

# Réaliser un fichier d'empreintes digitales

Rémi Belloeil<sup>(\*)</sup>

Comment réaliser une base de données d'empreintes digitales avec des élèves puis l'exploiter ? C'est ce que cet article va vous présenter.

## Le cadre

L'option « Méthodes et Pratiques Scientifiques » comporte un thème « investigations policières » ; c'est dans ce cadre que j'ai réalisé cette activité. Les conditions étant très variables d'un lycée à l'autre, je vais les décrire ici brièvement. Dans mon lycée, deux enseignants pris parmi les trois disciplines Mathématiques, Physique-Chimie et Sciences de la Vie et de la Terre (SVT) animent un ou deux groupes de 16 à 25 élèves à raison de 2h par semaines pendant 27 semaines (soit l'équivalent de 1h30 par semaine en moyenne sur 36 semaines), payées 45 minutes par groupe et par semaine toute l'année (c'est-à-dire la moitié du temps élèves). Le premier thème était « investigations policières » ; j'ai réalisé une base de données d'empreintes digitales dans un groupe et ma collègue une base concernant les fibres dans l'autre groupe, puis nous avons échangé les groupes ; ensuite nous avons travaillé alternativement sur le thème « Vision du monde ». En fait, nous sommes intervenus ma collègue de SVT et moi alternativement, sur des plages de 6 à 7 semaines consécutives.

## Le but

Il s'agissait de réaliser un fichier EXCEL dans lequel seraient décrites les empreintes des dix doigts de tous les élèves participants, soit 250 empreintes pour le premier groupe et 160 pour le deuxième, puis de retrouver une empreinte. La question intéressante est de savoir si on peut décrire une empreinte indépendamment du repère cartésien initial. Bien sûr, il y a l'utilisation des outils informatiques et en particulier du scanner et du tableur, mais aussi le traitement d'une image avec un logiciel (Mesurim), réutilisé dans la partie SVT, et une réflexion sur la confidentialité des données. Mais le but était aussi de mener un travail collectif où la part de chacun serait importante pour le groupe.

## L'organisation

Lors de la première séance pour introduire le thème nous avons regardé un épisode de Numbers (« Le coupable idéal ») où la question de l'orientation d'une empreinte digitale apparaît et les élèves ont posé des questions sur le sujet. À la seconde, le but a été annoncé et une réflexion collective a eu lieu sur la façon d'y arriver.

---

(\*) remi.belloeil@orange.fr

Les étapes ont été ainsi listées :

- Relever les empreintes avec un tampon encreur (RE),
- Scanner les empreintes (RS),
- Pour chaque image, relever les coordonnées de 6 minuties (points particuliers dans les empreintes digitales, essentiellement les points où une ligne se sépare en deux) (MC),
- Calculer les distances entre ces 6 points deux à deux, puis la somme de ces distances (CC),
- Parallèlement, créer des catégories pour décrire chaque empreinte (spirale à droite ou à gauche, ...) (CE),
- Réaliser un tableau de EXCEL où chaque ligne représentera une empreinte, avec le nom de l'image initiale, sa description, les coordonnées des 6 points, les distances entre eux, la somme de ces distance (BD),
- Trouver le traitement qui permettrait à partir des coordonnées de 6 points d'une empreinte déjà relevée, mais éventuellement dans un autre ordre et dans un autre repère, de retrouver le nom de l'image correspondante dans le tableau de EXCEL (RF),
- Pendant tout le traitement garantir que les images et les relevés ne seront pas diffusés et même empêcher tout lien entre une image et le propriétaire de l'empreinte ; s'assurer de la destruction de tous les fichiers informatiques et tous les supports papiers utilisés (PC).

Chacune de ces étapes a été confiée à un groupe en précisant que tous les élèves devaient participer à toutes les étapes et donc être formés par les « experts » des différents groupes.

Cela fait donc 8 étapes ; le problème est qu'elles interviennent successivement et qu'il faut trouver un moyen d'occuper tout le monde pour que le travail soit efficace. Par exemple, lorsqu'on effectue le relevé des empreintes à l'encre, même avec plusieurs tampons, seuls quelques élèves seront occupés et lorsqu'on scanne les images, il faudrait avoir plusieurs scanners identiques de préférence pour éviter l'embouteillage à ce poste. De plus, au moment où des élèves sont inoccupés, l'enseignant est débordé par les questions Comment marche le scanner ?, Comment nommer les images ?, Où les enregistrer ? Pour éviter ce problème, j'avais donné à chaque groupe un travail préparatoire à faire, ce qui amenait d'autres questions de la part des élèves. Dois-je préciser que les élèves sont loin d'être autonomes et peuvent « glander » alors qu'ils n'ont pas répondu à toutes les demandes de l'enseignant ? Gérer des élèves qui réalisent des tâches différentes avec des outils différents est assez difficile. Le lecteur aura compris que je ne recommande pas de reproduire à l'identique cette activité, mais je pense qu'il peut tirer des idées intéressantes et peut-être éviter certains problèmes.

## Le rôle et les problèmes de chaque groupe

### RE : Relever les empreintes avec un tampon Encreur

Le groupe doit prendre les empreintes des dix doigts de tous les élèves en s'assurant qu'elles sont chacune dans un petit carré d'environ 3 cm de côté avec des axes gradués en cm pour permettre les mesures ultérieures. Chacune ne doit être ni trop claire ni trop foncée, il faut les contrôler au fur et à mesure, rayer et recommencer si nécessaire. Le protocole interdit de marquer le nom ou la référence à un numéro identifiant (cette contrainte pourrait être contournée par un système d'anonymat lié à un tableau qui serait ensuite détruit). Les empreintes d'une même personne seront disposées aléatoirement sur la feuille. Il faut prévoir l'accès à un lavabo et du savon ! Comme ce groupe est surtout occupé au début, on peut lui confier comme deuxième tâche RF qui n'interviendra qu'à la fin. Un membre du groupe PC doit être détaché au contrôle de ce groupe dès le début.

### RS : Relevé des empreintes avec le Scanner

Ce groupe doit disposer de plusieurs scanners, identiques et connectés au réseau si possible et c'est le principal problème matériel. Il doit apprendre à utiliser les scanners, éventuellement former les autres. Il faut qu'il contrôle la qualité des images et s'assure qu'elles aient toutes des noms différents, et qu'il n'y ait pas de perte à ce niveau. En cas de mauvaise qualité du support il doit faire reprendre une empreinte si le protocole d'anonymat le permet. Un membre du groupe BD (Base de données) doit noter tous les noms d'images. Un membre du groupe PC (Protocole et Confidentialité) doit s'assurer qu'il n'y a aucun lien entre les noms d'images et les propriétaires des empreintes, il doit aussi surveiller dans quels répertoires sont stockées les images et s'assurer qu'elles ne seront pas à disposition de personnes extérieures à la classe puis qu'elles seront effacées dès que possible ; les supports des empreintes doivent être conservés et transmis au groupe CE (Catégorie des Empreintes) ou détruits par eux. Techniquement, il y a là des problèmes de droits d'accès à des répertoires avec la nécessité de transférer ces images au groupe MC (Détection des Minuties et relevé des Coordonnées).

### MC : Détection des Minuties et relevé des Coordonnées

Après s'être renseigné sur les minuties et sur le fonctionnement du logiciel Mesurim, ce groupe est responsable de la définition de 6 points et du relevé de leurs coordonnées à l'aide de ce logiciel. Il est possible de définir un repère sur une première image qui sera ensuite réutilisé pour les images suivantes, mais l'unité doit être la même pour toutes (mm, cm ou pixel ?). Je n'ai pas vu de moyens d'enregistrer directement les coordonnées dans un fichier. Il faut donc écrire le nom de l'image et noter les six coordonnées pour les transmettre au groupe CC chargé des calculs sur les coordonnées. Les membres de ce groupe doivent montrer aux autres comment procéder.

Le nombre de points a été limité à 6 pour ne pas permettre d'identification, être réalisable sur les empreintes et simplifier le traitement par l'uniformisation des



**BD : Réalisation de la Base de Données**

La réalisation de la table consiste essentiellement à écrire les titres dans la première colonne et pourrait être confiée au groupe précédent, mais le rôle de ce groupe est surtout de s'assurer de ne pas perdre des données en route, de s'assurer qu'il n'y aura pas de doublons, que toutes les images auront été soumises au relevé des coordonnées puis au calcul des distances et enfin apparaissent bien dans le fichier final.

**CE : Catégories des Empreintes**

À côté des caractérisations numériques des empreintes, il est intéressant, si le groupe d'élèves est nombreux, de rechercher un aspect qualitatif. On trouve assez facilement des informations sur cet aspect sur Internet. Si on considère seulement les distances, deux empreintes symétriques par rapport à un axe sont identiques et cet aspect qualitatif permet de les distinguer. En réalité, il a été difficile d'intégrer le travail de ce groupe dans le travail collectif, parce que cette distinction est moins évidente à comprendre et à transmettre.

**RF : Recherche dans le Fichier**

Pour que ce groupe travaille avant que le fichier ne soit constitué, il faut lui fournir un fichier fictif où les données sont obtenues à partir de tirages aléatoires (mais figés). À partir d'une empreinte de ce fichier, l'enseignant réalise un relevé fictif en changeant l'ordre des points et les coordonnées par un changement de repère (par une isométrie, une translation suffit). Cette ligne supplémentaire qui ne comporte pas les distances entre les points, est fournie au groupe qui doit retrouver la référence de l'empreinte utilisée. Rechercher plusieurs nombres dans différents champs est difficile, c'est pourquoi utiliser la somme comme unique indicateur est performant. Après avoir calculé les distances et leur somme, les élèves peuvent effectuer un tri sur la somme et retrouver l'empreinte dans le tableau, en cas d'ex aequo, il suffit de regarder les distances entre les coordonnées, mais cela n'a jamais été nécessaire.

Lorsque le fichier a été réalisé, chaque groupe s'est vu confier une ligne différente correspondant à une empreinte du fichier et a dû mener sa recherche, guidé par les membres du groupe RF. Et cela s'est réalisé sans problème.

**PC : Respect du Protocole et de la Confidentialité**

Une autorisation avait été demandée aux familles pour réaliser ce fichier en indiquant un certain nombre de garanties : anonymat des images, destruction des supports et des fichiers électroniques à la fin de l'activité. Le groupe PC avait pour mission de contrôler le processus de conserver les supports le temps nécessaire (prévoir des pochettes), de les détruire ainsi que de surveiller les fichiers électroniques créés, leur mise à disposition limitées aux seuls utilisateurs puis leur destruction (sous la direction de l'enseignant). Ceci s'est plutôt bien passé dans le premier groupe, le deuxième a été moins vigilant.

## Bilan

Avec le premier groupe de 25 élèves, les pertes ont été si nombreuses qu'il n'y a eu qu'une trentaine d'empreintes dans le fichier final ; ce qui est un certain échec pour le travail collectif. Le deuxième groupe a été mieux dirigé grâce à l'expérience du premier et a conduit à une centaine d'empreintes. Chaque groupe devait faire un rapport. Cela reste assez pauvre au regard de l'énergie développée par l'enseignant et des problèmes techniques à résoudre ! Cependant, les élèves ont incontestablement découvert une façon d'aborder un problème, de calculer à partir de coordonnées à l'aide d'un tableur et perçu le rôle du repère ainsi qu'une façon de s'en affranchir par le calcul des distances. Un travail directement à partir d'un fichier fictif est certainement plus reposant, moins salissant et peut-être aussi formateur pour l'aspect mathématique. Mais il fait disparaître la difficulté de se confronter à la réalité, l'utilisation d'un logiciel pour analyser une image, le problème du choix des minuties, ... Je ne doute pas que le lecteur saura tirer profit de cette expérience pour réaliser une activité où il reprendra une partie seulement de ce travail et conservera l'essentiel.

Le lecteur mathématicien qui souhaite plus de détail sur le traitement mathématique, consultera avec intérêt le travail réalisé dans le cadre d'un mémoire de Master de Didactique sur la Modélisation à l'adresse :  
<http://www.mathom.fr/mathom/FeteDeLaScience/FS2007/Complements/Empreintes%20digitales.pdf>