

Rapport final (31/03/2009) du projet IBISA (PEPS 2007)

(*Image-Based Identification/Search for Archeology*)

<http://dept-info.labri.fr/~sm/Projets/IBISA>



1 Description du projet

IBISA (*Image-Based Identification/Search for Archeology*) est un système logiciel (sous licence libre GNU *General Public License*) pensé par une équipe pluridisciplinaire de chercheurs en informatique et en archéologie, appartenant à 3 UMR du CNRS réparties sur 2 départements (ex-ST2I et ex-SHS) et 2 universités de Bordeaux (Bordeaux 1 et 3 – Michel de Montaigne).

1.1 Porteur

Sylvain Marchand est Maître de Conférences au LaBRI (CNRS UMR 5800, ex-département ST2I), Université Bordeaux 1, depuis septembre 2001, et habilité à diriger des recherches depuis décembre 2008. C'est un jeune chercheur en informatique, membre de l'équipe Image et Son, et qui s'intéresse aux représentations spectrales principalement pour les sons, mais aussi pour les images. Il est également membre de la Société archéologique de Bordeaux depuis 1999, et vice-président de sa section numismatique (Cercle Bertrand-Andrieu) depuis janvier 2004. Il y étudie les monnaies antiques (grecques et romaines).

En 2006, lors d'un exposé [Mar06] au Musée d'Aquitaine (Bordeaux) pour parler de la numismatique, science qui étudie l'Histoire à travers les monnaies, Sylvain Marchand a évoqué les compétences du LaBRI en analyse et traitement d'images, et a montré ce qu'elles pourraient apporter à la numismatique. Il a été question notamment d'extraction de contour, de recalage automatique, d'identification automatique, et de construction d'un représentant idéal (voir plus loin).

Suite à cet exposé, il a été contacté par Pierre Regaldo - Saint Blancard, membre du Service régional de l'archéologie (DRAC Aquitaine), qui l'a alors mis en relation avec deux UMR CNRS (ex-département SHS) de l'Université Michel de Montaigne – Bordeaux 3, qui ont des problématiques archéologiques *a priori* différentes, mais mettant en jeu quasiment les mêmes méthodes informatiques, étudiées au LaBRI. . .



FIG. 1 – Carreaux médiévaux provenant du château de Villandraut (XIV^e siècle).

1.2 Partenaires

• IRAMAT-CRP2A (CNRS UMR 5060, ex-département SHS, sections 32, 31, 15 et 05)

Françoise Bechtel (Professeure), **Ayed Ben Amara** (Maître de Conférences) et **Béatrice Cicutini** (doctorante) y étudient des carreaux estampés glaçurés [Eam81, Org04, AGPB⁺05, Cic06], qui décoraient les sols des monuments médiévaux prestigieux. Pour réaliser ces carreaux, on appliquait dans de l'argile support rouge une matrice portant un motif en relief. Celui-ci était donc imprimé en creux dans l'argile support. Il était ensuite rempli d'argile blanche, recouvert d'un mélange glaçurant, puis cuit. Beaucoup de carreaux subsistent (voir Figure 1), mais très peu de matrices.

• Ausonius (CNRS UMR 5607, ex-département SHS, section 32)

Alain Bresson (Professeur) et **Koray Konuk** (Chargé de Recherches CNRS) y étudient les monnaies grecques antiques [Bre05, BV08, Kon98, AKKM02]. **Jean-Pierre Bost** (Professeur émérite) s'intéresse quant à lui aux monnaies romaines [BM02]. Toutes ces monnaies furent jadis frappées à partir de coins, matrices portant des motifs en creux et laissant des empreintes avers et revers sur un flan métallique. De très nombreuses monnaies ont été retrouvées (voir Figure 2), mais rares sont les coins monétaires (matrices) qui sont parvenus jusqu'à nous.

Dans les deux cas, les archéologues constituent des bases de données d'images numériques des objets archéologiques, à des fins d'archivage ou de publication. Parallèlement, ils veulent pouvoir déterminer si deux objets – qui ont subi les outrages du temps (usure, casse, *etc.*) – proviennent de la même matrice, ou bien sont de même motif / style. À l'heure actuelle, la méthode utilisée est la comparaison visuelle des objets, reposant souvent sur l'interprétation subjective des motifs observés. Dans le cas de l'étude de plusieurs milliers d'objets, cela représente une tâche fastidieuse, engendrant erreurs (subjectivité et fatigue humaines aidant) et perte de temps...

Savoir si deux objets proviennent de la même matrice – qui subissait également une usure – donne une information chronologique importante. L'utilisation de motifs de styles semblables est aussi une indication importante. Dans le cas des monnaies, le lien entre les matrices avers / revers donne aussi des informations précieuses. Dans le cas d'une étude d'un ensemble de nombreux objets trouvés dans le même contexte (pavement dans le cas des



FIG. 2 – Monnaie romaine : denier de l’empereur Alexandre Sévère (avers et revers).

carreaux, trésor dans le cas des monnaies), on peut en apprendre beaucoup sur la manière dont ces objets ont été fabriqués (nombre total de matrices, ateliers de production) et échangés (contexte socio-économique), chronologiquement et géographiquement. Les liens entre objets et ensembles d’objets permettent de confirmer ou d’infirmer des hypothèses scientifiques, et ouvrent ainsi la voie à d’autres types de recherches (*e.g.* l’analyse des matériaux des objets).

- **LaBRI (CNRS UMR 5800, ex-département ST2I, section 07)**

Sylvain Marchand (Maître de Conférences HDR, porteur du projet) a une expertise des représentations spectrales pour les signaux, les sons et les images. Il a notamment étudié la transformée de Fourier-Mellin pour le recalage d’images [WZ00], l’a déjà appliquée à l’archéologie en développant un premier module logiciel de recalage d’images et de recherche basée image [Mar04]. Il avait également fait une étude préliminaire montrant que le facteur d’inter-corrélation reflétait bien la notion subjective de similarité entre images.

Pascal Desbarats (Maître de Conférences) est un spécialiste de l’analyse et du traitement d’images. Il a une expérience importante en recalage d’images, appliqué à l’imagerie médicale [dSDS⁺04, dSQDM05]. Il a également l’expérience du développement de logiciels conséquents.

Jacques-Olivier Lachaud (Professeur) a travaillé sur les modèles déformables [TL02, LT05], modèles qui s’avèrent être particulièrement bien adaptés au détourage des images de monnaies et devraient l’être également pour l’extraction des motifs des carreaux. **Anne Vialard** (Maître de Conférences) connaît bien ces modèles. Dans ce projet, ils ont fourni leur implémentation existante des modèles déformables et leur expertise du domaine pour aider à l’intégration de cette implémentation au sein du système logiciel.

1.3 Objectifs

Le projet IBISA a abouti à une maquette logicielle (sous licence libre GNU GPL) qui manipule une base d’images numériques (au maximum quelques milliers) d’objets archéologiques (carreaux estampés glaçurés médiévaux et monnaies antiques grecques et romaines). Ces objets ont comme particularités communes d’avoir été produits dans le passé par l’empreinte d’une matrice et d’avoir ensuite subi une usure au cours du temps.

Le système IBISA est entre autres en mesure de fournir une aide permettant de déterminer semi-automatiquement, d’après leurs images, si deux objets sont les mêmes, ou proviennent de la même matrice, ou présentent le même motif / style, ou bien sont vraiment différents.

Il est important de noter que le système est résistant aux changements de conditions de prise de vue, et peut identifier un même objet à partir d'images prises différemment. La correction des conditions de prise de vue (centrage, orientation, échelle) se fait également de manière automatique. Notamment, si une image *S* semblable à *C* a été trouvée dans la base, cela se fait en recalant *C* sur *S* (*via* le recalage d'images). Le détourage de l'image pour isoler la monnaie du fond ou le motif du carreau peut aussi être effectué semi-automatiquement (*via* les modèles déformables).

Pour chaque classe d'images provenant de la même matrice, un *représentant idéal* sera construit automatiquement. Ce représentant serait l'objet le plus complet et dans l'état de conservation optimal, déduit par le système en prenant le meilleur dans chaque objet présent dans la classe. Dans le cas des carreaux, il s'agit de construire le motif complet, obtenu à partir de fragments. Dans le cas des monnaies, il s'agit de reconstruire l'image d'une monnaie sans usure ou presque. Cela revient à retrouver le motif de la matrice originelle (disparue) à partir des objets étudiés (parvenus jusqu'à nous).

2 Actions entreprises et à venir

Le but premier de ce PEPS IBISA était de permettre la réalisation de la maquette du système logiciel. Son développement a pris plus de temps que prévu, mais nous disposons actuellement d'un prototype utilisable qui remplit la plupart des objectifs annoncés (voir Section 3).

2.1 Réunions de travail

Des réunions travail se sont tenues régulièrement, avec des réunions plus conséquentes tous les 6 mois environ (11 juin 2007, 14–15 janvier 2008, 15 octobre 2008) au CRP2A, à Ausonius ou au LaBRI, la dernière en date ayant été la plus importante, car ce fut l'occasion pour tous de discuter de la maquette logicielle avec le prestataire extérieur (Anthony Beurivé).

De plus, des réunions entre le porteur de projet et ce même prestataire se sont tenues à la même fréquence – tous les 6 mois environ (21 décembre 2007, 3 juillet 2008, 20 mars 2009) à Paris ou à Bordeaux.

2.2 Groupe de travail

Le porteur de projet a présenté le projet IBISA et ses aspects d'analyse d'image lors du groupe de travail de l'équipe Image et Son du LaBRI, le 10 janvier 2008.

2.3 Prise de contacts

Il a également pris contact avec Michel Amandry, Directeur du Département des monnaies, médailles et antiques de la Bibliothèque nationale de France, Paris. Ce dernier s'est montré enthousiaste quant à l'utilité du futur outil IBISA. Le projet lui a été présenté une première fois le 14 décembre 2007 à Paris. Une démonstration du logiciel sur des monnaies pourra bientôt lui être faite.

2.4 Constitution de bases d'images

Nous avons commencé à constituer une base d'images de référence. Grâce au CRP2A, nous avons déjà une importante base d'images de **carreaux médiévaux**, puisque nous disposons des images utilisées par Béatrice Cicuttini lors de sa thèse (notamment celles des carreaux du pavement du château de Villandraut, Aquitaine, XIV^e siècle). Ausonius nous a fourni en avant-première des images de planches de **monnaies grecques** (du trésor de Phaïstos, Crète, II^e siècle av. J.-C.) et nous devrions pouvoir faire la numérisation de **monnaies romaines** (et probablement nous associer à une publication sur l'étude des monnaies numérisées). Cependant, l'obtention de bases de données d'images d'objets archéologiques semble plus complexe que prévu initialement. Les musées sont souvent réticents à diffuser le contenu de leurs collections (ou du moins ils souhaitent attendre la publication de ces collections par leur propres moyens, ce qui prend beaucoup de temps). Des bases d'images existent, notamment celles des objets provenant de ventes aux enchères publiques passées, mais alors nous n'avons pas les droits sur ces images. . .

2.5 Conférence internationale VAST 2009 (septembre 2009)

La conférence Eurographics *International Symposium on Virtual Reality, Archeology and Cultural Heritage*, regroupant image et archéologie, semble toute indiquée pour la publication d'une présentation du projet IBISA (voir par exemple [ZKS08]). Nous soumettrons donc courant avril un article présentant l'outil informatique IBISA, les techniques d'analyse et de traitement d'images mises en œuvre, ainsi que les premiers résultats sur nos bases de données.

2.6 Acquisition d'images pour le calibrage du logiciel (avant fin 2009)

Pour les monnaies romaines, le Musée d'Éauze (Gers) pourrait nous donner accès à environ 1500 monnaies (à numériser) issues du même coin (même matrice), information précieuse pour calibrer le module de similarité de notre système logiciel. Jean-Pierre Bost a pris des contacts préliminaires dans ce sens. Nous avons d'ors et déjà le matériel pour assurer la numérisation des monnaies, qui sera réalisée en principe avant la fin de l'année 2009.

3 Outil informatique IBISA

Le logiciel IBISA est un outil informatique pour gérer des collections de photos numériques d'objets archéologiques. Il a été développé dans un contexte scientifique afin d'expérimenter différents algorithmes d'analyse d'images pour l'archéologie. L'objectif est de trouver de bons algorithmes pour, d'une part, isoler les objets individuels sur des photos qui en regroupent plusieurs, et d'autre part, déterminer les similarités entre les objets ainsi extraits. Toutes ces opérations se font au travers d'une interface graphique simple d'utilisation.

IBISA est développé sous licence GNU GPL, ce qui signifie entre autres choses que le code source est librement disponible, modifiable, échangeable. Sa taille modeste lui permet d'être maintenu par une seule personne ou un petit groupe de personnes. Il est écrit en langage C++ pour les traitements classiques sur les images, qui doivent être performants (rotations, changements d'échelle, etc.). L'interface graphique est écrite en Tcl/Tk, un environnement particulièrement simple à maîtriser pour ce type de tâche. Les informations sur les photos

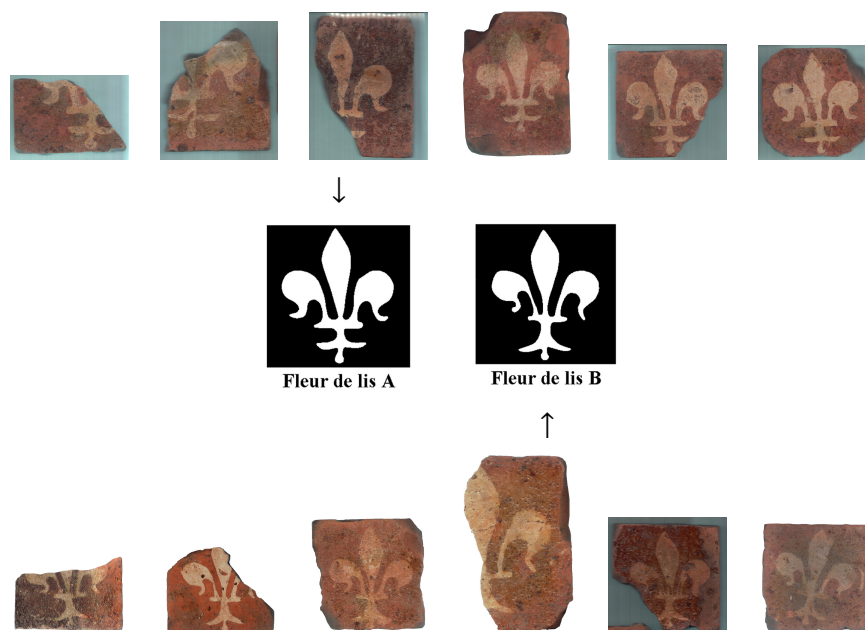


FIG. 3 – Étude de pavements : 2 types de motifs de lys. En haut des carreaux portant le lys de type A. En bas des carreaux portant le lys de type B. Au milieu, les représentants idéaux de ces 2 types de lys.

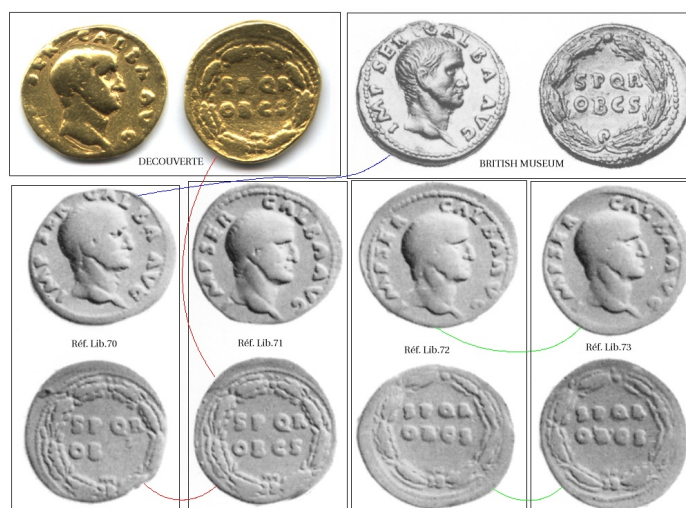


FIG. 4 – Étude de trésors : 6 monnaies en or (*aurei*) de l'empereur romain Galba. En haut à gauche se trouve l'exemplaire à étudier [PPS08]. En haut à droite se trouve un exemplaire du *British Museum* (en meilleur état de conservation) [Sea00]. En bas, une série de 4 exemplaires provenant d'un trésor découvert en Belgique [Thi72]. Il se trouve que : le revers de l'exemplaire à étudier provient du même coin que le numéro 71 du trésor ; le revers du numéro 71 provient du même coin que le numéro 70 (même si deux lettres manquent, car le coin était bouché lors de la frappe du 70) ; l'avvers du numéro 70, bien que plus usé, provient du même coin que l'exemplaire du *British Museum*.

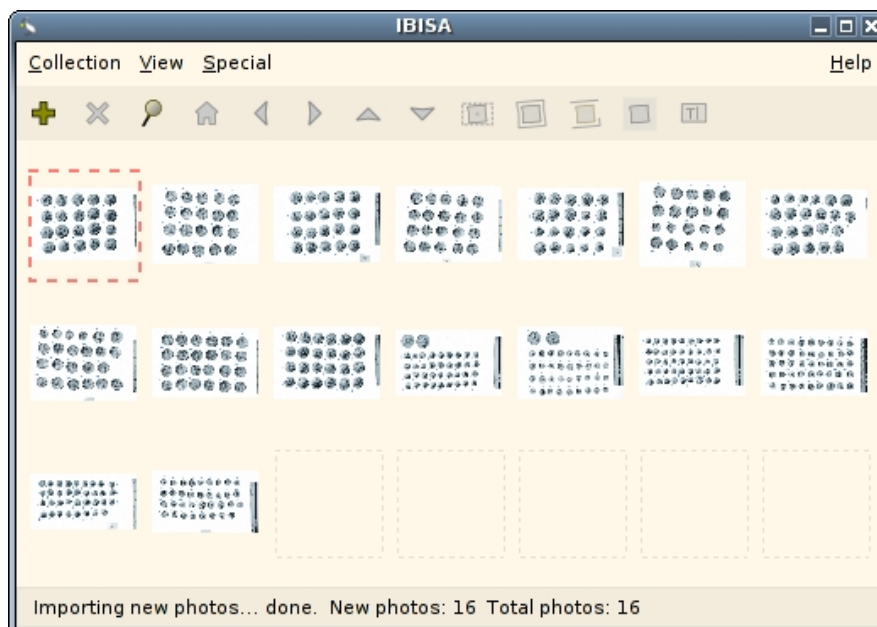


FIG. 5 – Logiciel IBISA (1/7) : Pour un utilisateur donné, les premières opérations au sein du logiciel consistent à importer des photos se trouvant sur son disque dur puis à extraire de ces photos les objets archéologiques individuels. Cette image montre l’interface du logiciel après que des photos de monnaies grecques y aient été importées.

sont enregistrées dans des bases de données SQLite portables. Enfin, dès qu’il s’agit d’expérimenter des algorithmes de traitement d’image sophistiqués, le logiciel fait appel à Octave, un logiciel très répandu pour le calcul scientifique. L’ensemble est organisé de façon simple et modulaire.

Le logiciel manipule une base d’images numériques (au maximum quelques milliers) d’objets archéologiques (carreaux / monnaies). Lorsqu’un nouvel objet est à étudier, son image est prise à partir d’un scanner à plat ou d’un appareil photo numérique. Il est alors possible :

- d’extraire l’objet du fond
(détourage : élimination du contexte, des ombres portées, etc.) ;
- de normaliser les conditions de prise de vue
(recalage : centrage [translation], alignement [rotation], mise à l’échelle [homothétie]) ;
- étudier les correspondances / liaisons de coins (similarité) ;

et il sera possible à brève échéance de déterminer automatiquement l’axe des coins. En revanche, la construction d’un représentant idéal, objectif annoncé mais clairement trop ambitieux, ne pourra se faire qu’à plus long terme.

3.1 Détourage

Le détourage permet d’isoler l’objet du fond de l’image. En effet, pour que le recalage d’images (Section 3.3) fonctionne correctement, il est indispensable de pouvoir faire abstraction du fond, en déterminant un masque m .

Les algorithmes généraux style “baguette magique” sont inadaptés (par exemple, pour



FIG. 6 – Logiciel IBISA (2/7) : L’une des photos une fois sélectionnée et agrandie, dans laquelle on a commencé à détourer un objet parmi d’autres (ligne rouge).



FIG. 7 – Logiciel IBISA (3/7) : L’objet une fois extrait.

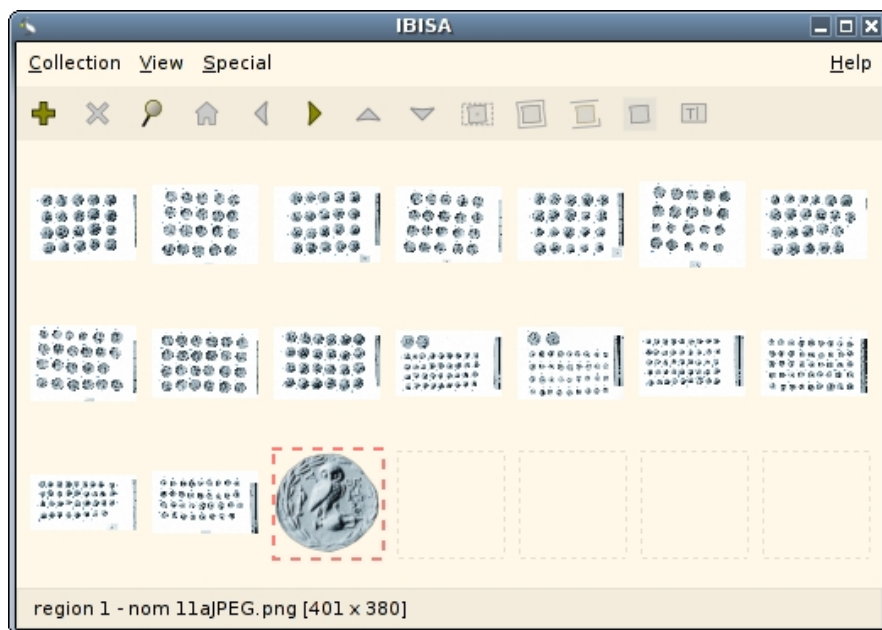


FIG. 8 – Logiciel IBISA (4/7) : L'ensemble initial de photos auquel l'objet détourné a été ajouté.

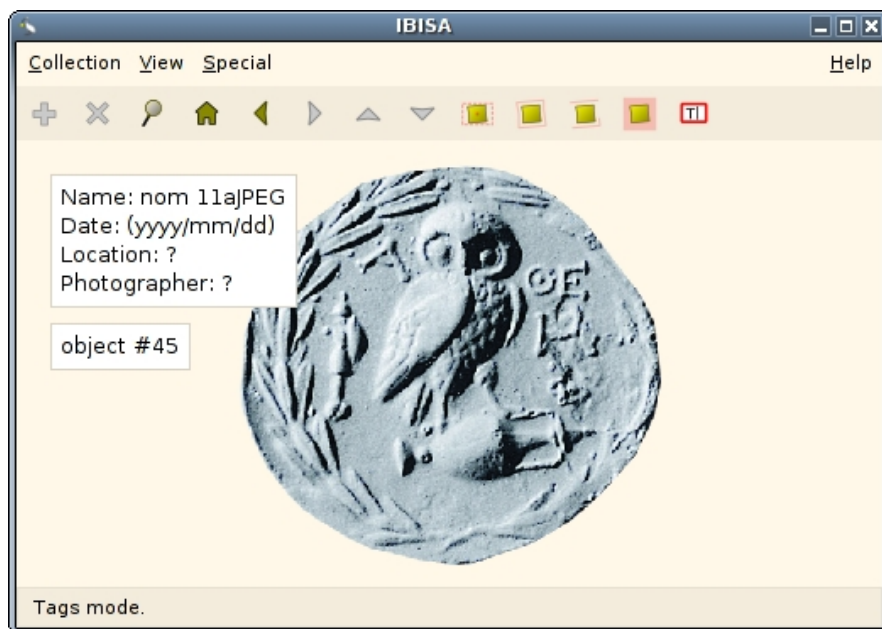


FIG. 9 – Logiciel IBISA (5/7) : Il est également possible de fournir des étiquettes (ou *tags*) à chaque photo ou objet.

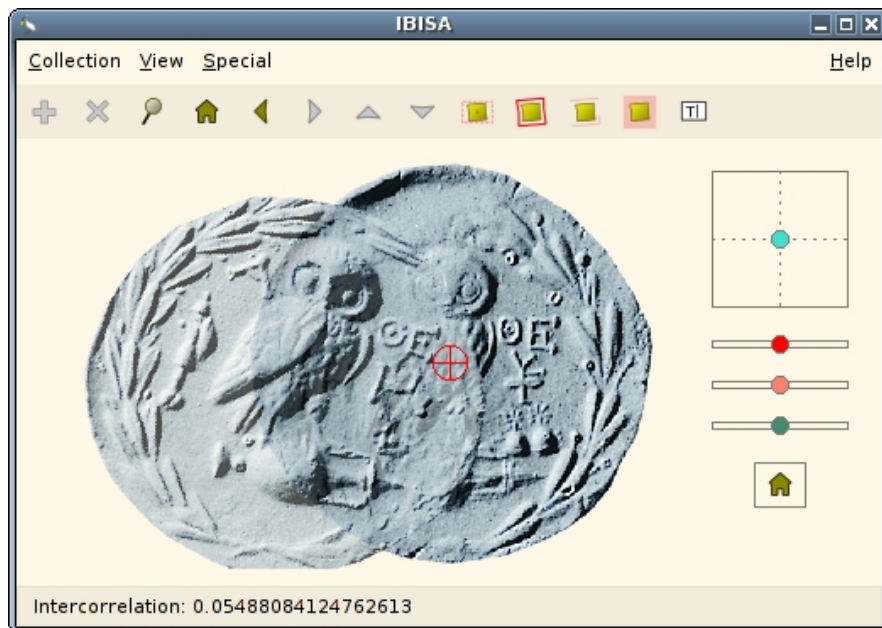


FIG. 10 – Logiciel IBISA (6/7) : Plus tard, on peut alors comparer les objets entre eux notamment en les ajustant manuellement grâce à un effet de transparence.



FIG. 11 – Logiciel IBISA (7/7) : Mais le plus sûr moyen d’obtenir de bons résultats et de passer par un algorithme de recalage automatique, disponible dans le menu de l’application. Pour cela, le logiciel Octave est appelé. Cette méthode permet à un spécialiste de modifier directement le code Octave de l’algorithme de recalage pendant que l’application IBISA fonctionne et, par conséquent, d’expérimenter ces changements de manière interactive.

une monnaie, du fait de l'ombre portée, soit le masque englobe la partie ombragée, soit il omet une partie de la monnaie). Ici, nous avons des connaissances *a priori* sur l'objet (par exemple, pour une monnaie, on cherche une forme presque ronde, de courbure faible).

Les modèles déformables par contours actifs (également connus sous le nom anglais de *snakes* [KWT88]) constituent une méthode semi-automatique pour le détourage. Pour une image à détourer donnée, il s'agit d'initialiser un contour assez proche du contour à extraire, puis de procéder à un certain nombre d'itérations de déformation du contour actif, afin de se rapprocher du contour idéal. On associe pour cela une énergie au contour, et à chaque itération le contour se déforme de façon à atteindre une position d'énergie minimale.

L'énergie totale est la somme des énergies internes au contour (énergie de tension et énergie de torsion) et des énergies externes (énergie provenant de l'image, de l'interaction avec l'utilisateur, voire de la ressemblance avec une forme prédéfinie. . .).

L'énergie interne au contour C délimité par les points $p(s)$ (avec $s \in [a; b]$) est donnée par :

$$E_{\text{interne}}(C) = \int_a^b \left(\alpha \left| \frac{\partial p}{\partial s} \right|^2 + \beta \left| \frac{\partial^2 p}{\partial s^2} \right|^2 \right) ds$$

où $\frac{\partial p}{\partial s}$ représente les variations de longueur (l'étirement du contour actif), $\frac{\partial^2 p}{\partial s^2}$ représente la courbure (la torsion du contour actif), et α et β sont des constantes arbitraires. À noter qu'un contour actif qui suit une courbe régulière a une énergie interne faible.

L'énergie externe (provenant ici de l'image) est donnée par :

$$E_{\text{externe}}(C) = \int_a^b -|\nabla g(p(s))|^2 ds$$

où $\nabla g(p(s))$ désigne le gradient de l'image g au point $p(s)$. Ainsi, un contour actif qui est positionné sur un contour de l'image a une énergie externe faible.

Il reste à trouver le contour qui minimise l'énergie totale

$$E_{\text{totale}}(C) = E_{\text{interne}}(C) + E_{\text{externe}}(C).$$

Cela revient à résoudre (Euler-Lagrange) :

$$-\alpha \cdot \underbrace{p''}_{\text{dérivée seconde}} + \beta \cdot \underbrace{p^{(4)}}_{\text{dérivée quatrième}} = \frac{\partial |\nabla g(p)|^2}{\partial p}$$

équation dont la résolution peut se faire par différences finies, en discrétisant la courbe p en des points p_i ($i = 1, \dots, N$), de préférence avec un pas de discrétisation h constant. On approche alors les premières dérivées classiquement par différences finies :

$$\begin{aligned} p'_i &= \frac{p_i - p_{i-1}}{h} \\ p''_i &= \frac{p'_{i+1} - p'_i}{h} = \frac{p_{i+1} - 2p_i + p_{i-1}}{h^2} \\ p^{(3)}_i &= \frac{p''_i - p''_{i-1}}{h} = \frac{p_{i+1} - 3p_i + 3p_{i-1} - p_{i-2}}{h^3} \\ p^{(4)}_i &= \frac{p^{(3)}_i - p^{(3)}_{i-1}}{h} = \frac{p_{i+2} - 4p_{i+1} + 6p_i - 4p_{i-1} + p_{i-2}}{h^4} \end{aligned}$$

et on aboutit aux équations suivantes :

$$\beta p_{i+2} + (-\alpha h^2 - 4\beta) p_{i+1} + (2\alpha h^2 + 6\beta) p_i + (-\alpha h^2 - 4\beta) p_{i-1} + \beta p_{i-2} = h^4 f_i$$

où f_i désigne la dérivée de la norme au carré du gradient au point p_i . En pratique, on considère $h = 1$. Ces équations s'écrivent sous forme matricielle ainsi :

$$A \cdot P = \lambda F$$

où $P = (p_i)_{i=1,\dots,N}$, $F = (f_i)_{i=1,\dots,N}$, et A est la matrice circulante $N \times N$ suivante :

$$A = \begin{pmatrix} 2\alpha + 6\beta & -\alpha - 4\beta & \beta & 0 & \cdots & 0 & \beta & -\alpha - 4\beta \\ -\alpha - 4\beta & 2\alpha + 6\beta & -\alpha - 4\beta & \beta & 0 & \cdots & 0 & \beta \\ \beta & -\alpha - 4\beta & 2\alpha + 6\beta & -\alpha - 4\beta & \beta & 0 & \cdots & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & \cdots & 0 & \beta & -\alpha - 4\beta & 2\alpha + 6\beta & -\alpha - 4\beta & \beta \\ \beta & 0 & \cdots & 0 & \beta & -\alpha - 4\beta & 2\alpha + 6\beta & -\alpha - 4\beta \\ -\alpha - 4\beta & \beta & 0 & \cdots & 0 & \beta & -\alpha - 4\beta & 2\alpha + 6\beta \end{pmatrix}.$$

Comme F dépend en fait de P , on doit résoudre l'équation matricielle de manière itérative. Comme

$$\frac{P_{t+\Delta_t} - P_t}{\Delta_t} \approx -A \cdot P_{t+\Delta_t} + \lambda F(P_t)$$

on obtient

$$P_{t+1} \approx (\lambda F(P_t) + \gamma P_t) \cdot (A + \gamma I_N)^{-1}$$

où I_N désigne la matrice identité de taille $N \times N$, et où γ et λ sont des constantes arbitraires, la première désignant l'inverse du pas de temps Δ_t et représentant l'inertie de la courbe aux déplacements, la seconde représentant l'influence des forces auxquelles est soumise la courbe.

Dans l'implémentation actuelle, le gradient est estimé par la méthode de Sobel, et sa norme au carré est :

$$|\nabla g|^2 \approx (S_x * g)^2 + (S_y * g)^2$$

où $*$ désigne la convolution et S_x et S_y sont les masques de Sobel (normalisés) pour les directions x et y :

$$S_x = \frac{1}{4} \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad \text{et} \quad S_y = \frac{1}{4} \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix};$$

puis F , sa dérivée, est approchée par différences finies (selon les deux directions x et y).

Le module logiciel de détournage par modèles déformables a été développé et testé. En pratique, des instabilités ont été constatées. Il semblerait que le paramétrage (α , β , γ , λ et h) des modèles déformables soit un point plus critique que prévu. Des expérimentations plus poussées sont en cours au LaBRI.

3.2 Transformations

Une transformation d'image est une fonction qui, à un point p de l'image source g , fait correspondre un point p' de l'image destination g' ainsi :

$$\begin{aligned} T : \quad p &\mapsto p' = T(p), \\ g &\mapsto g', g'(p') = g(p). \end{aligned}$$

En pratique, on obtient efficacement g' à partir de g en inversant T : en parcourant le domaine de p' , on calcule $g'(p') = g(T^{-1}(p'))$; mais $T^{-1}(p')$ n'est généralement pas un point de l'échantillonnage originel de g , d'où la nécessité de **reconstruire** l'image (par exemple, en pratique, par interpolation (bi)linéaire).

Dans la suite, on s'intéresse uniquement aux **transformations rigides** (translation, rotation, homothétie).

3.2.1 Transformation élémentaire 1 : translation

En notation complexe et en représentation cartésienne (bien adaptée à l'addition), une translation de vecteur (Δ_x, Δ_y) s'écrit simplement :

$$p' = p + t \quad \text{où} \quad p' = x' + iy', p = x + iy \text{ et } t = \Delta_x + i\Delta_y.$$

3.2.2 Transformation élémentaire 2 : rotation+homothétie

En notation complexe et en représentation polaire (bien adaptée à la multiplication), une rotation d'angle ϕ centrée à l'origine¹ combinée à une homothétie de rapport s s'écrit simplement :

$$p' = (\underbrace{s}_{\text{homothétie}} \cdot \underbrace{e^{i\phi}}_{\text{rotation}}) \cdot p \quad \text{où} \quad p' = \rho' e^{i\theta'} \text{ et } p = \rho e^{i\theta}.$$

Notons qu'en coordonnées log-polaires, la rotation+homothétie d'angle ϕ et de rapport s devient une **translation** de vecteur $(\log(s), \phi)$, puisque :

$$\begin{aligned} \log(p) &= \log(\rho) + i\theta, \\ \log(p') &= (\log(s) + i\phi) + \log(p). \end{aligned}$$

Notons également que cette translation est circulaire (périodique) selon l'axe θ (figurant des angles modulo 2π).

3.2.3 Propriétés spectrales

Le passage du domaine spatial (image $g(p)$, p désignant un point de l'espace) au domaine spectral (spectre $G(\omega)$, ω désignant une fréquence – également bidimensionnelle) est assuré par la transformée de Fourier 2D :

$$F : g(p) \mapsto G(\omega) = \int_p g(p) e^{-i\omega \cdot p}$$

¹Pour une rotation non centrée à l'origine, on peut toujours se ramener au cas centré à l'origine par une translation (du centre de la rotation).

et le passage inverse est assuré par la transformée de Fourier 2D inverse :

$$F^{-1} : G(\omega) \mapsto g(p) = \int_{\omega} G(\omega) e^{+i\omega \cdot p}$$

($\omega \cdot p$ désignant le produit scalaire de ω et p).

Énonçons alors les deux propriétés sur les spectres des images sur lesquelles repose l'algorithme de recalage d'images utilisé dans le projet (Section 3.3) :

1. **La translation ne modifie pas le module du spectre d'une image** (mais y ajoute simplement un déphasage). Cela reste vrai en pratique dans le cas discret, mais la translation – comme le spectre – devient périodique.
2. **Le spectre d'une image obtenue par rotation est la rotation du spectre de l'image de référence.** (*Idem pour l'homothétie, mais avec un rapport inverse.*) Cela reste vrai en pratique dans le cas discret, sous réserve d'appliquer un fenêtrage avant le calcul du spectre (pour limiter le phénomène de repliement spectral).

3.2.4 Transformations rigides

Toute transformation rigide (composition arbitraire de translations, rotations et homothéties) peut être vue comme **une** translation suivie d'**une** rotation+homothétie. En effet, avec cette convention, la composition est une opération interne, car en faisant subir au point p en séquence les transformations rigides 1 et 2 :

$$p_1 = (p + t_1) \cdot r_1 \quad \text{puis} \quad p_2 = (p_1 + t_2) \cdot r_2$$

on vérifie que p_2 est obtenu de p par une telle transformation, puisque :

$$p_2 = (p + \underbrace{t_1 + t_2/r_1}_{t_{12}}) \cdot \underbrace{r_1 r_2}_{r_{12}}.$$

Mais l'inverse d'une telle transformation

$$p' = (p + t) \cdot r$$

est donnée par

$$p = p'/r - t = p' \cdot r' + t'$$

c'est-à-dire la rotation inverse suivie de la translation inverse.

Pour pouvoir composer et inverser de manière élégante les transformations, il serait possible de passer au système de coordonnées homogènes (l'ajout pour le point p d'une troisième dimension – avec une valeur 1 – au couple cartésien (x, y) permet d'écrire toutes les transformations – y compris la translation – sous forme de produits matriciels).

3.3 Recalage

Le recalage d'images permet de s'affranchir des transformations rigides (translation, rotation et changement d'échelle) inhérentes aux différentes conditions de prise de vue, et ce en vue d'obtenir une mesure de similarité (Section 3.4) fiable.

3.3.1 Trouver une translation

La méthode employée pour trouver la translation est la corrélation de phase. Soit g une image, et g' obtenue par translation de vecteur t :

$$g'(p') = g(p' - t)$$

en considérant le domaine spectral :

$$G'(\omega) = G(\omega) \cdot e^{-i\omega t}.$$

La corrélation de phase est définie par :

$$\frac{G^*(\omega) \cdot G'(\omega)}{|G^*(\omega)| \cdot |G'(\omega)|} = \frac{G^*(\omega) \cdot G(\omega) \cdot e^{-i\omega t}}{|G(\omega)| \cdot |G(\omega)|} = e^{-i\omega t}$$

(car $G \cdot G^* = |G|^2$, G^* désignant le conjugué de G), dont la transformée de Fourier inverse est une impulsion en t . Estimer t revient alors à une recherche de maximum.

3.3.2 Trouver une rotation+homothétie

Soit g une image, et g' obtenue par rotation d'angle ϕ et homothétie de rapport s :

$$g'(p') = g(p'/r).$$

En coordonnées log-polaires, cela se ramène à une recherche de translation (voir précédemment), le couple (s, ϕ) étant donné par le passage en coordonnées cartésiennes du vecteur de translation trouvé.

3.3.3 Algorithme de recalage complet

L'algorithme de recalage d'images complet est le suivant :

1. recherche de la rotation+homothétie (estimation de (ϕ, s))
 - se ramène à une recherche de translation en passant en coordonnées log-polaires
 - **oubli des effets de la translation** en considérant les amplitudes des spectres
2. inversion de la rotation+homothétie
(rotation d'angle $-\phi$ et homothétie de rapport $1/s$)
3. recherche de translation (estimation de (Δ_x, Δ_y)) en absence de rotation et d'homothétie
4. inversion de la translation
(translation de vecteur $(-\Delta_x, -\Delta_y)$)
5. inter-corrélation entre l'image obtenue et la référence (voir en suivant).

Le recalage proposé repose sur la recherche de pics (maxima) dans la transformée de Fourier-Mellin. Nous serons bientôt en mesure de proposer une méthode pour améliorer la précision de l'estimation des pics. Un problème inattendu soulevé lors des tests est que le pic de plus grande amplitude ne correspond pas toujours à la plus forte similarité (ce qui nous a obligé à tester plusieurs pics candidats, ralentissant d'autant la méthode). Enfin, il a été noté une forte sensibilité de la méthode au masque obtenu par le détournage. Ce problème n'était pas explicitement signalé dans l'article de référence [WZ00], or une solution doit être

trouvée afin de pouvoir étudier des images de fragments d'objets (essentiel pour les carreaux médiévaux, souvent brisés).

Le module logiciel de recalage d'images a été développé et testé. La méthode de recalage par transformée de Fourier-Mellin donne des résultats satisfaisant, à condition d'utiliser en pratique certaines astuces comme :

- travailler avec un fond noir (à 0), en rajoutant une bordure (par exemple travailler avec des images de taille double, pour éviter tout débordement lors des transformations) ;
- effectuer un fenêtrage avant toute transformée (pour éviter les problèmes liés au repliement spectral) ;
- oublier les basses fréquences (élimination de la zone imprécise dans les images log-polaires, et amélioration la corrélation de phase) ;
- essayer aussi l'angle $\phi + \pi$ pour la rotation (car même si le passage en log-polaire est fait sur un intervalle de 2π , une ambiguïté modulo π subsiste – du fait de la propriété de symétrie Hermitienne des spectres des images réelles. . .) ;
- faire attention aux visions finie et périodique (par exemple dans le cas de la translation). . .
- faire bien attention aux centres des repères (domaines spatial et spectral) ; en effet le centre n'est correctement défini que pour des images de taille impaire, mais la méthode semble être plus précise pour des images de taille paire. . .

3.4 Similarité

La similarité est une mesure de la distance entre deux images. C'est un problème très difficile, au cœur du projet IBISA qui s'intéresse de surcroît à des images très semblables au premier abord. Un bon candidat pour la mesure de la similarité est le facteur d'inter-corrélation :

$$\begin{aligned}
 F(g_1, g_2) &= \frac{\sigma_{12}}{\sigma_1 \sigma_2} \\
 &= \frac{\sum_P (g_1(p) - \bar{g}_1) \cdot (g_2(p) - \bar{g}_2)}{\sqrt{(\sum_P (g_1(p) - \bar{g}_1)^2) \cdot (\sum_P (g_2(p) - \bar{g}_2)^2)}} \\
 &= \frac{\sum_P g_1(p) \cdot g_2(p) - \bar{g}_1 \cdot \bar{g}_2}{\sqrt{(\sum_P (g_1(p))^2 - (\bar{g}_1)^2) \cdot (\sum_P (g_2(p))^2 - (\bar{g}_2)^2)}}
 \end{aligned}$$

($\bar{g}_n$ désignant la moyenne de g_n , et σ_n l'écart-type), généralisable dans le cas de présence de masques m_1 et m_2 en remplaçant g_n par $m_n \cdot g_n$ et $\bar{g}_n$ par $\sum_P m_n(p) \cdot g_n(p) / \sum_P m_n(p)$.

Nous obtenons de bons résultats en remplaçant g_1 par sa version **recalée** sur g_2 . Nous sommes conscients qu'il faut encore valider en pratique la mesure de similarité (notamment l'influence des conditions de prise de vue, de l'usure des objets, etc.). Les 1500 objets (monnaies) issus de la même paire de matrices (coins) du trésor d'Éauze (numérisation à venir, voir plus haut) devrait nous y aider grandement.

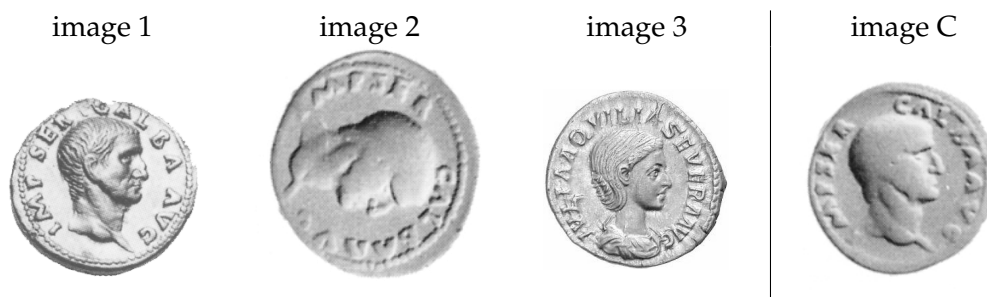


FIG. 12 – Quizz : l'image cible C étant à droite, classez les 3 images à sa gauche par ordre décroissant de similarité. (Réponse : 2, 1, 3 – 2 et C sont en fait issues de la même matrice, et 1 et C sont de même type.)

4 Bilan financier

Les 20 k€ alloués sur 2 ans au projet par l'ex-département ST2I du CNRS ont été utilisés de la façon suivante :

– modules logiciels (13,95 k€)	
– détournage	4,5 k€
– recalage	4,5 k€
– similarité + interface	4,95 k€
– missions (0,3 k€)	
– mission à Paris	0,15 k€
– réunion à Bordeaux	0,15 k€
– petit matériel (5,75 k€)	
– co-financement d'un portable (acheté par le CRP2A)	1,5 k€
– 4 disques durs (LaCIE Quadra D2) de grande capacité (1 To)	0,8 k€
– 2 scanners à plat format A4 (Epson Perfection V200)	0,17 k€
– 3 tablettes graphiques (Wacom Cintiq 12WX)	3 k€
– 1 Apple iPod Touch V2 16 Go	0,28 k€

total : 20 k€

Ceci respecte scrupuleusement le budget prévisionnel annoncé, avec 2 tiers pour la partie logicielle obtenue grâce à l'aide d'un prestataire, et 1 tiers pour du petit matériel informatique.

En fait, ce dernier tiers était initialement prévu pour les missions, qui seront finalement faites après mars 2009 sur des crédits Région et/ ou de recherche des laboratoires partenaires. Dans l'immédiat, il a été jugé utile que :

- chaque partenaire dispose d'un espace de stockage important pour les bases de données (disques durs externes) ;
- les partenaires pour la numismatique disposent d'un dispositif d'acquisition d'images de monnaies (scanners à plat) ;
- chaque partenaire dispose d'une tablette graphique permettant une plus grande facilité d'utilisation de l'interface du logiciel, notamment en ce qui concerne le détournage des motifs des carreaux (qui reste encore principalement manuel) ;
- la faisabilité d'une version du logiciel pour terminaux mobiles communiquants soit étudiée (développement sur Apple iPod Touch en vue d'utilisation sur un iPhone).

5 Conclusion et perspectives

Signalons que des membres de la proposition de projet initiale ne sont plus physiquement à Bordeaux. En effet, Koray Konuk est actuellement à l'Institut Français d'Études Anatoliennes, Istanbul. De plus, Jacques-Olivier Lachaud a été recruté comme Professeur à l'Université de Savoie et a donc quitté le LaBRI. Anne Vialard (Maître de Conférences au LaBRI) a pris le relais pour la partie détournage, mais a été contrainte de prendre un congé long en 2008. Béatrice Cicuttini a eu, elle, un heureux événement – et donc également un congé. Alain Bresson est parti à l'Université de Chicago, USA. Quant au porteur du projet, Sylvain Marchand, il a profité d'une délégation CNRS pour accepter des invitations outre-Atlantique (universités McGill au Canada et Stanford aux USA) dans la première moitié de l'année 2008.

De plus, nous remarquons que l'impossibilité de faire appel à un contractuel a rajouté une difficulté supplémentaire au développement de la maquette logicielle.

Il est clair quel projet IBISA a donc pris du retard, mais à ce jour nous disposons de l'outil informatique souhaité, capable de :

- détourner un objet dans une image, semi-automatiquement (tablette graphique et / ou modèles déformables *snakes*) ;
- recalcr une image source sur une image destination, manuellement (par transformations géométriques) ou automatiquement (par recherche de maxima dans la transformée de Fourier-Mellin) ;
- déterminer la similarité des deux images, manuellement (affichage par transparence) et / ou automatiquement (facteur d'inter-corrélation) ;

ce qui permet déjà en pratique de :

- gérer des bases de données d'images d'objets archéologiques ;
- rechercher une image à l'intérieur de ces bases à **partir d'une autre image** (dans le futur, cette recherche pourrait être initiée à distance, par exemple *via* un téléphone mobile) ;
- étudier exhaustivement les correspondances entre les images des bases de données (par exemple pour l'étude semi-automatique de pavements et de trésors monétaires – étude des correspondances / liaisons de coins, mais on peut aussi imaginer la vérification de contrefaçon ou la recherche d'objets volés) ;

même s'il nous reste à :

- généraliser la méthode de recalage au cas de fragments d'objets ;
- proposer une méthode de détournage adaptée aux motifs des carreaux médiévaux ;
- déterminer un représentant idéal d'une classe de plusieurs objets ;

et bien entendu à utiliser le logiciel dans des cas archéologiques concrets, ce qui devrait être fait dans un futur proche.

Le site du projet IBISA, qui évoluera même après l'arrêt du financement PEPS, est accessible en ligne à l'adresse :

<http://dept-info.labri.fr/~sm/Projets/IBISA>

6 Bibliographie

Références

- [AGPB⁺05] Ayed Ben Amara Ayed, Nicole Gourdon-Platel, Françoise Bechtel, Max Schvoerer, and Philippe Bon. Carreaux glaçurés provenant d'un château du Duc de Berry (Mehun-sur-Yèvre, fin XIV^e siècle) : Recherche d'indices techniques. *AchéoSciences*, pages 21–34, 2005.
- [AKKM02] Richard H. J. Ahston, Philip Kinns, Koray Konuk, and Andrew R. Meadows. The Hecatomnus Hoard (CH 5.17, 8.96, 9.387). *Coin Hoards*, 9 :95–158, pl. 6–20, 2002.
- [BM02] Jean-Pierre Bost and Thierry Martin. Un trésor d'aurei julio-claudiens à Montans (Tarn). *Trésors monétaires*, 20 :175–184, 2002.
- [Bre05] Alain Bresson. *Making, Moving, Managing, The New World of Ancient Economies (323-31 BCE)*, chapter Coinage and Money Supply in the Hellenistic World, pages 44–72. Oxbow Books, 2005.
- [BV08] Alain Bresson and Despoina Vallianou. *Le trésor de Phaïstos*. Bordeaux, 2008. À paraître (trésor de 600 monnaies hellénistiques).
- [Cic06] Béatrice Cicuttini. Les carreaux estampés de la forteresse médiévale de Blanquefort (XIV^e siècle, Gironde) : Étude comparée des différentes séries typologiques identifiées dans la collection. Master's thesis, IRAMAT-CRP2A, Université Michel de Montaigne – Bordeaux 3, juin 2006.
- [dSDS⁺04] Baudoin Denis de Senneville, Pascal Desbarats, Rares Salomir, Bruno Quesson, and Chrit T. W. Moonen. Correction of Accidental Patient Motion for On-line MR Thermometry. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 321, pages 637–644, 2004.
- [dSQDM05] Baudoin Denis de Senneville, Brunon Quesson, Pascal Desbarats, and Chrit T. W. Moonen. 3D Motion Estimation for On-Line MR Temperature Mapping. In *Proceedings of IEEE International Conference on Image Processing (ICIP)*, 2005.
- [Eam81] Elisabeth S. Eames. *Catalogue of Medieval Lead-Glazed Earthenware Tiles in the Department of Medieval and Later Antiquities*. British Museum Publications, London, 1981. Un volume de texte de 783 pages + 8 planches couleurs ; un volume d'illustrations, regroupant les motifs inventoriés dans les collections du British Museum. Soit l'inventaire de 13822 carreaux, et un répertoire de plus de 3000 motifs (qui constitue le volume 2).
- [Kon98] Koray Konuk. *The Early Coinage of Kaunos*, chapter Greek Numismatics in Memory of Martin Jessop Price, pages 197–223, pl. 47–50. London, 1998.
- [KWT88] Michael Kass, Andy Witkin, and Demetri Terzopoulos. Snakes : Active Contour Models. *International Journal of Computer Vision*, pages 321–331, 1988.
- [LT05] Jacques-Olivier Lachaud and Benjamin Taton. Deformable Model with a Complexity Independent from Image Resolution. *Computer Vision and Image Understanding*, 99(3) :453–475, 2005.

- [Mar04] Sylvain Marchand. Un outil informatique pour l'étude automatique de la correspondance des coins. In *Revue archéologique de Bordeaux*, volume XCV, page 253. Société archéologique de Bordeaux, 2004.
- [Mar06] Sylvain Marchand. Numismatique à la Société archéologique de Bordeaux. Cours public dans le cadre du cycle de conférences au Musée d'Aquitaine (Bordeaux) : Regards sur la Société archéologique de Bordeaux (1873–2005), 1^{er} février 2006.
- [Org04] Magali Orgeur. *Les carreaux de pavement des abbayes cisterciennes en Bourgogne (fin XII^e-fin XIV^e siècles)*. PhD thesis, Université de Bourgogne, juin 2004. Cette thèse comprend deux tomes de texte (809 p.), un tome de planches (400 p.) et un CD-ROM d'inventaire.
- [PPS08] Nicolas Parisot, Michel Prieur, and Laurent Schmitt. *Rome XXII*. Compagnie Générale de Bourse (CGB), Paris, 2008.
- [Sea00] David R. Sear. *Roman Coins and Their Values*, volume 1. Spink and Son Ltd, London, the millennium edition, 2000.
- [Thi72] Marcel Thirion. *Le trésor de Liberchies*. Pro Geminiaco a.s.b.l., Bruxelles, 1972.
- [TL02] Benjamin Taton and Jacques-Olivier Lachaud. Deformable Model with non-Euclidean Metrics. In A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, editors, *Proceedings of the European Conference on Computer Vision (ECCV)*, volume 2352 (part III) of *Lecture Notes in Computer Science (LNCS)*, pages 438–453. Springer, Berlin, 2002.
- [WZ00] George Wolberg and Siavash Zokai. Robust Image Registration Using Log-Polar Transform. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, septembre 2000.
- [ZKS08] Sebastian Zambanini, Martin Kampel, and Mario Schlapke. On the Use of Computer Vision for Numismatic Research. In *Proceedings of the 9th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST)*, pages 17–24, 2008.