la « topologie » principale du chiffre, mais par contre elle

(a) (b) (c)

FIGURE 6.16 – Cas difficiles avec leurs cartes des gradients binarisées : (a) chiffre « 9 » avec son trou ouvert ; (b) chiffre « 9 » avec son ouverture « bouchée » ; (c) chiffre « 6 » avec un bout de motif touché

n'est pas développée pour résister aux bruits qui modifient éventuellement la « topologie ». Par exemple, dans le cas où le chiffre est extrait d'une monnaie trop usée, nous ne pouvons pas obtenir une carte de bulles complète. Notre méthode est plutôt performante sur les monnaies neuves ou quasi-neuves, puisqu'il est plus facile d'obtenir des cartes des gradients propres. Concernant les monnaies moins usées, nous utilisons les opérations morphologiques pour que les gradients extérieurs du chiffre soient connectés. Cependant, ce paramétrage est assez délicat car il risque de boucher les bulles originales. Pour obtenir une meilleure performance, il pourrait donc être intéressant d'obtenir une carte des gradients plus robuste.

Il faut mentionner que ce sont plutôt les monnaies bien conservées que les clients de la société GENI souhaitent faire analyser. Les techniques de photographie que la société GENI cherche à mettre en œuvre correspondant donc à des conditions idéales pour la méthode proposée. Par ailleurs, dans le pipeline total (*cf.* Figure 6.2) de l'identification du type monétaire, à l'aide des informations *a priori* sur les millésimes des monnaies traitées, le taux de reconnaissance sera encore augmenté. Par exemple, le premier chiffre du millésime d'une monnaie « moderne » ne pourra être que « 1 » ou « 2 ».

Enfin, en considérant la variabilité et la disponibilité de données numismatiques réelles possédées par la société GENI, nous n'avons pas beaucoup étudié les méthodes basées sur l'apprentissage machine. Grâce à la qualité de photo et l'état de conservation de monnaies considérées, notre méthode est applicable dans le contexte réel pour la société qui n'a pas assez de données d'apprentissage pour le moment.

Chapitre 7

Gradation des monnaies

La gradation des monnaies par analyse d'images est le dernier problème abordé dans cette thèse. Malgré des critères universels décrits dans les systèmes de gradation comme l'échelle de Sheldon, ces derniers sont difficilement applicables de manière exacte. Ainsi, les grades attribués dépendent fortement des gradeurs, y compris dans le cadre des sociétés de gradation qui proposent une soi-disant gradation impartiale. Par ailleurs, l'objectivité de la gradation est influencée par d'autres facteurs. Par exemple, il est possible que les gradeurs tendent à donner des grades élevés à leurs monnaies préférées ou à celles soumises par des collectionneurs célèbres. Par conséquent, le grade d'une monnaie est considéré crédible si deux ou plusieurs gradeurs expérimentés donnent des avis similaires. Évidemment, il s'agit d'un travail qui requiert des interventions humaines coûteuses, et la société GENI souhaiterait faire une gradation plus objective et plus efficace que le processus humain à l'aide de l'analyse d'images et de l'intelligence artificielle.

Dans ce chapitre, nous proposons une méthode d'aide à la gradation des monnaies notamment pour les monnaies bien conservées. En effet, ce sont plutôt des belles monnaies bien conservées que les collectionneurs souhaitent faire grader pour les mettre en valeur. En outre, la différence visuelle entre des monnaies ayant des grades élevés proches est souvent très subtile à l'œil, alors que celle entre des monnaies de grades faibles proches est relativement évidente (*cf.* Figure 7.1). Dans le contexte de la gradation des monnaies bien conservées, nous supposons que les motifs en relief des monnaies considérées sont à 100% présents. Ainsi, la méthode proposée consiste à quantifier les « éléments inattendus » comme les rayures, les taches, les coups, *etc.*, en s'appuyant sur une monnaie « parfaite » comme référence. Dans la section 7.1, nous détaillons les approches pour détecter et quantifier les « éléments inattendus » causés par l'usure. Dans la section 7.2, nous présentons les expériences réalisées sur la base de données *20FrancsCOQ_Photos* et analysons les résultats obtenus. Enfin, la section 7.3 conclut notre méthode et donne des perspectives.

7.1 Méthodologie proposée

Les monnaies à grader dans le contexte de cette thèse sont censées comporter tous les « éléments attendus » (motifs en relief). Dans ce cas-là, ce sont principalement des « éléments inattendus » (rayures, taches, coups, *etc.*) qui déterminent le grade. D'autres facteurs contributifs existent, par exemple la patine ou la tranche, mais ne sont pas étudiés dans ce chapitre. D'une part, ils sont moins déterminants que les « éléments inattendus » ; d'autre part, nous n'avons pas assez de données pertinentes pour étudier tous les facteurs. Ainsi, la gradation des monnaies sera effectuée uniquement par quantification des « éléments inattendus ».

Grâce aux étapes que nous avons présentées dans les chapitres précédents, la monnaie à grader est segmentée et identifiée. Nous détectons et analysons les « éléments inattendus » sur la monnaie segmentée par comparaison avec une monnaie « parfaite » correspondant

Figure 7.1 – Exemples de monnaies *US Indian 2.5 Dollars* (revers) gradées par la société PCGS : (a) monnaies usées de grades faibles proches ; (b) monnaies bien conservées de grades élevés proches

à son type monétaire identifié. La comparaison est effectuée entre les deux monnaies recalées. La Figure 7.2 montre le pipeline total à partir de la photo professionnelle brute jusqu'à la gradation.

Concernant les « éléments inattendus », il en existe de plusieurs types d'un point def vue numismatique. Certains représentent un creux sur une surface plate, par exemple, un coup (*cf.* Figure 7.3 (a)) ou une rayure profonde (*cf.* Figure 7.3 (b)) ; d'autres correspondent à une zone plate de couleur différente par rapport aux alentours, par exemple, une tache (*cf.* Figure 7.3 (c)). Sur les photos professionnelles, les « éléments inattendus » sont visibles par une discontinuité de couleur par rapport à la couleur attendue de l'arrière-plan. En ce qui concerne les coups et les rayures, ils sont moins éclairés que la surface plate de la monnaie. Ceci est dû à la différence de hauteur. Ils apparaissent donc plus sombres dans l'image. Cependant, il est parfois difficile de distinguer le type d'un « élément inattendu » uniquement à partir de l'image. Une rayure est une discon-

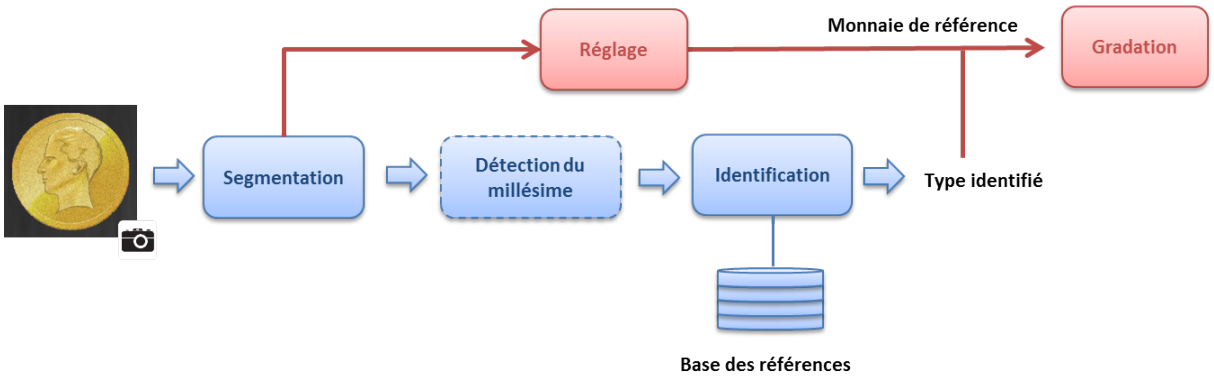


FIGURE 7.2 – Pipeline total de l'analyse d'images pour la gradation des monnaies

tinuité de couleur formant une ligne (courbe ou droite) de petite largeur, alors qu'une tache est une zone fermée qui présente une différence de couleur importante. Mais, une rayure courte profonde et une petite tache allongée sont visuellement très proches dans l'image. Par ailleurs, la texture du métal peut elle-même créer des petites variations de couleur partout dans la surface de la monnaie. Par conséquent, certains petits « éléments inattendus » peuvent être confondus avec ces non-uniformités intrinsèques de la monnaie.

Pour traiter toutes les difficultés mentionnées concernant les « éléments inattendus » dans l'image, nous les divisons en deux groupes et les traiterons différemment (*cf.* Figure 7.5) :

- **« Éléments inattendus » significatifs :** grosses rayures et grosses taches qui présentent une différence de couleur évidente par rapport à l'arrière-plan.

- **« Éléments inattendus » ténus :** micro-rayures ou micro-coups difficiles à repérer individuellement, mais l'ensemble rend la zone de la monnaie considérée plus « usée » qu'elle ne devrait l'être.

Le pipeline détaillé pour l'étape de la gradation des monnaies est montré à la Figure 7.4. Comme indiqué, nous recalons d'abord la monnaie à grader par rapport à une monnaie de référence « parfaite » correspondant à son type monétaire. Le paragraphe 7.1.1 détaille la méthode utilisée pour recaler les monnaies et le paragraphe 7.1.2 présente la méthode développée pour construire les monnaies « parfaites ». Puis, nous détectons les « éléments inattendus » significatifs par comparaison entre la monnaie « parfaite » construite et la monnaie à grader. L'algorithme est détaillé dans le paragraphe 7.1.3. Ensuite, nous détectons des zones censées être uniformes qui contiennent des « éléments inattendus » ténus. Cette partie est discutée dans le paragraphe 7.1.4. Au lieu de donner un seul grade numérique comme résultat, la sortie souhaitée par la société GENI se présente sous forme d'informations concrètes sur les deux groupes d'« éléments inattendus » quantifiés en fonction de leurs endroits (sur les motifs principaux ou le champ).

Ce résultat, appelé **Guide de Gradation (GG)**, servira d'indication objective pour les gradeurs. Cette partie est expliquée dans le paragraphe 7.1.5.

(a) (b) (c)

Figure 7.3 – Exemples de différents types « d'éléments inattendus » : (a) coups ; (b) rayures ; (c) taches

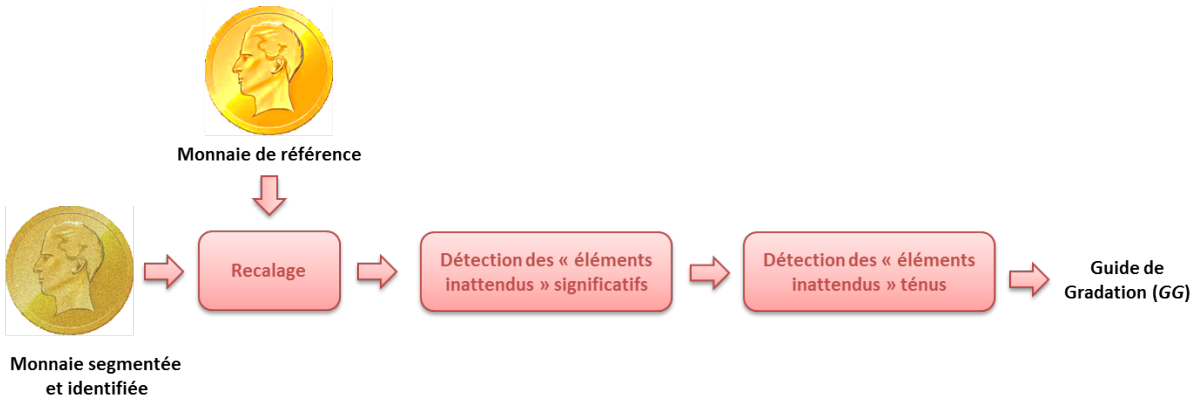


Figure 7.4 – Pipeline de la gradation des monnaies basée sur la quantification des « éléments inattendus »

7.1.1 Recalage des monnaies

Comme nous l'avons mentionné dans les chapitres précédents, les monnaies dans les photos professionnelles ont des orientations similaires mais pas nécessairement exactement les mêmes. Cette nuance d'orientation n'influence pas beaucoup sur l'identification du type monétaire. Par contre, le recalage est important pour comparer la monnaie à grader avec la monnaie de référence « parfaite ».

Soient I_k la monnaie à grader et I_R la monnaie de référence, l'objectif est de trouver une transformation qui minimise l'écart entre deux images après transformation selon un critère de similarité. Dans notre contexte, le recalage est effectué sur les cartes des gradients I_k^G et I_R^G des images respectives parce que ce sont les contours des reliefs impliqués dans les gradients que nous souhaitons mettre en correspondance. Les cartes des gradients sont obtenues à l'aide de l'opérateur Laplacien. Nous utilisons un filtre médian pour supprimer les petits gradients isolés afin que la plupart des gradients retenus se trouvent aux endroits des contours des reliefs. Nous notons $I_k^{G'}$ et $I_R^{G'}$ les cartes des gradients traitées.

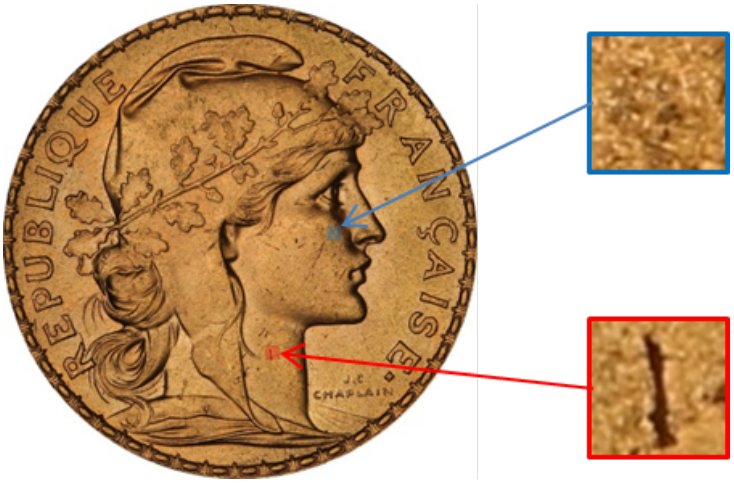


FIGURE 7.5 – Exemples « d'élément inattendu » significatif (zoomé et encadré en rouge) et « d'élément inattendu » ténu qui rend la surface « granuleuse » (zoomé et encadré en bleu)

Puisque les monnaies sont segmentées, il suffit d'appliquer une rotation de centre le centre de la monnaie et d'angle θ dans le domaine de la monnaie Ω. Nous notons $I_k^{G'}(\theta)$ la carte des gradients obtenue à l'aide de $I_k^{G'}$ tournant de θ dans le sens horaire. Parmi les nombreuses mesures de similarité qui ont été développées à ce jour, nous utilisons la *Sum of Squared Difference* (SSD), donnée par :

$$SSD(I_k^{G'}(\theta), I_R^{G'}) = \sum_{(x,y)\in\Omega} (I_R^{G'}(x,y) - I_k^{G'}(\theta)(x,y))^2 \tag{7.1}$$

Les gradients qui présentent les contours des reliefs respectifs des deux monnaies correspondront le mieux lorsque la SSD sera minimale. L'objectif est donc de trouver la rotation qui minimise la SSD entre les deux images, donnée par :

$$\theta^* = \underset{\theta\in[-\theta_0,\theta_0]}{argmin}\ SSD(I_k^{G'}(\theta), I_R^{G'}) \tag{7.2}$$

où θ_0 est une constante petite (expérimentalement, $\theta_0 = 10°$) puisque la différence des orientations originales est faible.

7.1.2 Construction des monnaies de référence

Dans notre contexte, une monnaie de référence « parfaite » ne devrait comporter aucune non-uniformité sur la surface aux endroits où il n'y a pas de relief dus aux motifs. Cependant, aucune monnaie réelle n'est exactement parfaite. Même sur les monnaies qui ont été gradées « Fleur de coin », nous pouvons souvent trouver des petits « éléments inattendus » qui échappent au gradeur. Par ailleurs, la surface des monnaies (sauf la surface miroir des « belles épreuves ») n'est jamais uniforme dans l'image à cause de la texture du métal. Ainsi, nous proposons de construire une monnaie « parfaite » synthétique sur

laquelle les non-uniformités se trouvent uniquement aux endroits des reliefs. Autrement dit, la monnaie « parfaite » idéale peut être considérée comme un masque binaire dont les pixels sont divisés en deux classes : 1) ceux qui appartiennent aux contours des reliefs et 2) ceux qui appartiennent aux régions censées être uniformes.

Nous construisons ces monnaies de référence « parfaites » synthétiques grâce aux images moyennes des monnaies réelles recalées selon les hypothèses suivantes :

- Les « éléments attendus » sont en commun pour les monnaies appartenant à un même type monétaire.

- Les « éléments inattendus » sont individuels pour chaque monnaie.

- Les contours des « élément attendus » sont plus sombres que l'arrière-plan, ainsi que certains « éléments inattendus ».

- Les contours des « élément attendus » présentent souvent un changement abrupt d'intensité, ainsi que de certains « éléments inattendus ».

Soit $I_i, i \in [1, N]$ l'une des N monnaies d'un même type monétaire, I_i^* est la monnaie recalée sur la première image I_1 à l'aide de la méthode de recalage que nous avons présentée précédemment (*cf.* Figure 7.6 (a)). Une monnaie « parfaite » en couleur est obtenue par la moyenne des monnaies recalées (*cf.* Figure 7.7 (a)) :

$$\hat{I} = \frac{1}{N} \sum_{i=1}^{N} I_i^*. \tag{7.3}$$

Cependant, la couleur moyenne n'a rien à voir avec une couleur de référence. La couleur des monnaies dépend également de la patine que nous étudierons dans le futur. Dans ce chapitre, ce sont uniquement les zones qui nous intéressent (zones avec des éléments attendus ou inattendus).

Ainsi, nous convertissons I_i^* en niveaux de gris pour calculer la carte des gradients I_i^G, en utilisant l'opérateur Laplacien (*cf.* Figure 7.6 (b)), et la monnaie seuillée I_i^T, obtenue par la méthode d'Ostu (*cf.* Figure 7.6 (c)). Plus précisément, la carte des gradients I_i^G présente les changements abrupts d'intensité, y compris des « grains » de la texture du métal. La monnaie seuillée I_i^T est composée des parties sombres de la monnaie. Puis, nous calculons la moyenne des cartes des gradients recalées (*cf.* Figure 7.7 (b))

$$\widehat{I^G} = \frac{1}{N} \sum_{i=1}^{N} I_i^G \tag{7.4}$$

et la moyenne des monnaies seuillées recalées (*cf.* Figure 7.7 (c))

$$\widehat{I^T} = \frac{1}{N} \sum_{i=1}^{N} I_i^T. \tag{7.5}$$

Sur $\widehat{I^G}$ et $\widehat{I^T}$, les éléments communs sont retenus alors que les éléments individuels sont filtrés. Il faut mentionner que $\widehat{I^G}$ ne comporte que les contours des « éléments attendus », alors que $\widehat{I^T}$ comporte les contours des « éléments attendus » ainsi que certaines régions moins éclairées (par exemple, les voisinages de l'œil et du nez dans la Figure 7.7 (a)). Grâce au fait que les monnaies sont photographiées sous les mêmes conditions, les régions sombres des I_i^* sont communes, et nous les considérons comme les non-uniformités attendues. Ainsi, les contours des « éléments attendus » et les régions sombres communes constituent la monnaie de référence que nous utiliserons, donnée par

$$M = \widehat{I^G}_B + \widehat{I^T}_B, \tag{7.6}$$

où $\widehat{I^G}_B$ et $\widehat{I^T}_B$ sont les images binarisées de $\widehat{I^G}$ et $\widehat{I^T}$. M est un masque binaire où tous les pixels blancs appartiennent aux contours des « éléments attendus » et aux régions sombres, et tous les pixels noirs appartiennent aux zones uniformes attendues (*cf.* Figure 7.8 (a)).

(a) (b) (c)

FIGURE 7.6 – Prétraitement d'une monnaie à grader : (a) monnaie en couleur recalée I_k^* ; (b) carte des gradients I_k^G ; (c) monnaie seuillée I_k^T

Grâce à la monnaie de référence construite, nous rechercherons les « éléments inattendus » de la monnaie à grader I_k sur les zones d'intérêt ZI_k censées être uniformes (*cf.* Figure 7.8 (b)) :

$$ZI_k = I_k^* \oplus M \tag{7.7}$$

Dans les paragraphes suivants, nous présentons respectivement la détection des « éléments inattendus » significatifs et ténus sur ZI_k.

La moyenne des cartes des gradients $\hat{I^G}$ sera utilisée comme la référence I_R^G pour recaler les monnaies à grader appartenant à son type monétaire.

(a) (b) (c)

Figure 7.7 – Images moyennes des monnaies : (a) monnaie moyenne en couleur $\hat{I}$; (b) moyenne des cartes des gradients $\widehat{I^G}$; (c) moyenne des monnaies seuillées $\widehat{I^T}$

7.1.3 Détection des « éléments inattendus » significatifs

Les « éléments inattendus » significatifs sont des grosses rayures ou des grosses taches qui présentent une différence évidente par rapport à l'arrière-plan, et que nous pouvons facilement repérer à l'œil.

Parmi les « éléments inattendus » significatifs dans la monnaie à grader I_k, ceux qui présentent des changements abrupts d'intensité, par exemple, une rayure, sont détecté via la carte des gradients I_k^G ; ceux qui apparaissent plus sombres que l'arrière-plan, par exemple, une tache noire, sont détectés dans la monnaie seuillée I_k^T. Évidemment certains « éléments inattendus » peuvent être repérés directement dans I_k^G et I_k^T. Cependant, les gradients sont peu performants pour détecter des taches complètes et leurs contours sont souvent discontinus. De même, le seuillage ne permet pas de détecter des rayures moins sombres. Ainsi, c'est la combinaison de I_k^G et I_k^T qui nous permet de détecter tous les « éléments inattendus » significatifs de manière relativement complète.

Notons $\bar{M}$ le complémentaire de M du domaine de la monnaie segmentée, e_k^G et e_k^T les « éléments inattendus » significatifs repérées respectivement sur la carte des gradients binarisée $I_k^{G_B}$ et la monnaie seuillée I_k^T, que n'appartiennent pas à M :

$$\begin{aligned} e_k^G &= I_k^{G_B} \cdot \bar{M} \\ e_k^T &= I_k^T \cdot \bar{M} \end{aligned} \tag{7.8}$$

Il faut noter qu'il existe des petits bruits partout sur e_k^G ainsi que sur e_k^T. Ces bruits sont soit causés par les « éléments inattendus » ténus, soit en provenance des surfaces intrinsèquement texturées de la monnaie. Nous filtrons ces petits bruits quelles que soient leurs causes, puisque dans cette étape nous nous focalisons sur les « éléments inattendus » significatifs. Il faut faire attention de ne pas supprimer les rayures longues de petite largeur. Ainsi, nous filtrons les composantes connexes dans e_k^G et e_k^T dont l'aire est inférieure à un seuil expérimental. $e_k^{G'}$ et $e_k^{T'}$ sont les images filtrées correspondantes qui contiennent

(a) (b)

FIGURE 7.8 – (a) Masque binaire de la monnaie de référence M ; (b) zone d'intérêt ZI_k d'une monnaie réelle

les composantes connexes importantes. Ainsi, la carte des « éléments inattendus » significatifs e_k est obtenue par

$$e_k = e_k^{G'} + e_k^{T'} \tag{7.9}$$

sur laquelle chaque composante connexe retenue présente un « élément inattendu » significatif. Ce processus est illustré à la Figure 7.9 avec le résultat d'une zone d'intérêt zoomée. Il faut noter que certains petits « éléments inattendus » mais quand même visibles à l'œil, peuvent être filtrés dans cette étape car il est difficile de les différencier du bruit de l'arrière-plan. Cependant, ces petits « éléments inattendus » seront pris en compte dans les « éléments inattendus » ténus que nous allons détecter dans le paragraphe suivant. D'autres résultats sont montrés au paragraphe 7.2.1.

7.1.4 Détection des « éléments inattendus » ténus

Les « éléments inattendus » ténus sont des petites rayures et des petits coups, sur les surfaces des monnaies, difficiles à détecter par la méthode proposée précédemment. Ils présentent pourtant une granulosité particulière dans la zone d'intérêt ZI_k. Pour ce faire, nous considérons que ZI_k est composée de deux parties :

- Zone granuleuse Z_g : partie de la zone d'intérêt visuellement granuleuse causée par les « éléments inattendus » ténus (*cf.* Figure 7.10 (a)).

- Zone lisse Z_s : partie de la zone d'intérêt quasi-uniforme qui ne comporte pas

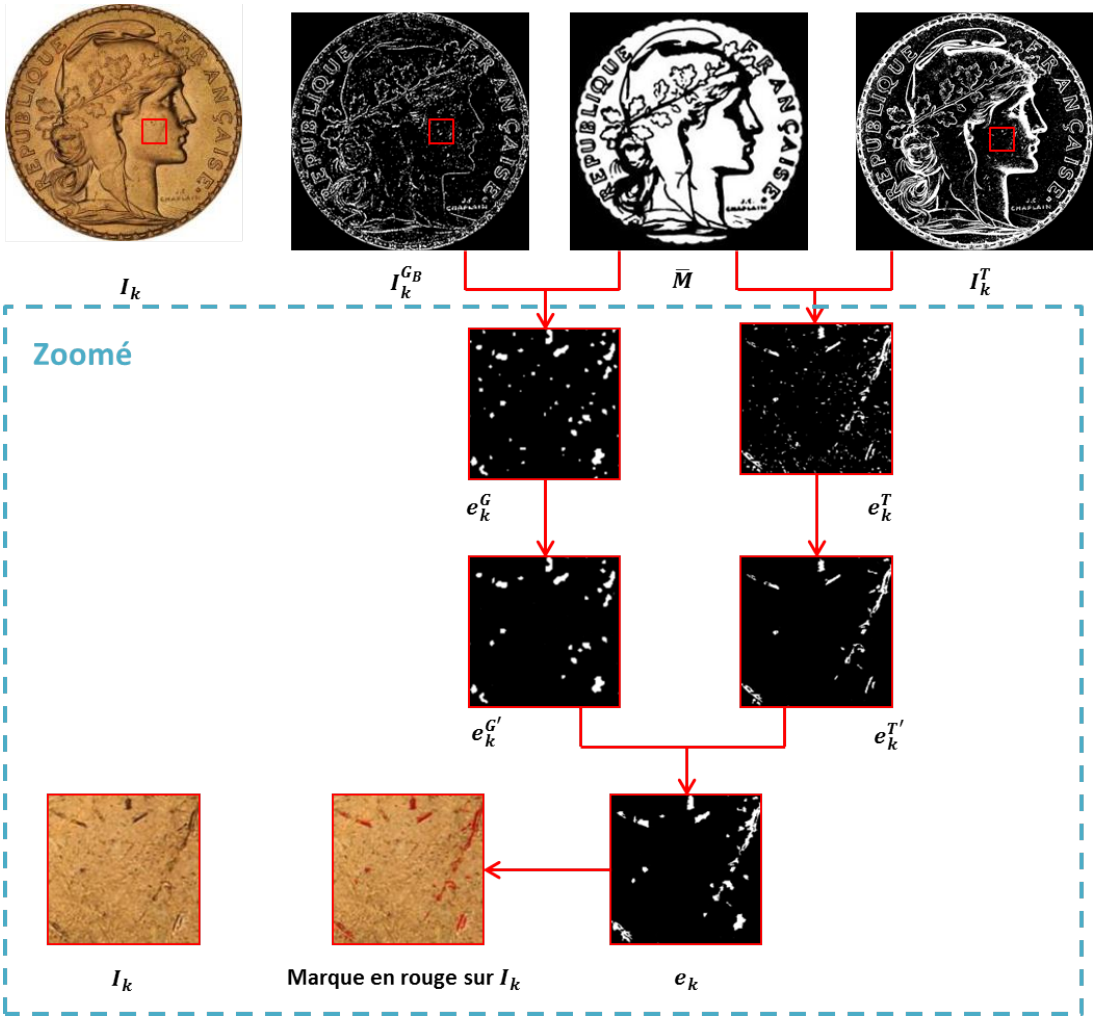


Figure 7.9 – Détection des « éléments inattendus » significatifs

« d'éléments inattendus » ténus dus à l'usure (*cf.* Figure 7.10 (b)).

L'objectif de cette étape est de quantifier la zone granuleuse de la monnaie qui contient les « éléments inattendus » ténus. Pour ce faire, nous divisons la zone d'intérêt en imagettes carrées (*patches*) de taille de $n \times n$ non superposées. Nous notons P l'ensemble des imagettes extraites de la zone d'intérêt.

Une imagette $p_i \in P$ censée comporter des « éléments inattendus » ténus est une imagette granuleuse, sinon une imagette lisse. A l'aide de données annotées par un expert, nous classifions chaque imagette dans l'une des deux classes « granuleuse » ou« lisse ». Enfin, l'ensemble des imagettes classifiées comme « granuleuses » P_g permet d'obtenir la zone granuleuse Z_g qui contient des « éléments inattendus » ténus.

Taille des imagettes

Pour la monnaie à grader, l'ensemble des imagettes découpées P est un sous-ensemble de la zone d'intérêt ZI, $P \subset ZI$. Plus les imagettes sont petites, plus P est proche de ZI (*cf.* Figure 7.11). Cependant, la granulosité de la zone est une observation portée sur une

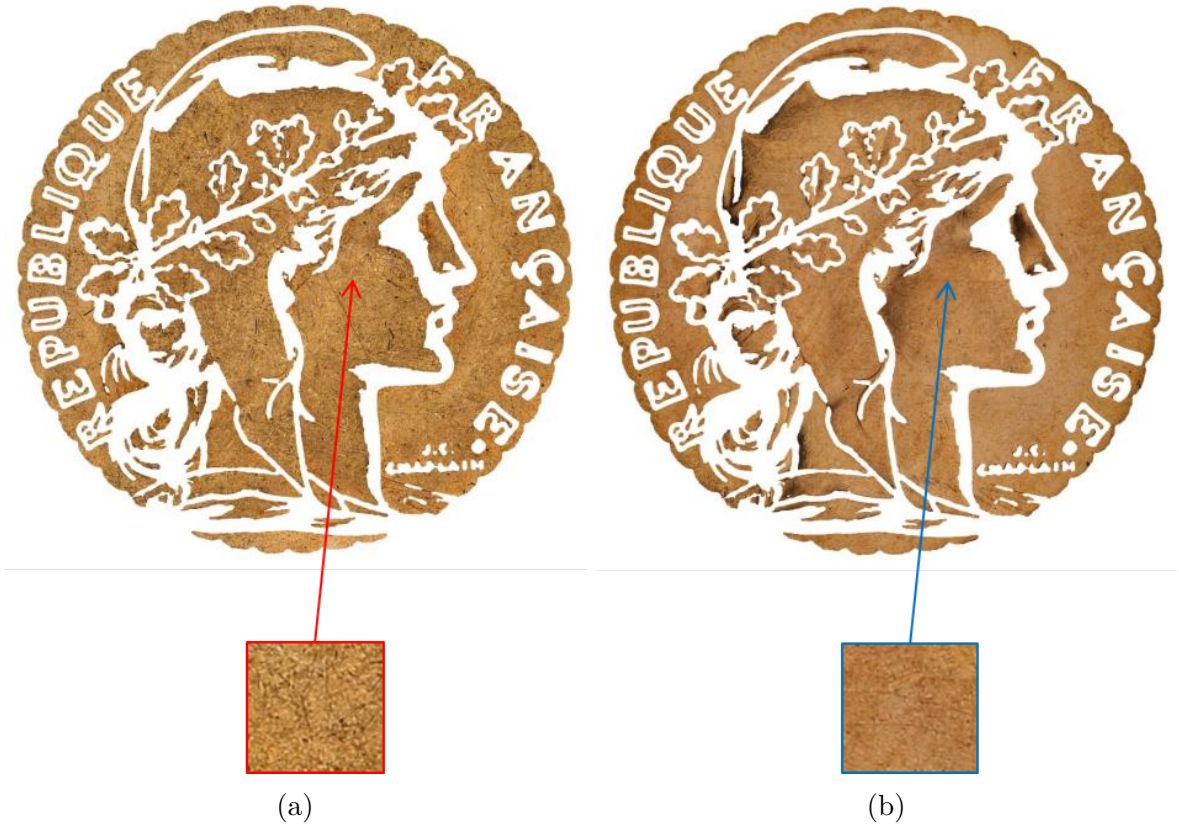


FIGURE 7.10 – Zones d'intérêt des monnaies avec des imagettes extraites zoomées : (a) imagette granuleuse ; (b) imagette lisse

surface assez grande. Si les imagettes sont trop petites, la granulosité sera moins présente (*cf.* Figure 7.12). Ainsi, il faut trouver un bon compromis pour la taille des imagettes. Au vu de la complexité du problème, une bonne représentation des « éléments inattendus » ténus est le plus important. Au vu des expériences menées, nous avons fixé la taille des imagettes à 64×64 pixels pour les monnaies segmentées normalisées à 2048×2048 pixels. Par conséquent, nous ne considérons pas certaines régions de ZI, notamment autour des contours des « éléments attendus », où une imagette de taille 64×64 pixels ne rentre pas.

Annotation des imagettes

La classification des imagettes découpées nécessite des données annotées, selon les deux classes « grauleuses » ou « lisses » par expert.

Contrairement à la monnaie « parfaite » synthétique (*cf.* Figure 7.7 (a)), aucune surface de monnaies réelles n'est exactement uniforme. Les imagettes lisses sont ainsi des imagettes légèrement non-uniformes à cause de la texture du métal, du velours de frappe, des petits coups de sac, *etc.* (*cf.* Chapitre 1). De manière générale, c'est la granulosité introduite par la circulation des monnaies que nous souhaitons détecter comme « éléments inattendus ». Par ailleurs, la classification des imagettes est indépendante des « éléments

(a) (b) (c)

Figure 7.11 – Division de la zone d'intérêt en imagettes de différentes tailles fixées : (a) 32 × 32 pixels ; (b) 48 × 48 pixels ; (c) 64 × 64 pixels

inattendus » significatifs. Par exemple, une imagette quasi-uniforme qui contient uniquement une rayure évidente devrait être considérée comme une imagette lisse.

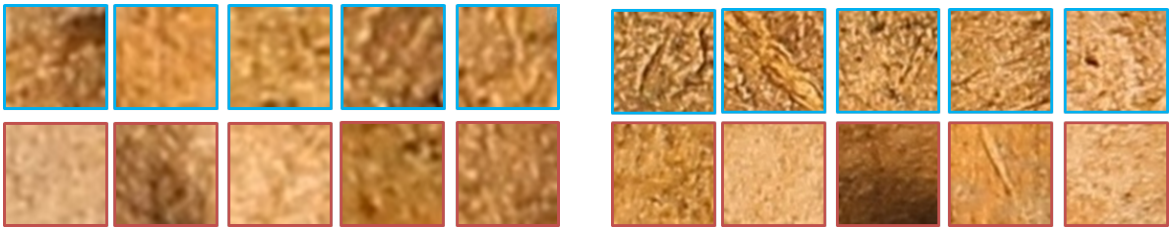
À cause de la complexité de l'annotation, un néophyte peut difficilement faire ce travail. Par ailleurs, même pour des gradeurs expérimentés, il s'agit d'une tâche difficile car ces derniers sont moins à l'aise pour analyser une monnaie sur une photo qu'à la main. Malgré ces contraintes, dans le cadre de cette étude, la société GENI nous a annoté des zones granuleuses sur la base de données *20FrancsCOQ_Photos*.

Notons $\boldsymbol{Z}_g$ et $\boldsymbol{Z}_s$ respectivement les ensembles de zones granuleuses et lisses de toutes les monnaies annotées. Nous extrayons les imagettes de taille fixée dans ces zones. Celles extraites de $\boldsymbol{Z}_g$ sont des imagettes granuleuses, notées $\boldsymbol{P}_g$; celles extraites de $\boldsymbol{Z}_s$ sont des imagettes lisses, notées $\boldsymbol{P}_s$.

Concernant les imagettes annotées, nous supposons que celles de $\boldsymbol{P}_g$ soient globalement moins uniformes que celles de $\boldsymbol{P}_s$. Cependant, ce n'est pas toujours les cas. Certaines imagettes de $\boldsymbol{P}_s$ apparaissent autant, voire moins, uniformes que celles de $\boldsymbol{P}_g$. La Figure 7.13 montre des exemples d'imagettes annotées en niveaux de gris dont l'intensité a été normalisée. Notre hypothèse est que la couleur n'apporte rien pour différencier la granulosité des surfaces. Nous proposons ce prétraitement afin d'avoir une illustration claire de la différence d'uniformité entre les deux classes.

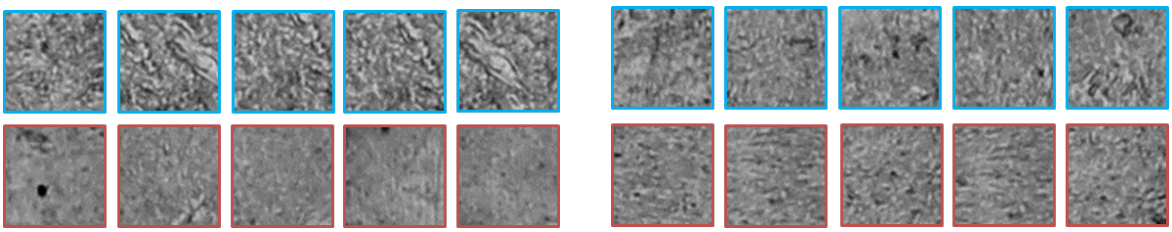
Écarts-types des imagettes

Intrinsèquement, la granulosité des imagettes en niveaux de gris est caractérisée par la variation des intensités (*cf.* Figure 7.13 (a)). Bien que les exemples de la Figure 7.13 (b) montrent des cas particuliers, nous supposons que la plupart des imagettes granuleuses possèdent une dispersion d'intensité plus élevée que les imagettes lisses, mesurée par



(a) (b)

FIGURE 7.12 – Exemples d'imagettes annotées de différentes tailles (redimensionnées à la même taille pour la visualisation ; imagettes granuleuses encadrées en bleu ; imagettes lisses encadrées en rouge) : (a) 32 × 32 pixels ; (b) 48 × 48 pixels



(a) (b)

FIGURE 7.13 – Exemples d'imagettes annotées prétraitées (imagettes granuleuses encadrées en bleu ; imagettes lisses encadrées en rouge) : (a) différence d'uniformité évidente ; (b) différence d'uniformité pas évidente

l'écart-type d'image :

$$\sigma_i = \sqrt{\frac{1}{n_p - 1} \sum_{1}^{n_p} (p_i(x, y) - \hat{p_i})^2} \tag{7.10}$$

où $\hat{p_i}$ est la moyenne des intensités de l'imagette p_i, et n_p le nombre de pixels contenus dans p_i. Une méthode naïve est de trouver un seuil optimal σ_{th} qui permet de séparer au mieux les deux classes sur les imagettes annotées. Pour la monnaie à grader, les zones granuleuses Z_g sont constituées par les imagettes découpées dont les écarts-types sont supérieurs à σ_{th} ; les zones lisses Z_l sont l'ensemble des imagettes découpées avec un écart-type inférieur ou égal à σ_{th}.

Cette approche naïve est performante pour classifier des imagettes quasi-uniformes et non-uniformes. Cependant, la granulosité que nous souhaitons détecter est beaucoup plus complexe et l'écart-type ne permet de capturer cette complexité. Par conséquent, cette approche n'arrive pas à distinguer les imagettes lisses mais visuellement non-uniformes, par exemple, celles qui présentent une texture du métal ou un velours de frappe avec un contraste évident, et les imagettes globalement lisses qui comportent un ou plusieurs « éléments inattendus » significatifs.

Apprentissage profond (*Deep Learning*)

Pour tenter de trouver une méthode efficace de classification des imagettes, et comme nous avons suffisamment de données et un étiquetage expert, nous avons décidé de faire appel à une intelligence artificielle de haut niveau qui permet de trouver les nuances que nous ne voyons pas à l'œil humain.

À ce jour, le *Deep Learning* et les réseaux de neurones convolutifs (*Convolutional Neural Networks* (CNN)) permet d'atteindre des résultats impressionnants en termes de classification d'images. Basé sur la structure basique de ANN (que nous avons testée dans le Chapitre 6), l'image d'entrée est passée au travers d'une succession de filtres de convolution pour extraire les caractéristiques et réduire la résolution de l'image. Puis, les images prétraitées par différents noyaux de convolution sont mises en plat comme entrée d'un perceptron multicouches entièrement connecté. La sortie de la dernière couche est la prédiction de la classe. Grâce à cette structure, les poids des filtres ainsi que ceux des neurones sont entraînés en même temps afin d'obtenir une meilleure classification. La Figure 7.14 montre la structure générale des CNN pour classifier des imagettes dans notre contexte.

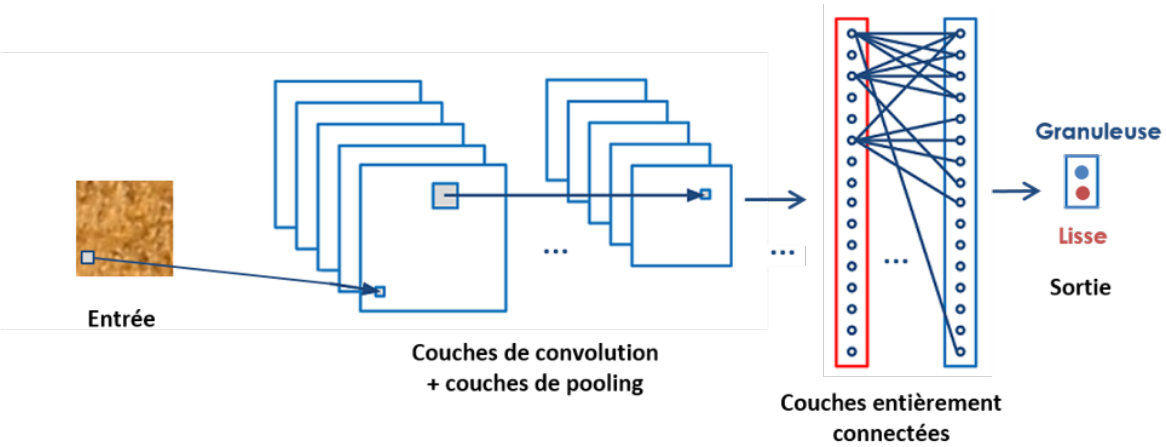


Figure 7.14 – Structure générale d'un réseau de neurones convolutionnels

Les performances extraordinaires de CNN en classification d'images dans différents contextes sont au prix d'un entraînement du modèle souvent coûteux pour fixer la structure et les paramètres associés, et qui nécessite un grand nombre de données annotées (généralement des milliers ou plus).

Cependant, dans la pratique, il est possible d'utiliser un modèle performant pré-entraîné et de l'adapter à l'application souhaitée en faisant du *Fine Tuning*. Le *Fine Tuning* consiste à profiter de la structure et des paramètres des premières couches d'un modèle optimisé par des experts en optimisant uniquement les paramètres des dernières couches pour une tâche similaire. Grâce à cette technique, nous pouvons utiliser les CNN de manière moins coûteuse en termes de temps et sans trop d'expertise.

Parmi tous les modèles CNN pré-entraînés performants sur une autre base de données, nous avons choisi d'utiliser le réseau *AlexNet* (Krizhevsky, Sutskever et Hinton 2012) qui est celui qui nous a donné les meilleurs résultats. Ce modèle prend en entrée des images en couleur de taille 256 × 256 pixels comme d'entrée. Nos imagettes annotées sont donc redimensionnées pour s'adapter à ce modèle. Grâce au *Fine Tuning*, nous adaptons les poids de la dernière couche du réseau. Le réseau entraîné sur les imagettes annotées est ensuite utilisé comme classifieur appliqué sur les imagettes découpées sur les zones d'intérêt des monnaies à grader.

7.1.5 Gradation résultante

Grâce aux étapes précédentes, nous obtenons la carte des « éléments inattendus » significatifs e et les imagettes granuleuses P_g qui contiennent les « éléments inattendus » ténus. Le résultat général de notre processus est donné par la surface des « éléments inattendus » significatifs, notée SS, le nombre des « éléments inattendus » significatifs, noté NS, et le pourcentage de la surface couverte par les imagettes granuleuses par rapport à la surface de zone d'intérêt ZI, noté RT,

$$RT = \frac{n_g \times n^2}{ZI} \tag{7.11}$$

où n_g est le nombre des imagettes granuleuses de taille de $n \times n$ dans P_g. Il faut noter que nous utilisons la surface absolue des « éléments inattendus » significatifs mais la surface relative des « éléments inattendus » ténus, parce que selon les numismates la gravité d'une grosse rayure est indépendante de la taille de la monnaie alors que l'influence des micro-rayures sont plutôt un pourcentage des surfaces qui ont l'air usées. Ainsi, la quantification générale des « éléments inattendus », appelée Guide de Gradation (GG), est donné par,

$$GG = \{SS, NS, RT\} \tag{7.12}$$

Selon les numismates, les endroits où se trouvent les « éléments inattendus » sont importants. Une rayure sur le motif principal, par exemple sur le portrait, est plus grave que la même rayure sur le champ. La société GENI souhaite que nous fournissions une quantification précise des « éléments inattendus » en fonction de leur position.

Pour chaque type monétaire donné, il est donc nécessaire d'annoter les endroits essentiels et les endroits moins importants. Ceci peut être fait par des experts, cependant, il est également possible de le réaliser de manière automatique en supposant qu'en général, les motifs importants sont dans le centre de la monnaie alors que le champ est la partie extérieure qui touche le bord. Pour cela, nous effectuons un étiquetage en composantes connexes de la zone d'intérêt (*cf.* Figure 7.15 (a)). Toutes les composantes connexes qui

touchent le bord de la monnaie sont étiquetés comme « peu importantes » (en vert à la Figure 7.15 (b)) et les autres comme « importantes » (en rouge à la Figure 7.15 (b)).

Il faut noter que cette classification est approximative. Par exemple, une feuille de la guirlande à la Figure 7.15 (b) est annotée dans la partie du champ à cause de contours non fermés. Cependant, de telles erreurs ne sont pas importantes pour les types monétaires courants. L'utilisation de l'annotation manuelle précise ou de l'annotation automatique approximative sera un choix à faire en fonction de la précision attendue. Les « éléments inattendus » détectés doivent alors être pondérés différemment en fonction de leur endroit.

Nous pouvons considérer que GG est un résultat intermédiaire pour aider les gradeurs en leur donnant des informations objectives sur les « éléments inattendus ». Concernant le grade numérique G_N à communiquer sur le marché, il s'agirait d'une fonction pondérée expérimentalement qui prendrait GG en compte,

$$G_N = f(GG, \boldsymbol{W}) \tag{7.13}$$

où $\boldsymbol{W}$ est l'ensemble des poids expérimentaux associés aux éléments de GG. Dans ce chapitre, nous nous concentrons uniquement sur la quantification des « éléments inattendus ». L'association de GG à un grade numérique sera étudiée postérieurement au sein de la société GENI.

(a) (b)

Figure 7.15 – (a) Étiquetage des composantes connexes de la zone d'intérêt et (b) classification automatique en deux zones d'importance

7.2 Expériences et discussions

Les expériences ont été effectuées sur la base de données *20FrancsCOQ_Photos* à l'aide des « éléments inattendus » annotés par un gradeur de la société GENI. Cette base de données est composée de 143 photos de l'avers et de 143 photos du revers. Les images sont déjà segmentées et recalées à l'aide des prétraitements précédemment proposés. Les monnaies de référence des deux côtés ont été construites à l'aide de toutes les images.

7.2.1 Détection des « éléments inattendus » significatifs

L'expert de la société GENI a annoté les « éléments inattendus » significatifs sur la base de données *20FrancsCOQ_Photos*. Cependant, cette annotation est seulement indicative et elle sert difficilement de vérité terrain pour réaliser une évaluation quantitative. En effet, tout d'abord, nous n'avons pas de critère standardisé sur de tels « éléments inattendus » significatifs qui n'ont aucun sens numismatique. Plus précisément, pour nous, ce sont des composantes connexes, qui représente une non-uniformité par rapport à l'arrière-plan, ayant une surface supérieure à un seuil expérimental permettant de séparer les bruits et les « éléments inattendus » ténus. Pour l'expert qui a annoté les données, il s'agit plutôt de défauts graves à l'œil même sur les contours en reliefs. Par ailleurs, la plupart des annotations ont été effectuées par un pinceau Photoshop de même épaisseur quelle que soit la taille réelle des « éléments inattendus » annotés, sauf pour certaines rayures ou taches extrêmement grosses. Ensuite, nous constatons que certaines rayures et taches évidentes ne sont pas annotées. Autrement dit, cette annotation est incomplète. En effet, dans la pratique le gradeur ne compte jamais un par un tous les « éléments inattendus ».

De manière générale, les résultats obtenus par notre méthode sur la détection des « éléments inattendus » significatifs est similaire à l'annotation de l'expert (*cf.* Figure 7.16). Une comparaison des résultats obtenus avec l'annotation de l'expert sur les imagettes zoomées est montré à la Figure 7.17, sur laquelle nous pouvons observer :

- La taille de la rayure détectée paraît plus correcte que celle annotée (*cf.* Figure 7.17 (a)).
- Les coups annotés sur les contours des reliefs ne sont pas détectés (*cf.* Figure 7.17 (b)).
- Un coup éclairé annoté n'est détecté que dans la partie sombre (*cf.* Figure 7.17 (c)).
- Une rayure annotée n'est détectée que dans la partie dans la zone d'intérêt (*cf.* Figure 7.17 (d)).
- Une rayure non annotée est détectée (*cf.* Figure 7.17 (e)).
- Des taches non annotées sont détectées (*cf.* Figure 7.17 (f)).
- Un poil est détecté, mais il s'agit d'un bruit au lieu d'un défaut (*cf.* Figure 7.17 (g)).

La méthode proposée basée sur seuillage et gradients est robuste pour détecter les

« éléments inattendus » abrupts et foncés, même certains d'entre eux ont échappé au gradeur. Cependant, elle est moins performante pour détecter les « éléments inattendus » ayant une couleur similaire à l'arrière-plan, par exemple un coup bien éclairé.

(a) (b)

Figure 7.16 – Exemple d'une monnaie avec détection des « éléments inattendus » significatifs détectés (en rouge) : (a) l'annotation de l'expert ; (b) le résultat obtenu

(a) (b) (c) (d) (e) (f) (g) (h)

Figure 7.17 – Comparaison des « éléments inattendus » significatifs détectés avec l'annotation de l'expert (en haut : annotation ; en bas : résultats obtenus)

7.2.2 Classification des imagettes annotées

Sur toutes les données annotées de la base *20FrancsCOQ_Photos*, nous extrayons des échantillons d'imagettes de taille fixée (64×64 pixels) de manière aléatoire dans les zones respectivement granuleuses et lisses. Pour avoir des données quasi-équilibrées en deux classes et sur toutes les monnaies, nous échantillonnons autant d'imagettes granuleuses que d'imagettes lisses par monnaie si la monnaie comporte deux zones différentes. Il faut noter que certaines monnaies ne contiennent qu'un type de zone. Notons N_g et N_s

TABLE 7.1 – Résultats de classification sur la base de 2865 imagettes annotées

Méthode	**Prétraitement**	**Classification (%)**
AlexNet	Couleur	85,7
AlexNet	Niveaux de gris	83,2
AlexNet	Niveaux de gris norm.	79,8
Ecart-type	Niveaux de gris norm.	76,8

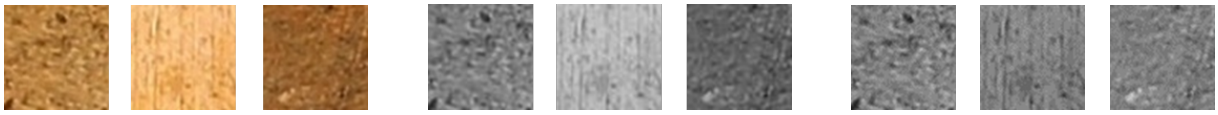
respectivement les nombres totaux d'imagettes annotées granuleuses et lisses, nous avons $N_g \approx N_s$.

La première base d'imagettes de test contient 2865 imagettes annotées ($N_g = 1380$ et $N_s = 1485$). Sur cette base nous testons le seuillage par écart-type et le *Deep Learning* à l'aide du modèle pré-entraîné *AlexNet*. Pour le *Deep Learning*, nous effectuons un *Fine Tuning* pour optimiser la dernière couche de neurones. La base d'apprentissage contient 2149 imagettes et la base de test 716 imagettes. Les distributions des imagettes par classe dans la base d'apprentissage et dans la base de test sont identiques à celles de la base totale.

Nous avons effectué des tests sur :

- les imagettes en couleur (sans prétraitement) (*cf.* Figure 7.18 (a)) ;
- les imagettes en niveaux de gris (*cf.* Figure 7.18 (b)) ;
- les imagettes en niveaux de gris normalisés (*cf.* Figure 7.18 (c)).

La Table 7.1 montre les résultats obtenus en validation croisée sur ces 2865 imagettes annotées. Le *Deep Learning* est plus performant que la méthode naïve basée sur l'écart-type d'image parce que la granulosité ne correspond pas seulement à l'uniformité des pixels. Les imagettes en couleur donnent les meilleurs taux de classification et il semble que la normalisation diminue légèrement la discrimination entre les classes.



(a) (b) (c)

FIGURE 7.18 – Exemples d'imagettes testées : (a) en couleur (sans prétraitement) ; (b) en niveaux de gris ; (b) en niveaux de gris normalisés

Pour évaluer la performance de la détection des « éléments inattendus » ténus sur les monnaies entières, après avoir classifié les imagettes découpées à l'aide du classifieur entraîné, nous les replaçons sur la monnaie pour les comparer avec les zones annotées par le gradeur (*cf.* Figure 7.19). Il faut mentionner que les imagettes utilisées pour l'entraînement et les imagettes à classifier ne sont évidemment pas les mêmes mais elles peuvent

partiellement être superposées car elles sont extraites des mêmes monnaies. Le meilleur taux de classification que nous obtenons est aux alentours de 60%, ceci est obtenu par le réseau en utilisant les imagettes en niveaux de gris normalisés. Ce résultat est beaucoup plus faible que la classification des imagettes elles-mêmes. Les causes peuvent être multiples :

- La base de test sur toutes les monnaies (des dizaines de milliers d'imagettes) est plus grande que celle des expériences précédentes (des milliers).

- Les zones annotées ne respectent pas la division en imagettes. Certaines imagettes découpées à classifier sont à cheval sur les deux types de zones.

- L'annotation grossière faite par l'expert rend les imagettes annotées bruitées. Plus précisément, l'annotation a été réalisée à l'aide de l'outil pinceau de Photoshop. Pour ne pas y passer trop de temps, l'expert qui a annoté ces monnaies a sélectionné un pinceau d'une grande épaisseur. Ainsi, les zones annotées comme granuleuse apparaissent souvent plus grandes que les endroits où se trouvent réellement les « éléments inattendus » ténus. Par ailleurs, l'expert n'a annoté de manière générale que les grandes zones. Les « éléments inattendus » ténus isolés n'ont pas été considérés lors de l'annotation mais sont repérés lors de la classification des imagettes dans les zones.

Par ailleurs, il est possible que la taille de la base d'apprentissage que nous avons utilisée ne soit pas suffisante au vu de la complexité du problème. La Table 7.2 montre les résultats d'expériences en utilisant des bases d'imagettes annotées de tailles différentes. Chaque base contient autant d'imagettes granuleuses que d'imagettes lisses ($N_g = N_s$). Les bases d'apprentissage et de test contiennent respectivement deux tiers et un tiers des données. Nous pourrions nous attendre à ce que plus nous utilisons de données pour l'apprentissage, meilleurs soient les résultats. Cependant, au contraire, nous obtenons un taux de classification qui diminue si nous entraînons le réseau avec trop d'imagettes. Nous observons que de 2000 à 3000 imagettes, le taux de classification augmente légèrement mais chute ensuite si le nombre d'imagettes annotées augmente à 5000. Il semble que la séparation entre les deux classes devienne moins évidente probablement à cause de l'introduction de données bruitées. Avec la base qui contient 9000 imagettes annotées, les taux de classification sont un peu meilleurs que ceux avec 5000 imagettes, mais restent plus faibles qu'avec 3000 imagettes.

En outre, lorsque le nombre de données est élevé, les résultats obtenus par les différents types d'imagettes sont proches. Cette observation peut s'expliquer le fait que la granulosité est indépendante de sa couleur ou de son intensité globale. Notons qu'en utilisant la couleur, nous nous apercevons notamment sur la Figure 7.19, que l'utilisation des imagettes en couleur tend à extraire des zones granuleuses partout sur le champ à tort.

(a) (b) (c) (d)

FIGURE 7.19 – Zones granuleuses (marquées en gris) sur trois exemples de monnaies : (a) annotation manuelle faite par la société GENI ; (b) (c) (d) zones granuleuses détectées par le réseau entraîné en utilisant respectivement les imagettes en couleur (b), en niveaux de gris (c) et en niveaux de gris normalisés (d)

Dans les expériences précédentes, la distribution des données entre classes est toujours équilibrée ou quasi-équilibrée. Cependant, dans la base de données *20FrancsCOQ_Photos*, la surface totale des zones granuleuses $\boldsymbol{Z}_g$ est plus petite que celles des zones lisses $\boldsymbol{Z}_s$. Ainsi, l'utilisation des données équilibrées entre les classes rend $\boldsymbol{P}_g$ plus représentatif que $\boldsymbol{P}_s$. En effet, nous testons sur des bases qui contiennent plus d'imagettes lisses que d'imagettes granuleuses ($N_g < N_s$). Les résultats sont montrés à la Table 7.3. Selon les résultats, l'influence de la distribution des données est moins importante que celle de la taille de données.

Enfin, dans nos expériences, les vrais positifs sont des imagettes de test appartenant à $\boldsymbol{P}_g$ classifiées comme granuleuses, alors que les vrais négatifs sont celles appartenant à $\boldsymbol{P}_s$ classifiées en lisses. Au contraire, les faux positifs et les faux négatifs sont respectivement les imagettes dont les classifications ne correspondent pas à leurs annotations. La Figure 7.20 montre des exemples des imagettes correctement classifiées et mal classifiées en utilisant les zones annotées par l'expert comme vérité terrain. Cependant, l'expert de

Table 7.2 – Taux de classification (%) obtenus avec *AlexNet* en faisant du *Fine Tuning* en utilisant un nombre croissant d'imagettes annotées

N. img	**N_g**	**N_s**	**Couleur**	**Niveaux de gris**	**Niveaux de gris norm.**
2000	1000	1000	85,3	84,4	83,3
3000	1500	1500	86,2	84,2	82,8
5000	2500	2500	73,5	74,2	73,3
9000	4500	4500	76,5	75,7	75,6

Table 7.3 – Taux de classification (%) obtenus avec *AlexNet* en faisant du *Fine Tuning* en utilisant différentes distributions d'imagettes annotées

N. img	**N_g**	**N_s**	**Couleur**	**Niveaux de gris**	**Niveaux de gris norm.**
3000	500	2500	85,0	85,2	84,1
3000	1000	2000	85,3	84,4	83,3
3000	1500	1500	86,2	84,2	82,8

la société GENI précise que la première imagette annotée comme lisse à la Figure 7.20 (b) devraient être une imagette granuleuse mal annotée, alors que certains faux positifs ou faux négatifs semblent avoir raison mais qu'il faut vérifier sur les monnaies entières. La Figure 7.21 montre les exemples « ambigus » vus par nos classifieurs (taux de prédiction vers 50% par classe), sur lesquels nous trouvons différentes textures et différents niveaux d'uniformité.

Après échange avec la société GENI, nous avons décidé de ne pas retravailler les annotations. D'une part, les gradeurs n'ont pas l'habitude de marquer tous les « éléments inattendus ». La gradation se fait plutôt sur une impression globale de la monnaie. D'autre part, les « éléments inattendus », notamment pour ceux qui sont ténus, sont vraiment difficiles à distinguer sur une seule l'image. Ainsi, il est difficile d'avoir de bons résultats en s'appuyant sur des vérités terrains bruitées. Pour obtenir une annotation précise, il faudrait que le gradeur annote sur la photo en observant attentivement la monnaie réelle à l'aide d'une loupe. Cependant, un tel travail est extrêmement coûteux et peu réalisable. L'annotation « grossière » de la base de données *20FrancsCOQ_Photos* a déjà pris environ 15 jours.

Pour conclure, malgré les difficultés, la méthode proposée arrive à détecter des zones granuleuses qui correspondent approximativement à l'annotation faite par le gradeur. Cependant, à cause de la précision des annotations et sans doute du nombre de données limité, les classifieurs que nous avons optimisés sur la base d'apprentissage en utilisant une partie des données annotées semblent « sur-apprendre » et ne pas trop généraliser à de nouvelles données. L'utilisation des imagettes en niveaux de gris normalisés semble réduire le surapprentissage sur les données actuelles. Les résultats sont encourageants selon les experts de la société GENI. Par exemple, à la Figure 7.19, même si les zones granuleuses

(a) (b)

FIGURE 7.20 – Exemples d'imagettes de test (imagettes annotées granuleuses encadrées en bleu ; imagettes annotées lisses encadrées en rouge) : (a) imagettes bien classifiées ; (b) imagettes mal classifiées

FIGURE 7.21 – Exemples d'imagettes « ambiguës »

détectées ne sont pas exactement aux mêmes endroits que l'annotation, elles sont dans des endroits similaires. En outre, la quantité relative de zones granuleuses correspond bien à celle de l'œil du gradeur.

Hormis les contraintes techniques, l'annotation des zones granuleuses est aussi un élément subjectif. Le gradeur qui a annoté la base de données *20FrancsCOQ_Photos* ne souhaite pas que son annotation soit diffusée. En fait, l'objectif industriel de cette étude n'est pas d'avoir une détection qui correspond au mieux à l'annotation manuelle, mais plutôt une détection ouverte sur laquelle les gradeurs professionnels seront largement d'accord. L'annotation manuelle, dans cette première étude de la gradation des monnaies, sert d'indication pour entraîner la machine n'ayant aucune connaissance mais n'a pas vocation à être diffusée.

7.3 Conclusion

Dans ce chapitre, nous avons proposé une première méthode automatique pour grader les monnaies bien conservées à partir d'images. La monnaie à grader est recalée à la monnaie de référence de son type monétaire. Cette référence est obtenue grâce à la moyenne des monnaies recalées pour localiser les contours des « éléments attendus » en commun. Les « éléments inattendus » significatifs sont détectés par comparaison avec la monnaie de référence. La principale difficulté est de détecter les « éléments inattendus » ténus qui sont complètement mélangés dans la texture de l'arrière-plan. Grâce au *Fine Tuning* sur un modèle performant du *Deep Learning*, *Alexnet*, nous avons pu détecter de manière approximative ces « éléments inattendus » ténus comme des zones granuleuses.

Le résultat de notre méthode fournit un Guide de Gradation (GG) qui donne une quantification des « éléments inattendus », plus précisément, le nombre et la surface des « éléments inattendus » significatifs et le pourcentage des surfaces des zones granuleuses sur différentes zones (sur les motifs importants ou sur le champ) de la monnaie. Ce résultat sert de guide indicatif et relativement objectif pour le gradeur. Le grade numérique montré aux collectionneurs sera une somme pondérée calculée à partir de GG. La société GENI ne nous a pas fourni de monnaies gradées avec un score numérique car selon eux, les grades attribués de manière conventionnelle sont probablement incorrects. Bien que l'association de GG à un grade numérique reste à étudier, il n'y a pas de raison de donner un grade plus élevé à une monnaie, qu'à une autre, si ses « éléments inattendus » sont plus nombreux que ceux de l'autre monnaie. Par ailleurs, dans le cas où le collectionneur met en doute le grade attribué à sa monnaie, au lieu de proposer simplement un grade plus élevé que le précédent pour lui fait plaisir, nous pouvons vérifier le grade et lui expliquer de manière concrète et convaincante à l'aide de GG.

L'annotation des données est un problème crucial pour cette étude. Une annotation précise est un travail extrêmement coûteux dont la qualité est souvent perturbée par la fatigue et l'incertitude du gradeur. A cause de la subjectivité, une annotation manuelle ne sera jamais précise à 100%. Pour avoir des annotations relativement moins subjectives, nous pourrons faire annoter les données par différents experts de manière indépendante mais au prix d'une intervention humaine coûteuse. Nous pourrons également utiliser la méthode proposée pour raffiner l'annotation. Grâce à l'intelligence artificielle, nous pourrons peut-être trouver des critères « cachés » à partir d'une annotation « grossière ». Puis, le gradeur pourra réévaluer son annotation originale en prenant compte celle obtenue par notre méthode afin de proposer une annotation raffinée, et cette dernière nous permettra d'avoir des données d'apprentissage moins bruitées.

Le travail actuel a été validé sur la base de données *20FrancsCOQ_Photos* mais il faudra que nous testions notre méthode sur des bases de données plus grandes qui contiennent d'autres types monétaires. Ces bases de données seront construites au fur à mesure au sein de la société GENI.

Conclusion générale

Résumé et extension du travail

Au cours de cette thèse, nous avons développé chacune des étapes permettant de grader les monnaies à partir de leurs photos.

Segmentation des monnaies

Nous avons commencé par extraire l'objet d'intérêt dans l'image, dans notre contexte la monnaie. Dans la littérature, une segmentation « grossière » des monnaies à l'aide de techniques basiques du traitement d'images est suffisante, par exemple, pour la reconnaissance automatique des monnaies. Cependant, pour avoir des illustrations parfaites afin de les intégrer dans les catalogues de vente, une segmentation précise est nécessaire. Un tel détourage précis est en principe réalisé par les numismates « à la main » en utilisant, par exemple, un outil comme Photoshop.

Dans ce contexte, nous avons proposé une méthode de segmentation des monnaies de manière précise, pour des monnaies rondes, mais également pour des monnaies de forme régulière, comme les monnaies à dents (sinusoïdales) et les monnaies à pans (polygonales). Pour ce faire, nous avons conçu un modèle paramétrique qui peut présenter toutes les formes possibles des monnaies prises en considération. À partir d'une segmentation « grossière », qui est proche de la région effective de la monnaie grâce à une qualité de photo professionnelle, nous obtenons un modèle géométrique estimé dont les paramètres sont dans les voisinages des valeurs optimisées. En ce qui concerne l'optimisation du modèle, et donc l'ajustement du modèle à la forme géométrique parfaite de la monnaie « grossièrement » segmentée, l'approche proposée consiste à minimiser une énergie d'homogénéité. La segmentation précise est enfin obtenue par le modèle optimisé. En même temps, les paramètres optimisés du modèle décrivent la forme de la monnaie et sont utilisées par la suite pour la reconnaissance.

Extension

Le modèle proposé qui contient 6 paramètres permet de segmenter des monnaies en trois formes régulières courantes : ronde, polygonale et sinusoïdale, et celles avec un trou rond ou carré au milieu. Une extension de ce travail consisterait à étendre le modèle à d'autres formes régulières, par exemple des monnaies polygonales à pans courbés (*cf.* Figure 4.19) ou à coins arrondis (*cf.* Figure 4.20). Dans ce cas-là, nous pourrions ajouter des paramètres supplémentaires pour décrire le rayon de la courbure du pan et du coin. Avec des connaissances *a priori*, nous pourrions également modéliser la striure ou la cannelure de la bordure des monnaies quasi-rondes. Cependant, augmenter la complexité du modèle imposerait de proposer de nouvelles stratégies pour l'optimiser de manière robuste et efficace.

Il faut mentionner qu'une modélisation précise ne s'applique que sur les monnaies « modernes » qui ont des formes géométriques données et dont les numismates souhaitent extraire les formes parfaites. Pour les monnaies anciennes de forme irrégulière et les monnaies avec un gros défaut sur la tranche (*cf.* Figure 4.17 (c)), la méthode n'est pas adaptée. Ainsi, il pourrait être intéressant d'arriver à distinguer automatiquement les monnaies « modélisables » et les monnaies « non-modélisables » afin d'appliquer des solutions de segmentation respectives.

Identification du type monétaire

Après avoir segmenté la monnaie à partir de la photo, nous avons proposé une méthode permettant d'apparier la monnaie segmentée au type monétaire le plus proche dans une base de références. La similarité calculée entre la monnaie requête et la monnaie de référence est basée, d'une part, sur des caractéristiques locales des motifs en relief, et d'autre part, sur les motifs en relief entre les caractéristiques locales appariées, autrement dit des caractéristiques semi-globales. Profitant du fait que dans les photos professionnelles, les orientations des monnaies sont très similaires, nous n'apparions que les points ayant des caractéristiques locales similaires dans des positions similaires. Nous avons également proposé une mesure de similarité semi-globale basée sur la construction de deux graphes quasi-identiques qui représentent la configuration géométrique intrinsèque des caractéristiques locales appariées. Selon les expériences menées, la méthode proposée est généralement fiable et robuste pour l'identification du type monétaire, même si les monnaies requêtes sont usées. Par ailleurs, dans le cas où les types monétaires à identifier ont une similarité-interclasse élevée, autrement dit les caractéristiques locales sont très similaires, le score de similarité semi-global que nous avons proposé est robuste et permet de distinguer deux monnaies de types différents s'il existe au moins de légères nuances dans leurs motifs.

Extension

Notons toutefois que cette méthode n'est pas invariante à la rotation. En effet, nous avons supposé que toutes les monnaies considérées étaient orientées de la même manière, ce qui correspond à notre cas. Bien que cette méthode soit adaptée à l'application réelle, il pourrait être intéressant d'étendre la méthode proposée à n'importe quelle position de la monnaie. Au lieu de rechercher l'appariement des caractéristiques locales dans un voisinage, nous pourrions libérer un peu la contrainte géométrique en recherchant dans une zone plus grande.

Par ailleurs, la méthode proposée n'utilise pas d'apprentissage automatique par manque de données pertinentes, notamment pour les monnaies rares. Dans le cadre de notre méthode, il suffit d'un exemple par type monétaire. Cependant, dans le contexte réel, si la

base de références ne contient pas le type monétaire correspondant à la monnaie requête, le système ne donnera en sortie que la monnaie la plus proche au sens des similarités proposées. Les critères de rejet restent donc à étudier. Ensuite, dans les expériences menées, les monnaies de référence sont des monnaies réelles dont les photos sont de même qualité que les monnaies requêtes. L'une des extensions de ce travail serait d'étudier les résultats obtenus à l'aide de monnaies de référence construites comme des monnaies moyennes, des monnaies synthétiques, des monnaies dessinées, *etc.* En effet, la base de données des images de monnaies de la société GENI étant encore en cours de construction, pour le moment la base de références de photos professionnelles pour l'identification du type monétaire est loin d'être complète. Il sera intéressant d'utiliser les références « temporaires » trouvées ailleurs pour réaliser l'identification du type monétaire avec un taux de réussite acceptable. Puis, nous pourrions remplacer les références par des photos professionnelles au fur à mesure de ma constitution de la base de données. Pour les types monétaires dont nous avons déjà assez de photos professionnelles, nous pouvons construire une référence « parfaite » si elle donne une meilleure performance. En outre, la référence « parfaite » d'un type monétaire sera utilisée pour la gradation.

Enfin, avec l'augmentation du nombre de données dans la base de références, une recherche exhaustive est coûteuse en termes de temps de calcul. L'idée serait d'utiliser les connaissances *a priori* comme le poids (car le système de prise de photos de la société GENI devrait également peser la monnaie en même temps) et les caractéristiques de forme obtenues lors de la segmentation pour présélectionner des types monétaires potentiels. Il pourrait alors être intéressant de définir des stratégies de réduction du nombre de types monétaires avant même de faire l'analyse des images.

Reconnaissance des chiffres extraits des monnaies

Dans l'objectif d'identifier les types monétaires dont les motifs sont extrêmement similaires, mais dont les millésimes sont différents, nous avons proposé une méthode de reconnaissance des chiffres extraits des monnaies afin de reconnaître les millésimes des monnaies. Pour ce faire, en supposant qu'un millésime contient toujours quatre chiffres, nous recherchons d'abord la zone de la date dans la monnaie à l'aide d'une fenêtre glissante en comparant avec un histogramme normalisé, et nous la découpons en quatre imagettes de chiffres. Grâce à des positions et des formes relativement prédictibles concernant la zone de la date, nous avons réussi à obtenir des imagettes de chiffres avec des bruits acceptables, même sur des monnaies usées. La principale difficulté posée par la reconnaissance des caractères extraits des monnaies est un premier plan difficile à segmenter puisque qu'il est de la même couleur que l'arrière-plan. Afin de pallier ce problème, nous convertissons l'imagette de chiffre en carte des gradients qui contient les contours du chiffre. Puis, à partir de la carte des gradients, nous extrayons l'ensemble des composantes connexes

(bulles) dans l'enveloppe convexe du chiffre qui n'appartiennent pas au chiffre sur une image intermédiaire que nous appelons carte de bulles. Ensuite, nous filtrons cette carte de bulles. Le chiffre est reconnu à l'aide de caractéristiques sur les bulles grâce à des connaissances *a priori* sur la « topologie » des dix chiffres arabes. Deux stratégies ont été proposées pour reconnaître les chiffres par la carte de bulles filtrée : la première se base sur un système de votes qui calcule la somme pondérée des bulles retenues ; la seconde est réalisée par comparaison des graphes de Reeb construits à l'aide des traitements sur la carte des bulles. Selon les expériences menées, notre méthode a obtenu un taux de réussite élevé sur les chiffres extraits des monnaies avec des contours nets. Cependant, lorsque les chiffres sont extraits de monnaies usées sur lesquelles nous ne pouvons pas extraire des contours corrects sur la carte des gradients, le taux de reconnaissance chute. La version utilisant le système de votes est beaucoup plus robuste que celle utilisant le graphe de Reeb si les données sont bruitées. Il faut mentionner que notre méthode ne nécessite pas d'apprentissage machine. Elle s'adapte bien au contexte réel parce que pour le moment la société GENI n'a pas beaucoup de données et la plupart des monnaies à traiter sont neuves ou quasi-neuves. Pendant cette étude, nous avons également publié la base de données *CoinNUMS* qui contient des chiffres extraits des monnaies en espérant que cette base de données attirera l'attention de chercheurs spécialisés en reconnaissance de caractères.

Extension

Bien que la méthode proposée soit robuste à certains types de bruits qui ne modifient pas la « topologie » supposée du chiffre, elle n'a pas été développée pour les chiffres avec des polices « exotiques », notamment ceux qui ont des creux attendus fermés ou des trous attendus ouverts. Par ailleurs, la performance de notre méthode dépend largement de la qualité de la carte des gradients obtenue. Une autre limite est que la méthode développée pour les dix chiffres est difficile à transposer aux lettres ou autres caractères étrangers.

Pour obtenir une meilleure performance et pour étendre à d'autres caractères, il existe plusieurs directions que nous pourrions étudier. Premièrement, pour traiter des chiffres avec des polices « exotiques », nous pourrions inclure des caractéristiques de contours (plat ou courbé). Deuxièmement, il serait intéressant d'étudier des méthodes plus sophistiquées pour extraire une carte des gradients « propre », même sur les monnaies usées. Troisièmement, comme le *Deep Learning* a donné des résultats impressionnants en reconnaissance de caractères, nous pourrions lancer des expériences afin de trouver un modèle optimisé et les meilleures données d'entrée (imagettes originales, cartes de bulles, caractéristiques extraites, *etc.*) du réseau pour ce problème.

Gradation des monnaies

Finalement, nous avons élaboré une méthode permettant l'aide à la gradation de monnaies bien conservées en quantifiant les « éléments inattendus ». Pour ce faire, la monnaie segmentée est comparée avec une monnaie de référence « parfaite » de son type monétaire identifié après recalage. La monnaie « parfaite » est un masque binaire qui ne contient que les non-uniformités attendues, et qui est obtenue par moyenne des monnaies recalées. Nous cherchons les « éléments inattendus » sur certaines zones d'intérêt de la monnaie.

A cause de la complexité des « éléments inattendus » présents dans l'image, nous les avons séparés en « éléments inattendus » significatifs et « éléments inattendus » ténus. Pour les « éléments inattendus » significatifs, nous les détectons en combinant une détection des gradients avec un seuillage, en supposant qu'ils sont sombres ou présentent des changements abrupts sur des zones quasi-uniformes. En ce qui concerne les « éléments inattendus » ténus qui ne peuvent pas être repérés individuellement, nous détectons les zones qui contiennent ces « éléments inattendus » ténus par imagettes. Cette partie est très difficile car il est impossible pour des non-experts de distinguer les zones contenant les « éléments inattendus » ténus et des zones contenant des non-uniformités attendues comme la texture du métal. Même les experts numismates ne sont pas toujours sûrs de leurs annotations. À l'aide de données grossièrement annotées par la société GENI, nous utilisons le *Deep Learning* avec un *Fine Tuning* du modèle *AlexNet*. À la fin, le résultat de gradation, Guide de Gradation (GG), donne la surface et le nombre « d'éléments inattendus » significatifs et le pourcentage des zones qui contiennent des « éléments inattendus » ténus, en fonction de l'importance de leurs endroits. Selon les expériences menées, nous avons obtenu des résultats qui ressemblent à l'annotation de l'expert. Cependant, nous avons également repéré certaines données annotées discutables en analysant les résultats.

Extension

La gradation des monnaies par analyse d'images est un domaine complètement ouvert qui est peu étudié dans la littérature.

Une extension directe de la méthode proposée serait de grader des monnaies usées en rajoutant une quantification des « éléments attendus ». Cette étape serait également réalisée par comparaison avec une monnaie « parfaite ». Il faut noter que certains « éléments inattendus » sur les contours des reliefs n'ont pas été pris en compte par la méthode présentée. Cependant, ils seront détectés lors de la quantification des « éléments attendus » par une analyse permettant de distinguer les « éléments attendus » manquants et les « éléments inattendus » introduits.

Selon l'expert, un défaut de la profondeur des « éléments inattendus » significatifs est beaucoup plus grave qu'un défaut de leur taille. Cependant, il est difficile d'obtenir des informations de profondeur sur des images 2D. De nouvelles approches d'acquisition

seraient étudiées. Nous pourrions utiliser un scanner 3D, comme dans le travail de Debroutelle et al. (2017) sur les tessons céramiques, mais le temps d'acquisition et la taille des données semblent moins correspondre à la pratique. Par ailleurs, il faut également étudier si la précision permet de détecter les défauts creusés. Une proposition basée sur le système de prise de photos déjà développé est de photographier une monnaie deux fois avec des angles de lumière légèrement différents. L'hypothèse est que la couleur sera différente pour les « éléments inattendus » creusés, mais restera presque la même pour les « éléments inattendus » plats.

Pour les « éléments inattendus » ténus, l'annotation de manière précise sur les photos est un véritable défi, même pour les experts. En fait, la granulosité des zones de la monnaie n'est pas une observation binaire. La classification en deux classes est une solution approximative qui génère l'incertitude de l'expert sur certaines zones ambiguës. Nous pourrions tester une classification en trois classes, par exemple les zones granuleuses, les zones lisses et les zones ambiguës, pour regarder si les résultats obtenus correspondent mieux à l'annotation de l'expert.

Enfin, le grade d'une monnaie ne dépend pas seulement des « éléments attendus » et des « élément inattendus ». D'autres facteurs contributifs, comme la patine, seront étudiés dans le futur. Il faut que nous étudiions également l'association de tous les facteurs contributifs quantifiés à un grade numérique dans l'échelle de Sheldon.

Perspectives pour la numismatique par analyse d'images

Perspectives au sein de la société GENI

La construction de bases de données avec des images pertinentes est cruciale pour toutes les étapes étudiées dans cette thèse. Au cours de cette thèse, nous avons travaillé avec le photographe et les experts numismates de la société GENI pour développer un système de prise de vues photographiques fiable et performant. Grâce aux images de qualité acquises par ce système, nos méthodes proposées peuvent atteindre une performance décente dans l'application réelle. Cependant, à cause du manque d'échantillons de monnaie et du temps pour le développement de ce système, nous n'avons pas encore de base de données d'images de taille exploitable pour toutes les étapes. Ces bases de données d'images seront complétées au fur à mesure.

La société GENI a développé plusieurs bases de données numismatiques relativement complètes comme celles des monnaies « modernes » françaises (plus de 8000 références) ainsi que celles des euros (plus de 10000 références). Il est possible d'accéder à ces bases de données d'informations grâce à un site internet ainsi qu'une application smartphone. Cependant, les images prises par le système de prise de photos n'ont pas été incluses.

Dans le futur, nous allons travailler avec la société GENI pour définir un protocole

standard sur la base de données d'images concernant le prétraitement, l'annotation, le format, *etc.* qui permettra non seulement de faciliter nos traitements automatiques, mais aussi de combiner les bases de données numismatiques existantes de l'entreprise. Dès que les bases de données d'informations et la base d'images seront combinées, nous pourrons profiter des descriptions détaillées sur les types monétaires comme connaissances *a priori* pour obtenir de meilleures performances.

L'implémentation des algorithmes dans l'environnement réel est importante pour que les experts numismates puissent profiter directement de ce que nous avons développé au cours de cette thèse. Il s'agit de développer un logiciel interne ou des bibliothèques compatibles avec le logiciel de gestion des données numismatiques utilisé par la société GENI. Les codes seront optimisés pour diminuer le temps de calcul.

Perspectives pour le traitement des monnaies anciennes

Le travail présenté dans cette thèse se concentre sur les monnaies « modernes ». Dans la communauté des chercheurs en traitement d'images, ce sont cependant plus les monnaies anciennes qui attirent l'intérêt. Par ailleurs, selon les numismates de la société GENI, le marché de la numismatique antique est passé devant le marché des monnaies « modernes » en termes de volume et de poids économique.

Concernant l'identification du type monétaire, les défis sont multiples car ces monnaies ont été frappées au marteau et comportent des niveaux d'usure élevés. Par ailleurs, chaque exemplaire est unique. Il n'existe pas de standardisation comme dans le cas des monnaies « modernes ». Par conséquent, les difficultés d'authentification ne peuvent que se multiplier. Concernant la gradation, une monnaie ancienne ne se grade pas du tout de la même façon qu'une monnaie moderne. La seule gradation professionnelle qui a été tentée est celle faite par la société américaine NGC, mais les résultats ne correspondent pas aux attentes des collectionneurs. Cependant, grâce à la frappe au marteau « personnalisée », l'identification du coin monétaire utilisé lors de la frappe pourraient permettre d'identifier l'origine et l'évolution de la frappe. Ce sujet attire non seulement les numismates et les collectionneurs, mais aussi les historiens, les économistes, les géographes, les muséologues *etc.*, qui souhaitent étudier les mondes antiques (projet ERC SILVER (*Silver Isotopes and the Rise of Money*)).

L'extension de notre travail aux monnaies anciennes impliquerait donc des travaux sur les sujets suivants :

- Identification des monnaies anciennes frappées par les mêmes coins monétaires : la variation des motifs des monnaies anciennes appartenant à un même type monétaire dépend des coins monétaires utilisés. Nous pourrions étendre notre méthode d'identification aux monnaies anciennes frappées par le même coin afin de relier les monnaies aux coins monétaires, voire aux ateliers où elles ont été frappées.

- Détection de l'usure des coins : pour les monnaies anciennes, les « éléments inattendus » pourraient provenir aussi des coins. Il s'agira de distinguer les creux inattendus provenant de l'usure de la monnaie après la frappe de ceux provenant de la frappe afin de tracer l'utilisation des coins et de dater les monnaies associées.

- Reconnaissance des caractères extraits des monnaies anciennes : les légendes sont souvent plus discriminantes que les motifs pour identifier le type monétaire ancien. À partir des premières études de l'état de l'art sur cette problématique, l'utilisation du *Deep Learning* pourrait se révéler avantageuse en supposant que nous ayons suffisamment de données pour entraîner le réseau.

- Classification des motifs principaux sur les monnaies anciennes : il s'agit de la classification automatique d'objets, par exemple, des différentes chouettes sur les monnaies anciennes (*cf.* Figure ci-dessous). Une difficulté est que l'œil humain a du mal à distinguer les objets d'intérêt de part le fait que les reliefs sont extrêmement bruités. Ce problème est plus difficile encore que la reconnaissance des caractères extraits des monnaies. Cette étude permettrait de construire automatiquement un inventaire des motifs utilisés dans les monnaies anciennes (les gémeaux, les animaux, les navires, les têtes, *etc.*) en fonction de la géographie et de la culture.

Différentes chouettes sur les monnaies anciennes

Bibliographie

Adams, Rolf, et Leanne Bischof. 1994. « Seeded region growing ». *IEEE Transactions on Pattern Analysis and Machine Intelligence* **16** (*6*) : page(s): 641–647.

Agrawal, Motilal, Kurt Konolige et Morten Rufus Blas. 2008. « Censure : Center surround extremas for realtime feature detection and matching ». In *European Conference on Computer Vision,* page(s): 102–115.

Anwar, Hafeez, Sebastian Zambanini et Martin Kampel. 2013. « Supporting ancient coin classification by image-based reverse side symbol recognition ». In *International Conference on Computer Analysis of Images and Patterns,* page(s): 17–25.

Anwar, Hafeez, Sebastian Zambanini et Martin Kampel. 2014. « A rotation-invariant bag of visual words model for symbols based ancient coin classification ». In *IEEE International Conference on Image Processing,* page(s): 5257–5261.

Anwar, Hafeez, Sebastian Zambanini et Martin Kampel. 2015a. « Coarse-grained ancient coin classification using image-based reverse side motif recognition ». *Machine Vision and Applications* **26** (*2-3*) : page(s): 295–304.

Anwar, Hafeez, Sebastian Zambanini et Martin Kampel. 2015b. « Efficient scale-and rotation-invariant encoding of visual words for image classification ». *IEEE Signal Processing Letters* **22** (*10*) : page(s): 1762–1765.

Anwar, Hafeez, Sebastian Zambanini, Martin Kampel et Klaus Vondrovec. 2015. « Ancient coin classification using reverse motif recognition : Image-based classification of roman republican coins ». *IEEE Signal Processing Magazine* **32** (*4*) : page(s): 64–74.

Arandjelovic, Ognjen. 2010. « Automatic attribution of ancient Roman imperial coins ». In *IEEE Conference on Computer Vision and Pattern Recognition,* page(s): 1728–1734.

Arandjelović, Ognjen. 2012. « Reading ancient coins : automatically identifying denarii using obverse legend seeded retrieval ». In *European Conference on Computer Vision,* page(s): 317–330.

Atherton, Tim J, et Darren J Kerbyson. 1999. « Size invariant circle detection ». *Image and Vision computing* **17** (*11*) : page(s): 795–803.

Ballard, Dana H. 1981. « Generalizing the Hough transform to detect arbitrary shapes ». *Pattern Recognition* **13** (*2*) : page(s): 111–122.

Bassett, Rick, Ping Gallivan, Xiang Gao, Eric Heinen et Akarsh Sakalaspur. 2002. « Development of an automated coin grader : a progress report ». In *Mid-Atlantic Student Workshop on Programming Languages and Systems.*

Bay, Herbert, Tinne Tuytelaars et Luc Van Gool. 2006. « Surf : Speeded up robust features ». In *European conference on computer vision,* page(s): 404–417.

Belongie, Serge, Jitendra Malik et Jan Puzicha. 2002. « Shape matching and object recognition using shape contexts ». *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24** (*4*) : page(s): 509–522.

Benligiray, Burak, Halil Ibrahim Cakir, Cihan Topal et Cuneyt Akinlar. 2015. « Counting Turkish Coins with a Calibrated Camera ». In *International Conference on Image Analysis and Processing,* page(s): 216–226.

Beucher, Serge, et Fernand Meyer. 1992. « The morphological approach to segmentation : the watershed transformation ». *Optical Engineering* **34** : page(s): 433–433.

Boykov, Yuri Y, et M-P Jolly. 2001. « Interactive graph cuts for optimal boundary & region segmentation of objects in ND images ». In *IEEE International Conference on Computer Vision,* 1 : page(s): 105–112.

Calonder, Michael, Vincent Lepetit, Christoph Strecha et Pascal Fua. 2010. « Brief : Binary robust independent elementary features ». In *European conference on computer vision,* page(s): 778–792.

Campos, Teo de, Bodla Rakesh Babu et Manik Varma. 2009. « Character recognition in natural images ». In *International Conference on Computer Vision Theory and Applications.*

Cerutti, Guillaume, Laure Tougne, Antoine Vacavant et Didier Coquin. 2011. « A parametric active polygon for leaf segmentation and shape estimation ». In *International Symposium on Visual Computing,* page(s): 202–213.

Chan, Tony F, et Luminita A Vese. 2001. « Active contours without edges ». *IEEE Transactions on Image Processing* **10** (*2*) : page(s): 266–277.

Chen, Huahua. 2010. « Chinese coin recognition based on unwrapped image and rotation invariant template matching ». In *nternational Conference on Intelligent Networks and Intelligent Systems,* page(s): 5–7.

Cheng, Yizong. 1995. « Mean shift, mode seeking, and clustering ». *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17** (*8*) : page(s): 790–799.

Ciregan, Dan, Ueli Meier et Jürgen Schmidhuber. 2012. « Multi-column deep neural networks for image classification ». In *IEEE conference on Computer Vision and Pattern Recognition,* page(s): 3642–3649.

Comaniciu, Dorin, et Peter Meer. 2002. « Mean shift : A robust approach toward feature space analysis ». *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24** (*5*) : page(s): 603–619.

Coustaty, Mickael, Rudolf Pareti, Nicole Vincent et Jean-Marc Ogier. 2011. « Towards historical document indexing : extraction of drop cap letters ». *International Journal on Document Analysis and Recognition (IJDAR)* **14** (*3*) : page(s): 243–254.

Crain, Louis M. 1993. *Automated coin grading system.* US Patent 5,224,176, juin.

Cuhaj, George, et Thomas Micheal. 2015. *Standard Catalog of Word coins, 1801-1900.* Krause Publications.

Debroutelle, Teddy, Sylvie Treuillet, Aladine Chetouani, Matthieu Exbrayat, Lionel Martin et Sebastien Jesset. 2017. « Automatic classification of ceramic sherds with relief motifs ». *Journal of Electronic Imaging* **26** (*2*) : page(s): 023010.

Desrousseaux, Stéphane, et Michel Prieur. 2014. *Le Franc 10 : les Monnaies Françaises.* Les Chevau-légers.

Dong, Liju, Ge Yu, Philip Ogunbona et Wanqing Li. 2008. « An efficient iterative algorithm for image thresholding ». *Pattern Recognition Letters* **29** (*9*) : page(s): 1311–1316.

Falcão, Alexandre X, Jorge Stolfi et Roberto de Alencar Lotufo. 2004. « The image foresting transform : Theory, algorithms, and applications ». *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26** (*1*) : page(s): 19–29.

Fare, Callum, et Ognjen Arandjelović. 2017. « Ancient Roman Coin Retrieval : A Systematic Examination of the Effects of Coin Grade ». In *European Conference on Information Retrieval,* page(s): 410–423.

Felzenszwalb, Pedro F, Ross B Girshick, David McAllester et Deva Ramanan. 2010. « Object detection with discriminatively trained part-based models ». *IEEE Transactions on Pattern Analysis and Machine Intelligence* **32** (*9*) : page(s): 1627–1645.

Felzenszwalb, Pedro F, et Daniel P Huttenlocher. 2004. « Efficient graph-based image segmentation ». *International Journal of Computer Vision* **59** (*2*) : page(s): 167–181.

Felzenszwalb, Pedro F, et Daniel P Huttenlocher. 2005. « Pictorial structures for object recognition ». *International Journal of Computer Vision* **61** (*1*) : page(s): 55–79.

Fukumi, Minoru, Sigeru Omatu, Fumiaki Takeda et Toshihisa Kosaka. 1992. « Rotation-invariant neural pattern recognition system with application to coin recognition ». *IEEE Transactions on Neural Networks* **3** (*2*) : page(s): 272–279.

Glasbey, Chris A. 1993. « An analysis of histogram-based thresholding algorithms ». *CVGIP : Graphical Models and Image Processing* **55** (*6*) : page(s): 532–537.

Grady, Leo. 2006. « Random walks for image segmentation ». *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28** (*11*) : page(s): 1768–1783.

Haperin, James. 1990. *How to Grade U.S. Coins.*

Heath, Mike, Sudeep Sarkar, Thomas Sanocki et Kevin Bowyer. 1996. « Comparison of edge detectors : a methodology and initial study ». In *IEEE International Conference on Computer Vision,* page(s): 143–148.

Hollard, Dominique. 1991. « L'illustration numismatique au XIXe siècle ». *Revue numismatique* **6** (*33*) : page(s): 7–42.

Huber, Reinhold, Herbert Ramoser, Konrad Mayer, Harald Penz et Michael Rubik. 2005. « Classification of coins using an eigenspace approach ». *Pattern Recognition Letters* **26** (*1*) : page(s): 61–75.

Huber-Mörk, Reinhold, Sebastian Zambanini, Maia Zaharieva et Martin Kampel. 2011. « Identification of ancient coins based on fusion of shape and local features ». *Machine Vision and Applications* **22** (*6*) : page(s): 983–994.

Hundelshausen, Felix von, et Rahul Sukthankar. 2012. « D-Nets : beyond patch-based image descriptors ». In *IEEE Conference on Computer Vision and Pattern Recognition,* page(s): 2941–2948.

Joulin, Armand, Francis Bach et Jean Ponce. 2010. « Discriminative clustering for image co-segmentation ». In *IEEE Conference on Computer Vision and Pattern Recognition,* page(s): 1943–1950.

Judkins, Maggie, et Thomas Micheal. 2016a. *Standard Catalog of Word coins, 1901-2000.* Krause Publications.

Judkins, Maggie, et Thomas Micheal. 2016b. *Standard Catalog of Word coins, 2001-Date.* Krause Publications.

Kampel, Martin, et Maia Zaharieva. 2008. « Recognizing ancient coins based on local features ». In *International Symposium on Visual Computing,* page(s): 11–22.

Karatzas, Dimosthenis, Lluis Gomez-Bigorda, Anguelos Nicolaou, Suman Ghosh, Andrew Bagdanov, Masakazu Iwamura, Jiri Matas, Lukas Neumann, Vijay Ramaseshan Chandrasekhar, Shijian Lu et al. 2015. « ICDAR 2015 competition on robust reading ». In *International Conference on Document Analysis and Recognition,* page(s): 1156–1160.

Karatzas, Dimosthenis, Faisal Shafait, Seiichi Uchida, Masakazu Iwamura, Lluis Gomez i Bigorda, Sergi Robles Mestre, Joan Mas, David Fernandez Mota, Jon Almazan Almazan et Lluis Pere De Las Heras. 2013. « ICDAR 2013 robust reading competition ». In *International Conference on Document Analysis and Recognition,* page(s): 1484–1493.

Kass, Michael, Andrew Witkin et Demetri Terzopoulos. 1988. « Snakes : Active contour models ». *International Journal of Computer Vision* **1** (*4*) : page(s): 321–331.

Kavelar, Albert, Sebastian Zambanini et Martin Kampel. 2012. « Word detection applied to images of ancient roman coins ». In *International Conference on Virtual Systems and Multimedia,* page(s): 577–580.

Kavelar, Albert, Sebastian Zambanini et Martin Kampel. 2014. « Reading the legends of Roman Republican coins ». *Journal on Computing and Cultural Heritage* **7** (*1*) : page(s): 5.

Kavelar, Albert, Sebastian Zambanini, Martin Kampel, Klaus Vondrovec et Kathrin Siegl. 2013. « The ILAC-project : supporting ancient coin classification by means of image analysis ». In *XXIV International CIPA Symposium.*

Khashman, Adnan, Boran Sekeroglu et Kamil Dimililer. 2006. « Intelligent coin identification system ». In *IEEE International Conference on Control Applications,* page(s): 1226–1230.

Khashman, Adnan, Boran Sekeroglu et Kamil Dimililer. 2007. « Rotated coin recognition using neural networks ». *Analysis and Design of Intelligent Systems using Soft Computing Techniques :* page(s): 290–297.

Kim, Jongpil, et Vladimir Pavlovic. 2014. « Improving ancient roman coin recognition with alignment and spatial encoding ». In *Workshop at the European Conference on Computer Vision,* page(s): 149–164.

Krizhevsky, Alex, Ilya Sutskever et Geoffrey E Hinton. 2012. « Imagenet classification with deep convolutional neural networks ». In *Advances in Neural Information Processing Systems,* page(s): 1097–1105.

LeCun, Yann, Léon Bottou, Yoshua Bengio et Patrick Haffner. 1998. « Gradient-based learning applied to document recognition ». *Proceedings of the IEEE* **86** (*11*) : page(s): 2278–2324.

Leutenegger, Stefan, Margarita Chli et Roland Y Siegwart. 2011. « BRISK : Binary robust invariant scalable keypoints ». In *IEEE International Conference on Computer Vision,* page(s): 2548–2555.

Liu, Ce, Jenny Yuen et Antonio Torralba. 2011. « Sift flow : Dense correspondence across scenes and its applications ». *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33** (*5*) : page(s): 978–994.

Long, Jonathan, Evan Shelhamer et Trevor Darrell. 2015. « Fully convolutional networks for semantic segmentation ». In *IEEE Conference on Computer Vision and Pattern Recognition,* page(s): 3431–3440.

Lowe, David G. 2004. « Distinctive image features from scale-invariant keypoints ». *International Journal of Computer Vision* **60** (*2*) : page(s): 91–110.

Lucas, Simon M, Alex Panaretos, Luis Sosa, Anthony Tang, Shirley Wong et Robert Young. 2003. « ICDAR 2003 robust reading competitions ». In *International Conference on Document Analysis and Recognition,* page(s): 682–687.

Maaten, LJP van der, et PJ Boon. 2006. « Reliable Classification of Partially occluded Coins ». In *MUSCLE CIS Coin Competition Workshop.*

Marr, David, et Ellen Hildreth. 1980. « Theory of edge detection ». *Proceedings of the Royal Society of London* **207** (*1167*) : page(s): 187–217.

Merton, Henry A. 1990. *Method and system for objectively grading and identifying coins.* US Patent 4,899,392, février.

Mivcuvsik, Branislav, et Allan Hanbury. 2006. « Automatic image segmentation by positioning a seed ». In *European Conference on Computer Vision,* page(s): 468–480.

Mortensen, Eric N, et William A Barrett. 1995. « Intelligent scissors for image composition ». In *Conference on Computer Graphics and Interactive Techniques,* page(s): 191–198.

Nölle, Michael, Harald Penz, Michael Rubik, Konrad Mayer, Igor Holländer et Reinhard Granec. 2003. « Dagobert-a new coin recognition and sorting system ». In *International Conference on Digital Image Computing-Techniques and Applications.*

Nölle, Michael, Michael Rubik et Allan Hanbury. 2006. « Results of the MUSCLE CIS coin competition 2006 ». In *MUSCLE CIS Coin Competition Workshop,* page(s): 1–5.

Otsu, Nobuyuki. 1975. « A threshold selection method from gray-level histograms ». *Automatica* **11** (*285-296*) : page(s): 23–27.

Pan, Xingyu, Kitti Puritat et Laure Tougne. 2014. « A new coin segmentation and graph-based identification method for numismatic application ». In *International Symposium on Visual Computing,* page(s): 185–195.

Pan, Xingyu, et Laure Tougne. 2016. « Topology-based character recognition method for coin date detection ». In *IEEE International Conference on Image Analysis and Processing,* t. 2.

Pan, Xingyu, et Laure Tougne. 2017. « A new database of digits extracted from coins with hard-to-segment foreground for optical character recognition evaluation ». *Frontiers in ICT* **4** : page(s): 9.

Pan, Xingyu, et Laure Tougne. 2018. « Image analysis and Deep Learning for aiding professional coin grading ». In *International Conference on Image Video Processing and Artificial Intelligence.*

Pavlidis, T, et SL Horowitz. 1974. « Picture segmentation by a directed split-and-merge procedure ». In *Proc. Second International Joint Conference on Pattern Recognition.*

Reisert, Marco, Olaf Ronneberger et Hans Burkhardt. 2007. « A fast and reliable coin recognition system ». In *Joint Pattern Recognition Symposium,* page(s): 415–424.

Rother, Carsten, Vladimir Kolmogorov et Andrew Blake. 2004. « Grabcut : Interactive foreground extraction using iterated graph cuts ». In *ACM Transactions on Graphics,* 23 : page(s): 309–314. 3.

Rublee, Ethan, Vincent Rabaud, Kurt Konolige et Gary Bradski. 2011. « ORB : An efficient alternative to SIFT or SURF ». In *IEEE International Conference on Computer Vision,* page(s): 2564–2571.

Sezgin, Mehmet, et al. 2004. « Survey over image thresholding techniques and quantitative performance evaluation ». *Journal of Electronic Imaging* **13** (*1*) : page(s): 146–168.

Shahab, Asif, Faisal Shafait et Andreas Dengel. 2011. « ICDAR 2011 robust reading competition challenge 2 : Reading text in scene images ». In *International Conference on Document Analysis and Recognition,* page(s): 1491–1496.

Shanbhag, Abhijit G. 1994. « Utilization of information measure as a means of image thresholding ». *Graphical Models and Image Processing* **56** (*5*) : page(s): 414–419.

Sheldon, William, Dorothy Paschal et Walter Breen. 1958. *Penny Whimsy.* Harper & Row.

Shen, Linlin, Sen Jia, Zhen Ji et Wen-Sheng Chen. 2009. « Statictics of Gabor features for coin recognition ». In *IEEE International Workshop on Imaging Systems and Techniques,* page(s): 295–298.

Shen, Linlin, Sen Jia, Zhen Ji et Wen-Sheng Chen. 2011. « Extracting local texture features for image-based coin recognition ». *IET Image Processing* **5** (*5*) : page(s): 394–401.

Shi, Jianbo, et Jitendra Malik. 2000. « Normalized cuts and image segmentation ». *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22** (*8*) : page(s): 888–905.

Shi, Jichuan, Nilanjan Ray et Hong Zhang. 2012. « Shape based local thresholding for binarization of document images ». *Pattern Recognition Letters* **33** (*1*) : page(s): 24–32.

Thome, Nicolas, Antoine Vacavant, Lionel Robinault et Serge Miguet. 2011. « A cognitive and video-based approach for multinational license plate recognition ». *Machine Vision and Applications* **22** (*2*) : page(s): 389–407.

Van Der Maaten, L, et E Postma. 2006. *Towards automatic coin classification.*

Van Der Maaten, Laurens J, et P Poon. 2006. « Coin-o-matic : A fast system for reliable coin classification ». In *MUSCLE CIS Coin Competition Workshop,* page(s): 7–18.

VC, Hough Paul. 1962. *Method and means for recognizing complex patterns.* US Patent 3,069,654, décembre.

Vincent, Luc, et Pierre Soille. 1991. « Watersheds in digital spaces : an efficient algorithm based on immersion simulations ». *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13** (*6*) : page(s): 583–598.

Yanowitz, Shimon D, et Alfred M Bruckstein. 1989. « A new method for image segmentation ». *Computer Vision, Graphics, and Image Processing* **46** (*1*) : page(s): 82–95.

Zaharieva, Maia, Reinhold Huber-Mörk, Michael Nölle et Martin Kampel. 2007. « On Ancient Coin Classification ». In *International Symposium on Virtual Reality, Archaeology and Cultural Heritage,* page(s): 55–62.

Zaharieva, Maia, Martin Kampel et Sebastian Zambanini. 2007. *Image based recognition of coins–an overview of the coins project.*

Zambanini, Sebastian, et Martin Kampel. 2009. « Robust Automatic Segmentation of Ancient Coins ». In *International Conference on Computer Vision Theory and Applications,* page(s): 273–276.

Zambanini, Sebastian, et Martin Kampel. 2011. « Automatic coin classification by image matching ». *International conference on Virtual Reality, Archaeology and Cultural Heritage :* page(s): 65–72.

Zambanini, Sebastian, et Martin Kampel. 2012. « Coarse-to-fine correspondence search for classifying ancient coins ». In *Asian Conference on Computer Vision,* page(s): 25–36.

Zambanini, Sebastian, et Martin Kampel. 2013. « A local image descriptor robust to illumination changes ». In *Scandinavian Conference on Image Analysis,* page(s): 11–21.

Zambanini, Sebastian, Albert Kavelar et Martin Kampel. 2013. « Improving ancient roman coin classification by fusing exemplar-based classification and legend recognition ». In *International Conference on Image Analysis and Processing,* page(s): 149–158.

Zambanini, Sebastian, Albert Kavelar et Martin Kampel. 2014. « Classifying ancient coins by local feature matching and pairwise geometric consistency evaluation ». In *International Conference on Pattern Recognition,* page(s): 3032–3037. IEEE.

Zhang, Zhao, William V Stoecker et Randy H Moss. 2000. « Border detection on digitized skin tumor images ». *IEEE Transactions on Medical Imaging* **19** (*11*) : page(s): 1128–1143.

Acronymes

AB Assez Beau. 16

ADD *Angle-distance-distribution*. 45, 46

ANN *Artificial Neural Networks*. 149, 151

B Beau. 14, 16

BE Belle épreuve. 13, 16

BRIEF *Binary Robust Independent Elementary Features*. 119

CNN *Convolutional Neural Networks*. 172, 173

DCSM *Deviation from Circular Shape Matching*. 48

DoG *Difference of Gaussian*. 37

DPM *Deformable Part Model*. 51

EADD *Edge-angle-distance-distribution*. 45

EDD *Edge-distance-distribution*. 45, 46

FAST *Features from Accelerated Segment Test*. 119, 120

FDC Fleur de Coin. 13, 14, 16

FF Facteur de Forme. 41

GG Guide de Gradation. 162, 173, 174, 182, 189

HoG *Histogram of oriented Gradients*. 49, 53

LBDH *locally-biased Directional Histogram*. 51

LBP *Local Binary Patterns*. 57

LIDRIC *local image descriptor robust to illumination changes*. 49, 50

LoG *Laplacien of Gaussian*. 37

ORB *Oriented FAST and Rotated BRIEF*. 119–122

SIFT *Scale-Invariant Feature Transform*. 49–51, 53, 57, 119, 120

SPL Splendide. 14, 16

SSD *Sum of Squared Difference*. 163

SSE *Similarity Score by graph Edges*. 129, 131–135

SSN *Similarity Score by graph Nodes*. 129, 131–133

SUP Superbe. 14, 16, 57

SVM *Support Vector Machine*. 53

TB Très Beau. 14, 16, 57

TB Très Très Beau. 14, 16, 57

TS Topologie Simplifiée. 147, 149

Glossaire

belle épreuve Qualité d'une monnaie la plus élevée qui caractérise une monnaie dont les motifs sont mats et le fond est comme un miroir. 13, 25, 207

champ Partie plate de la monnaie sur laquelle aucun motif n'a été gravé. 12, 24, 55, 161, 173, 174, 178, 182

coin Morceau de métal sur lequel est gravée en creux les motifs d'une monnaie. 10, 191, 192

coups de sac Chocs nombreux d'aspect bien particuliers causés par les chocs entre les monnaies stockées et transportées en sac. 12, 169

Essai Monnaie frappée en très faible quantité avec des coins neufs et destinée à présenter un nouveau type monétaire. 10, 61, 71, 105, 209, 210

Fleur de Coin État de conservation d'une monnaie qui doit être parfait et ne présenter aucun défaut visible. 13, 16

millésime Date d'émission d'une monnaie gravée sur un côté de la monnaie. 4, 12, 28, 29, 74, 76, 77, 79, 82, 83, 139–141, 153, 155, 187

velours de frappe Ridules à la surface d'une monnaie qui diminue au fur et à mesure par la circulation, en lui donnant un aspect « satiné » mat. 12, 169

Table des figures

1.1 Bourse numismatique à Paris, le 12 Octobre 2013 9
1.2 Monnaie de Paris, institut monétaire national de la France 10
1.3 Coins des monnaies en euro de la Finlande 11
1.4 Différentes pièces de collection : (a) Essai 5 Francs Napoléon II (légende « ESSAI » encadrée en rouge) ; (b) 2 Euros *D-Day* commémorative ; (c) jeton de jeux 11
1.5 Principaux éléments d'une monnaie 13
1.6 Variantes d'une 2 euros commémorative avec les étoiles de différentes tailles 13
1.7 Exemples de monnaies « belles épreuves » 14
1.8 Exemples de monnaies *US Washington Quarter Dollar* : (a) sans patine ; (b) avec patine 14
1.9 Exemples de monnaies gradées suivant l'échelle française 17
1.10 Dr. Willam H. Sheldon, inventeur de l'échelle de Sheldon 18
1.11 Monnaie *Flowing hair 1 Cent* au grade PO-1 (*basal state*) 19
1.12 Sociétés de gradation professionnelle : PCGS, NGC, ICG et ANACS 22
1.13 Monnaies gradées scellées dans des coques de protection 22
1.14 Illustrations des monnaies *Liberty Bell Half Dollar* : (a) issue d'un scanner ; (b) photo amateur ; (c) photo de qualité réalisée par la société NGC 24
1.15 Installations destinées à la photographie numismatique : (a) installation amateur ; (b) tente de lumière et support illuminé ; (c) lumière axiale . . . 24
1.16 Monnaies photographiées à l'aide d'une lumière axiale (encadrées en rouge) et non-axiale : (a) monnaie courante ; (b) belle épreuve 25
1.17 Photos numismatiques professionnelles : (a) PCGS TrueView ; (b) NGC Photo Vision (monnaie sous coque) 26
1.18 Logo de la société GENI 26
1.19 Site web de la société GENI (www.geni.expert) 27
1.20 Coques de protection du modèle « américain » 28
1.21 Conception de la coque de protection par la société GENI 29

2.1 Différence de formes entre des monnaies anciennes (en haut) et des monnaies « modernes » (en bas) 34

2.2 Motifs non-standardisés dans les monnaies anciennes appartenant au même type monétaire (Zambanini et Kampel 2012) 35
2.3 Schéma général de la reconnaissance d'objets (en bleu), les boîtes rouges représentent le schéma de gradation des monnaies utilisé pour cette thèse. . 35
2.4 Exemples de techniques de segmentation d'image : (a) par seuillage : seuil adaptatif (Yanowitz et Bruckstein 1989) ; (b) par régions : *meanshift* (Comaniciu et Meer 2002) ; (c) par contours : contours actifs *level-set* ; (d) par graphe : de plus court chemin (Falcão, Stolfi et Alencar Lotufo 2004) (e) par approche évoluée : modèle paramétrique + carte de dissimilarité de couleur + contours actifs guidés (Cerutti et al. 2011) 38
2.5 Exemples de monnaies bimétalliques segmentées à l'aide de la méthode d'Otsu : (a) monnaie globale avec une intensité différente de l'arrière-plan ; (b) monnaie avec une partie dont l'intensité est proche de l'arrière-plan . . 40
2.6 Exemples de monnaies segmentées par transformée de Hough : (a) monnaies « modernes » (Reisert, Ronneberger et Burkhardt 2007) ; (b) monnaies anciennes (Arandjelovic 2010) 41
2.7 Facteur de Forme (FF) des régions segmentées obtenues à l'aide de différents seuils T (Zambanini et Kampel 2009) 42
2.8 Exemples de monnaies anciennes segmentées par la méthode de Zambanini et Kampel (2009) : (a) monnaie bien segmentée ; (b) monnaie mal segmentée à cause des ombres 42
2.9 Monnaies physiquement similaires : (a) Thaï 10 Baht (environ 20 centimes en euro) ; (b) 2 euros 43
2.10 Répartition de la carte des contours (Van Der Maaten et Poon 2006) : (a) EDD ; (b) ADD et (c) EADD 46
2.11 Processus de construction du descripteur fondé sur les directions des gradients (cartes des directions des gradients et représentations binaires des descripteurs correspondant uniquement aux boîtes englobantes en haut à gauche) (Reisert, Ronneberger et Burkhardt 2007) 47
2.12 Exemples d'eigenimages construites à partir des contours lissés (Huber et al. 2005) 48
2.13 Description de forme d'une monnaie ancienne à l'aide d'une courbe normalisée (Huber-Mörk et al. 2011) 49
2.14 Points d'intérêt SIFT appariés (Kampel et Zaharieva 2008) : (a) sur deux monnaies appartenant au même type monétaire ; (b) sur deux monnaies appartenant à des types monétaires différents 50
2.15 Schéma général du fonctionnement d'un sac de mots visuels (Anwar, Zambanini et Kampel 2013) 51

2.16 (a) Légendes comportant le même texte (*E PLURIBUS UNUM*) avec des apparences variées sur de différents types monétaires « modernes » ; (b) légendes avec des textes différents qui différencient les types monétaires similaires. 53
2.17 Capture d'écran de *Photograde*^{TM} en ligne : images de référence représentant un *Lincoln Cent* avec le grade MS-65, MS-66 et MS-67. 56
2.18 Avers de l'Essai 10 centimes Merley à 24 pans (bord zoomé encadré en rouge) 61
2.19 Avers de l'Essai 1 livre à 12 pans (bord zoomé encadré en rouge) 61

3.1 Exemples de 65 types monétaires extraits de la base de données constituées par Arandjelovic (2010) 66
3.2 Exemples d'images de la base *MUSCLE CIS* 2006 67
3.3 Exemples d'images de monnaies de la République romaine appartenant à deux types monétaires (un type sur chaque ligne) 68
3.4 Exemples de monnaies de la base *USA_Grading* de grades différents appartenant à deux types monétaires (un type sur chaque ligne) 70
3.5 Exemples de monnaies trop usées « sans motif » dans *Photograde*^{TM}, exclues de la base de données *USA_Grading* 70
3.6 Éclairage non-uniforme sur des monnaies de grades proches du même type monétaire 71
3.7 Exemples de monnaies de la base de données *GENI_CoinPhotos* 72
3.8 Exemples de monnaies photographiées au laboratoire LIRIS 73
3.9 Exemples de monnaies semi-synthétiques de formes variées à partir d'une vraie photo 73
3.10 Exemples de monnaies segmentées et recalées de *20FrancsCOQ_Photos* . . 74
3.11 Annotation des « éléments inattendus » sous des calques superposés à l'aide de Photoshop 75
3.12 Capture d'écran de l'utilisation de l'outil interactif *CoinLegendExtract* . . . 79
3.13 Extraction de légendes horizontales à l'aide de *CoinLegendExtract* : (a) image originale, (b) image corrigée par rotation, (c) image correspondante en coordonnées polaires. 79
3.14 Zones candidates susceptibles de contenir la date dans la sous-base considérée de *USA_Grading* 80
3.15 Pipeline de l'extraction automatique des chiffres de la monnaie 81
3.16 Exemples de dates extraites : les bonnes ou quasi-bonnes encadrées en vert (a-f) ; les mauvaises encadrées en rouge (g-h) 82
3.17 Exemples d'images de la base de données *CoinNUMS* : (a) *CoinNUMS_geni* ; (b) *CoinNUMS_pcgs_m* ; (c) *CoinNUMS_pcgs_a* 83

3.18 Exemples d'imagettes de chiffres : (a) images synthétiques « idéales » des chiffres en police de « type creusé » et exemples réels : (b) chiffres avec un premier plan distinct ; (c) chiffres inclinés ; (d) chiffres légèrement découpés ; (e) chiffres avec du bruit externe ; (f) chiffre avec du bruit interne 85

3.19 Capture d'écran du site où *CoinNUMS* est disponible 86

4.1 Modèles de formes différentes : (a) forme polygonale ; (b) forme sinusoïdale ; (c) forme ronde 92

4.2 Processus de la méthode de segmentation composée de trois étapes principales : segmentation « grossière », estimation du modèle et calcul du modèle optimal 93

4.3 Qualité de la segmentation « grossière » : (a) monnaie originale (synthétique) de forme sinusoïdale ; (b) segmentation exacte ; (c) segmentation approximative ; (d) segmentation trop bruitée 93

4.4 Différentes étapes de la segmentation « grossière » : (a) images originales ; (b) images seuillées I^T ; (c) cartes des contours I^C ; (d) cartes des probabilités I^P ; (e) résultats de la segmentation « grossière » I^S 95

4.5 Bord bruité sur I^S zoomée 95

4.6 Recherche du segment le plus long et du segment plus court sur une région polygonale et une région sinusoïdale 97

4.7 Estimation de N_e à l'aide du rapport $r(N)$ des segments simulés 99

4.8 Estimation de la rotation de la monnaie 100

4.9 Estimation de la forme de la monnaie : (a) Ψ_1 ; (b) Ψ_2 ; (c) $|Area(\Psi_1) - Area(\Psi_2)|$ 101

4.10 Graphique des expériences sur r_S 102

4.11 Énergie du modèle avec différentes configurations : (a) $E_{ext} > 0, E_{int} > 0, E > 0$;(b)$E_{ext} = 0, E_{int} > 0, E > 0$; (c) $E_{ext} > 0, E_{int} = 0, E > 0$; (d)$E_{ext} = 0, E_{int} = 0, E = 0$ 102

4.12 Courbe $N - E_{min}$ 103

4.13 Exemples de monnaies trouées françaises : (a) monnaie avec un trou rond (Essai Michelin Théodore 5 centimes) ; (b) monnaie avec un trou carré (Essai Michelin Théodore 4 centimes) 105

4.14 Processus de segmentation des monnaies trouées 106

4.15 Exemples de monnaies segmentées (repérées par un contour rouge) de la base *LIRIS_coinPhotos* à l'aide du modèle optimal : (a) monnaie à 16 pans, (b) à 22 pans, et (c) à 24 pans 109

4.16 Exemples de monnaies segmentées de la base *GENI_CoinPhotos* (première ligne : images originales ; deuxième ligne : segmentations « grossières » zoomées ; troisième ligne : segmentations précises repérées par un contour vert) : (a) monnaie ronde, (b) monnaie polygonale, (c) monnaie sinusoïdale, (d) monnaie trouée 110
4.17 Exemples de monnaies particulières dans la base *GENI_CoinPhotos* (en haut : images originales ; en bas : segmentations zoomées repérées par un contour rouge) : (a) monnaie striée, (b) monnaie cannelée (c) monnaie quasi-ronde 111
4.18 Monnaie segmentée d'un arrière-plan noir avec des bords zoomés encadrés en rouge 112
4.19 Segmentation d'une monnaie polygonale ayant des pans courbés (repérée par un contour rouge) 113
4.20 Segmentation d'une monnaie polygonale ayant des coins arrondis (repérée par un contour rouge) 113

5.1 Schéma global du calcul des scores de similarité 118
5.2 Pipeline de l'identification du type monétaire 119
5.3 Points d'intérêts détectés par différents algorithmes sur la même image . . 120
5.4 Détection et appariement des points d'intérêt ORB sur deux monnaies *US Barber Quarter obv.* (points rouges : points d'intérêt détectés non appariés ; points jaunes : points d'intérêt appariés) : (a) une paire des points d'intérêt appariés sur les nez du portrait ; (b) appariement naïf de tous les points d'intérêt en utilisant un seuil de distance. 121
5.5 Comparaison entre le nombre de « vraies » correspondances selon différents paramètres 122
5.6 « Vraies » correspondances obtenues entre deux monnaies en utilisant les mêmes paramètres (a) sur les images originales et (b) sur les cartes des gradients associées 123
5.7 Contrainte de voisinage 124
5.8 Contrainte de bordure 124
5.9 Score de similarité non discriminant dans certains cas : (a) deux monnaies appartenant à des types monétaires différents ; (b) deux monnaies appartenant au même type monétaire 125
5.10 Points d'intérêts ORB retenus 126
5.11 Construction de deux graphes quasi-identiques à l'aide des triangulations de Delaunay 127
5.12 Construction du descripteur d'arête avec $s = 3$ 128

5.13 Comparaison des arêtes correspondantes : bandes centrales (en rouge) prises en compte 129
5.14 Arêtes similaires marquées en orange (arêtes courtes non considérées) . . . 130
5.15 Résultats de classification des k-plus proches voisins sur la base *USA_Grading* 132
5.16 Influence de la rotation des monnaies sur le taux d'identification en utilisant SSE 133
5.17 Exemples de monnaies issues des classes ambiguës dont les nuances sont marquées en rouge : *Barber10.rev* et *Seated10.rev* 134

6.1 Périodes d'émission différentes de types monétaires similaires : *DrapDolSE.rev* (encadrés en bleu) : 1795-1798 ; *DrapDolLE.rev* (encadré en rouge) : 1798-1803 139
6.2 Pipeline de l'identification du type monétaire prenant en compte la détection du millésime 140
6.3 Exemples de chiffres binaires et leur enveloppe convexe (rouge) (a) et (c) ; avec leur carte de bulle associée (b) et (d) 141
6.4 Construction de la carte de bulles 142
6.5 Pipeline de l'utilisation du système de votes 144
6.6 Graphes de Reeb générés de dix chiffres bien segmentés en police courante 146
6.7 Exemples de graphes de Reeb générés de chiffres originalement en police de « type creusé » : (a) chiffres non bruités (b) chiffre bruité 146
6.8 Exemples de Topologies Simplifiées (TS) (encadrées en rouge) des chiffres bruités 147
6.9 Exemples de polices courantes de chiffres binaires et leur enveloppe convexe (en rouge) : (a) chiffres « 1 » avec des « topologies » différentes ; (b) chiffres appartenant aux différentes classes avec des « topologies » similaires 148
6.10 Pipeline de la comparaison des graphes de Reeb 148
6.11 Exemples de la base de données *Chars74K* : *Chars74K_EnglishImg* (en haut) ; *Chars74K_EnglishHnd* (en bas) 150
6.12 Exemples de chiffres synthétiques testés (en haut) et de cartes des gradients binarisées obtenues sur la base *CoinNUMS_pcgs_a* (en bas) 151
6.13 (a) Chiffres extraits de monnaies usées (b) cartes des gradients binarisées « illisibles » 151
6.14 Résultats obtenus sur les différents tests (ANN1 : ANN avec 100 images d'apprentissage ; ANN2 : ANN avec 200 images d'apprentissage) 152
6.15 Matrices de confusion 153
6.16 Cas difficiles avec leurs cartes des gradients binarisées : (a) chiffre « 9 » avec son trou ouvert ; (b) chiffre « 9 » avec son ouverture « bouchée » ; (c) chiffre « 6 » avec un bout de motif touché 155

7.1 Exemples de monnaies *US Indian 2.5 Dollars* (revers) gradées par la société PCGS : (a) monnaies usées de grades faibles proches ; (b) monnaies bien conservées de grades élevés proches 160
7.2 Pipeline total de l'analyse d'images pour la gradation des monnaies 161
7.3 Exemples de différents types « d'éléments inattendus » : (a) coups ; (b) rayures ; (c) taches 162
7.4 Pipeline de la gradation des monnaies basée sur la quantification des « éléments inattendus » 162
7.5 Exemples « d'élément inattendu » significatif (zoomé et encadré en rouge) et « d'élément inattendu » ténu qui rend la surface « granuleuse » (zoomé et encadré en bleu) 163
7.6 Prétraitement d'une monnaie à grader : (a) monnaie en couleur recalée I_k^* ; (b) carte des gradients I_k^G ; (c) monnaie seuillée I_k^T 165
7.7 Images moyennes des monnaies : (a) monnaie moyenne en couleur $\hat{I}$; (b) moyenne des cartes des gradients $\widehat{I^G}$; (c) moyenne des monnaies seuillées $\widehat{I^T}$ 166
7.8 (a) Masque binaire de la monnaie de référence M ; (b) zone d'intérêt ZI_k d'une monnaie réelle 167
7.9 Détection des « éléments inattendus » significatifs 168
7.10 Zones d'intérêt des monnaies avec des imagettes extraites zoomées : (a) imagette granuleuse ; (b) imagette lisse 169
7.11 Division de la zone d'intérêt en imagettes de différentes tailles fixées : (a) 32×32 pixels ; (b) 48×48 pixels ; (c) 64×64 pixels 170
7.12 Exemples d'imagettes annotées de différentes tailles (redimensionnées à la même taille pour la visualisation ; imagettes granuleuses encadrées en bleu ; imagettes lisses encadrées en rouge) : (a) 32×32 pixels ; (b) 48×48 pixels 171
7.13 Exemples d'imagettes annotées prétraitées (imagettes granuleuses encadrées en bleu ; imagettes lisses encadrées en rouge) : (a) différence d'uniformité évidente ; (b) différence d'uniformité pas évidente 171
7.14 Structure générale d'un réseau de neurones convolutionnels 172
7.15 (a) Étiquetage des composantes connexes de la zone d'intérêt et (b) classification automatique en deux zones d'importance 174
7.16 Exemple d'une monnaie avec détection des « éléments inattendus » significatifs détectés (en rouge) : (a) l'annotation de l'expert ; (b) le résultat obtenu 176
7.17 Comparaison des « éléments inattendus » significatifs détectés avec l'annotation de l'expert (en haut : annotation ; en bas : résultats obtenus) 176
7.18 Exemples d'imagettes testées : (a) en couleur (sans prétraitement) ; (b) en niveaux de gris ; (b) en niveaux de gris normalisés 177

7.19 Zones granuleuses (marquées en gris) sur trois exemples de monnaies : (a) annotation manuelle faite par la société GENI ; (b) (c) (d) zones granuleuses détectées par le réseau entraîné en utilisant respectivement les imagettes en couleur (b), en niveaux de gris (c) et en niveaux de gris normalisés (d) 179
7.20 Exemples d'imagettes de test (imagettes annotées granuleuses encadrées en bleu ; imagettes annotées lisses encadrées en rouge) : (a) imagettes bien classifiées ; (b) imagettes mal classifiées 181
7.21 Exemples d'imagettes « ambiguës » 181

Liste des tableaux

1.1 Catalogue des prix (€) pour 10 Francs Cérès 15
1.2 Système de gradation français à 7 échelles 17
1.3 Termes équivalents dans les systèmes de gradation européens 17
1.4 Echelle de Sheldon de gradation 19
1.5 Echelle de Sheldon de rareté 20
1.6 Critères de gradation destinés aux monnaies spécifiques 21

3.1 Comparaison des bases de données numismatiques 76
3.2 Propriétés visuelles des différentes sous-bases de CoinNUMS 85
3.3 Distribution de données dans les sous-bases de *CoinNUMS* 86

4.1 Performance concernant l'estimation de forme des monnaies sur des images de taille 700 × 700 pixels environ de la base *LIRIS_CoinPhotos* 108
4.2 Temps de calcul sur *LIRIS_CoinPhotos* (AMD Turion II 2,4GHz, 2G RAM, Win7 32bits) 108

5.1 Récapitulatif des divers tests faits sur la base *USA_Grading2* 131
5.2 Résultats sur différentes sous-bases de *USA_Grading* 133
5.3 Extraits de résultats sur des « classes ambiguës » 134

6.1 Vote d'une bulle en fonction de son étiquette 145
6.2 Distribution de données dans différentes bases de test 151

7.1 Résultats de classification sur la base de 2865 imagettes annotées 177
7.2 Taux de classification (%) obtenus avec *AlexNet* en faisant du *Fine Tuning* en utilisant un nombre croissant d'imagettes annotées 180
7.3 Taux de classification (%) obtenus avec *AlexNet* en faisant du *Fine Tuning* en utilisant différentes distributions d'imagettes annotées 180