

Etude de l'effet de l'amodiaquine sur les globules rouges infectés par le paludisme dans les images multispectrales

Olivier K. BAGUI¹, William YAVO², Dominique TANO² et Jeremie T. ZOUEU^{1*}

¹Laboratoire d'Instrumentation Image et Spectroscopie, Institut National Polytechnique Houphouët-Boigny,
BP 1093 Yamoussoukro, Côte d'Ivoire

²UFR Pharmacie, INPH, 01 BPV 34 Abidjan, Côte d'Ivoire

* Correspondance, courriel : jeremie.zoueu@inphb.edu.ci

Résumé

Nous avons montré que l'amodiaquine se fixe préférentiellement dans les globules rouges infectés et nous avons également été capable de retrouver les concentrations du médicament dans les cellules, à partir d'images multispectrales. Une scène d'images multispectrales de frottis sanguins non marqués est obtenue à partir d'une culture de plasmodium falciparum, à laquelle nous avons ajouté différentes doses d'amodiaquine (AQ) ; Les images sont obtenues à partir d'enregistrement de treize capteurs spectraux couplés à un microscope multimodal et multispectral. Ces images sont ensuite utilisées pour étudier l'interaction du médicament avec les globules rouges en fonction des concentrations, en ayant recours à des techniques d'analyse multivariée telles que la classification hiérarchique, la méthode des k-moyennes et l'analyse en composante principale. Les résultats obtenus nous montrent que l'imagerie multispectrale est un atout majeur pour le diagnostic médical et peut donc constituer une technique de routine pour l'étude de nouvelles molécules antipaludiques, notamment issues de la pharmacopée africaine.

Mots-clés : *imagerie multispectrale, classification hiérarchique, analyse en composante principale, k-moyennes.*

Abstract

Study of the Effect of Amodiaquine on Red Blood Cells Infected by Malaria in Multispectral Images

We have demonstrated that amodiaquine binds preferentially in the infected red blood cells and we have also been able to retrieve the drug concentrations in the cells. A scene of multispectral images of free labelled blood smears have been obtained from plasmodium falciparum culture to which we have added different quantities of amodiaquine; the images have been recorded from thirteen spectral sensors coupled with a multimodal and multispectral microscope. These images have been used to study the interaction of the drug with the red blood cells as function of drug concentration by the use of multivariate statistical analysis techniques such as hierarchical classification, k-means method and principal component analysis techniques. The results show that the multispectral imagery is a key technique in medical technologies and can therefore be used as routine method for antimalarial drug design, especially in African traditional pharmacopeia study.

Keywords : *multispectral imagery; hierarchical classification; principal component analysis; k-means.*

1. Introduction

Un pixel dans une image multi-spectrale [1] capturée par un microscope à transmission peut être considéré comme un mélange de signatures optiques de plusieurs chromophores. Il est donc possible d'extraire les paramètres optiques de ce mélange correspondant aux quantités relatives à ces éléments. L'un des avantages de l'imagerie multi-spectrale réside dans le fait que plusieurs constituants différents peuvent être caractérisés et identifiés simultanément par leurs signatures spectrales. Parmi des paramètres quantitatifs tels que la taille de cellules, la forme, la texture, l'analyse des spectres obtenus par imagerie multi-spectrale est un facteur important pouvant être utilisé en plus, pour permettre la détection et l'identification appropriées de cellules dans les images aussi bien que l'effet du médicament. L'infection des globules rouges par le plasmodium est un problème clinique important en Afrique, qui est la zone la plus endémique; la recherche de nouvelles molécules antipaludéennes issues de la pharmacopée africaine ou par des méthodes modernes est abondante. Considérant que l'étape-clé de l'action des molécules anti-malariques réside en leur association covalente ou non avec l'hème, Nous proposons une étude basée sur les spectres d'absorptions des globules rouges et utilisant la classification hiérarchique ascendante (CHA), la méthode des k-moyennes et l'analyse en composantes principales (ACP), pour montrer qu'il est possible par imagerie multi-spectrale d'étudier l'interaction de différentes concentrations d'amodiaquine (AQ) sur les globules rouges infectés. Cette technique pourra ensuite être utilisée comme un instrument de routine.

2. Matériel et méthodes

2-1. Analyse en Composante Principale (ACP)

L'ACP, développée en France dans les années 1960 par J- P. Benzécri, est une méthode de statistique exploratoire permettant de décrire un grand tableau de données de type individus / variables. Lorsque les individus sont décrits par un nombre important de 5 variables, aucune représentation graphique simple ne permet de visualiser le nuage de points formés par les données. L'ACP propose une représentation dans un espace de dimension réduite, permettant ainsi de mettre en évidence d'éventuelles structures au sein des données. Pour cela, nous recherchons les sous-espaces dans lesquels la projection du nuage déforme le moins possible le nuage initial [2]. Les données sont sous la forme d'un tableau à n lignes et p colonnes, que l'on stocke sous la forme d'une matrice X de taille $n \times p$.

2-2. Classification Hiérarchique Ascendante (CHA)

Comme les autres méthodes de l'analyse des données, dont elle fait partie, la classification hiérarchique a pour but d'obtenir une représentation schématique simple d'un tableau rectangulaire de données dont les colonnes, suivant l'usage, sont des descripteurs de l'ensemble des observations, placées en lignes [3]. Il existe deux types de méthodes de classification hiérarchique:

- la méthode hiérarchique descendante;
- La méthode hiérarchique ascendante.

L'étude que nous avons effectuée est basée sur les méthodes hiérarchiques ascendantes. Initialement, le nombre de pixels reflète le nombre de classes. Les pixels les plus similaires sont par la suite répartis en groupes d'observations homogènes, chaque groupe étant bien différencié des autres [4,5]. Cependant, cet objectif est plus raffiné ; on veut, en général, obtenir des sections à l'intérieur des groupes principaux, puis des subdivisions plus petites de ces sections, et ainsi de suite.

En bref, on désire avoir une hiérarchie, c'est-à-dire une suite de partitions "emboîtées", de plus en plus fines, sur l'ensemble d'observations initiales. Une telle hiérarchie peut être résumée par un arbre hiérarchique appelé dendrogramme.

2-3. Méthode des « K-moyennes »

La méthode des K-moyennes est une technique de quantification vectorielle, c'est aussi un algorithme de minimisation alternée qui étant donné un entier K, va chercher à séparer un ensemble de points en K classes [6,7]. La particularité de la méthode des "K-moyennes" c'est que le nombre de classes doit être spécifié préalablement. La méthode la plus utilisée pour estimer ce nombre c'est de mener une classification hiérarchique sur un échantillon représentatif de l'ensemble des individus. Pour représenter chacune des K classes C_k , on utilise la moyenne (ou la moyenne pondérée) π_k de ses points, appelée centroïde (ou centre de masse). Chacune des C composantes du vecteur π_k est calculée par *l'équation (1)*:

$$\pi_{kj} = \frac{1}{|C_k|} \sum_{O_i \in C_i} P_{ij} \tag{1}$$

Dans le cas de données numériques, cela donne un sens géométrique et statistique à la méthode. L'inertie intra-classe constitue le critère à optimiser. Elle est définie comme la moyenne des carrés des distances des objets de la classe au centre de gravité de celle-ci. On cherche ainsi à construire des classes compactes. L'inertie intra classe associée à la classe C_k s'écrit comme *l'équation (2)*

$$I_k = \frac{1}{|C_k|} \sum_{O_i \in C_i} d^2(O_i, \pi_{kj}) \tag{2}$$

L'objectif est alors de minimiser la somme de l'inertie intra-classe sur l'ensemble des classes. L'algorithme procède en deux étapes : dans la première phase, on réassigne tous les objets au centroïde le plus proche, et dans la deuxième phase, on recalcule les centroïdes des classes qui ont été modifiées. Pour mesurer la proximité entre un centroïde et un objet, on calculera une distance entre ces deux vecteurs. On pourra utiliser, par exemple, la distance euclidienne. Les deux phases sont itérativement répétées jusqu'à ce qu'un critère d'arrêt soit atteint (par exemple, si aucune modification n'a eu lieu, ou si le nombre maximum d'itérations a été atteint). Les principaux problèmes de l'approche des k-moyennes comme des autres approches partitionnelles, sont l'influence de la partition initiale (qui est souvent choisie de façon aléatoire), et le choix du paramètre K qui n'est pas toujours évident. Soit $P_0 = \{C_1, \dots, C_k\}$

2-4. Préparation des échantillons

La technique de micro test de l'OMS a été utilisée pour mesurer l'inhibition de la maturation des schizontes par microscopie. Les érythrocytes infectés lavés ont été mises en suspension dans du RPMI 1640 avec 10 % de sérum humain, HEPES 25 mM, NaHCO₃ 25 mM et à un taux d'hématocrite de 1,5 %. Cinquante microlitres du mélange de sang-suspension ont été distribués dans chaque puits des plaques à 96 puits pour culture de tissus pré-dosés et incubés à 37 ° C dans des bocal de bougie durant 42 heures selon la méthode standard. Les concentrations finales étaient réparties entre 12,5 et 1600 nM pour l'amodiaquine. Après la période d'incubation, les parasites ont été recueillis et les frottis sanguins sans coloration ont été préparés. Le nombre de schizontes, définie comme schizontes ayant plus de 3 noyaux, a été prise en compte par 200 parasites asexués. Les isolats de moins de 20 % de schizontes dans les échantillons de contrôle et sans médicaments ont été exclus.

Les résultats ont été exprimés comme 50 % des valeurs de concentration inhibitrice (CI50). Les valeurs seuils de résistance in vitro à la chloroquine et l'amodiaquine ont été fixées à 60 nM et 100 nM respectivement. Sept différentes concentrations a l'amodiaquine ont été utilisées dans cette étude (C1=12,5 ; C2=25 ; C3=50 ; C4=100 ; C5=200 ; C7=800 ; C7=1600 mol.L⁻¹).

2-5. Acquisition et prétraitement des images

Le système d'acquisition de données a été décrit précédemment [8-10], Les sources lumineuses classiques, en transmission et en réflexion ont été enlevé et remplacé, pour tous les modes, par un ensemble de 13 DELs avec des longueurs d'onde allant de 375nm à 940nm. Une source en anneau est fixée au système à l'aide d'une fibre optique pour le mode diffusion. Les oculaires mécaniques initiaux sont de même substitué par une caméra monochrome CMOS à 12 - bit (2592 x 1944, Guppy - 503B, Vision Allied Technology, avec un capteur de MT9P031 de Micron / Aptina), avec une taille de pixel de 2,2 µm x 2.2 µm, qui est utilisée pour acquérir des images. Pour les modes de transmission, de réflexion et de diffusion, le système acquière automatiquement un total de 39 images spectrales (13 images par mode), pour la même scène, en utilisant une carte d'acquisition de données (NI-DAQ) couplée à un ordinateur, qui contrôle les intensités de courant pour les réglages fins. Nous avons dans ce travail, utilisé le mode Transmission pour visualiser l'effet du médicament, car il révèle mieux la présence des substances à l'intérieur de la cellule. Les treize sources lumineuses sont suffisantes pour étudier les propriétés optiques des cellules biologiques [8]. Pour effectuer les mesures de nos images multi spectrales avec les différents modes, la camera capture les images transmises à travers l'objectif. L'intensité lumineuse $I_{xy\lambda}$ enregistrée par chacun des canaux de la camera est estimée à l'aide du modèle suivant [10] de *l'équation (3)*:

$$I_{xy\lambda} = \int_0^{\infty} E_{xy\lambda} T_{xy\lambda} S_{xy\lambda} d\lambda \quad (3)$$

$E_{xy\lambda}$ le spectre d'émission (ou irradiance) de la source lumineuse utilisée ; $T_{xy\lambda}$ le spectre optique de l'échantillon analysé ; $S_{xy\lambda}$ la sensibilité spectrale du capteur ; x et y sont les coordonnées des pixels ; λ la longueur d'onde

Cette expression traduit le fait que la mesure effectuée par chacun des canaux de la camera, représente une valeur moyenne effectuée sur le produit de convolution du spectre optique réel attendu de la mesure et des contributions de l'ensemble du système d'acquisition. Pour obtenir des images spectrales en transmission, nous procédons comme suit pour chaque longueur d'onde :

- Capture de l'image transmise par l'échantillon sur lame ($S_{xy\lambda}$);
- Capture de l'image de la lame vide ($R_{xy\lambda}$); mesure dite de référence;
- Capture de l'image en obstruant le passage de la lumière ($B_{xy\lambda}$) : mesure dite du bruit de fond
- L'image finale $IT_{xy\lambda}$ s'obtient comme dans *l'équation (4)*:

$$IT_{xy\lambda} = \frac{S_{xy\lambda} - B_{xy\lambda}}{R_{xy\lambda} - B_{xy\lambda}} \quad (4)$$

3. Résultats

Différents frottis sanguins non colorés ont été analysés dans cette étude, pour chaque échantillon, nous avons procédé à l'extraction des intensités moyennes des pixels dans chaque globule rouge pour chaque image et chaque concentration que nous avons regroupée dans une base commune. Une première étude visant à connaître le nombre de classes dans notre base, fondée sur le critère de la silhouette révèle l'existence de deux grands groupes (**Figure 1**). Notons que la valeur de la silhouette pour chaque point est une mesure de la façon dont ce point est similaire aux points de son groupe par rapport à des points dans d'autres groupes [11]. La valeur moyenne maximale de la silhouette correspond au nombre de classe réel.

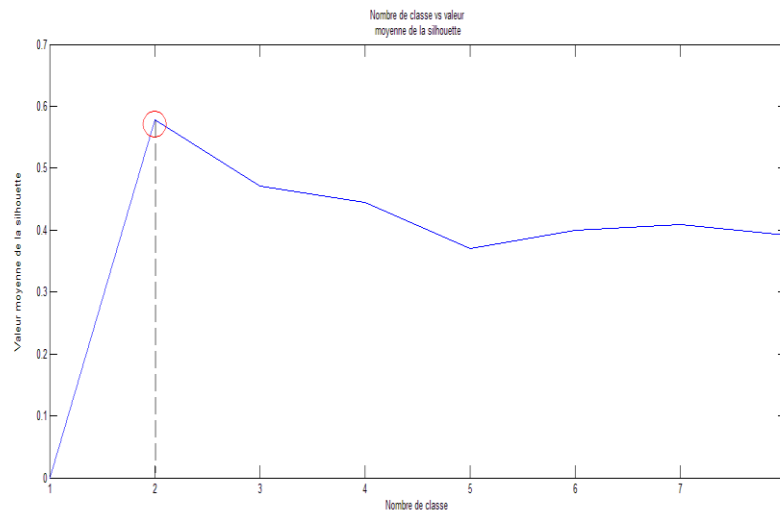


Figure 1 : représentation des valeurs moyenne de la silhouette en fonction du nombre de classes

La représentation de la silhouette révèle l'existence de deux classes dans notre base. Nous procédons ensuite à une répartition en deux classes à partir des deux techniques que sont la classification hiérarchique ascendante (CHA) (**Figure 2**) et la méthode des K-moyennes, (**Figure 3**) afin de différencier les éléments des deux classes et confirmer qu'il s'agit des mêmes éléments.

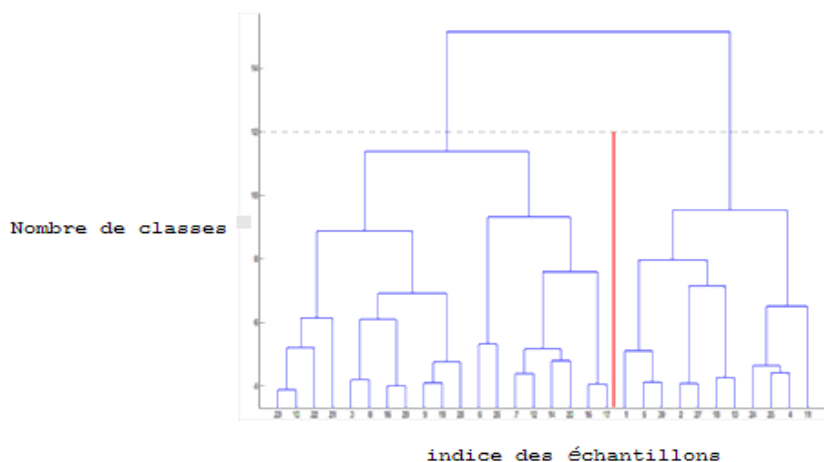


Figure 2 : dendrogramme résumant la classification hiérarchique ascendante

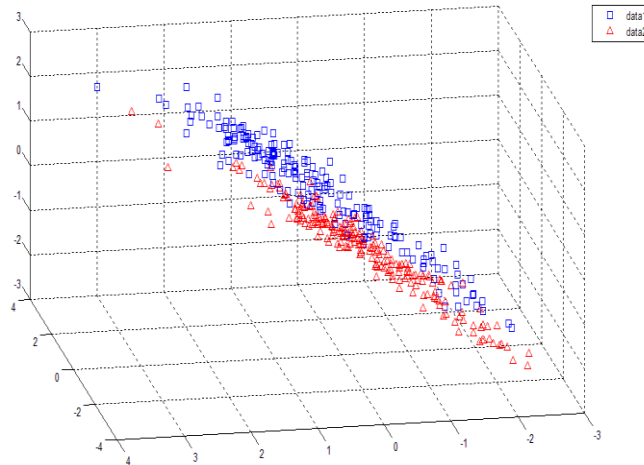


Figure 3 : représentation des points de la base selon les deux classes par la méthode des K-moyennes

L'analyse des spectres moyens d'absorption de ces deux groupes de données nous montre qu'il s'agit pour le premier groupe d'érythrocytes non-infectés (**Figure 4**) et pour le second, d'érythrocytes dont les spectres sont différents des précédents (**Figure 5**).

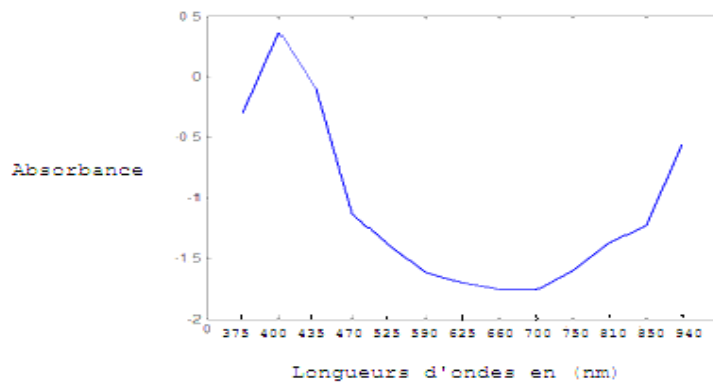


Figure 4 : spectre d'absorption du groupe 1

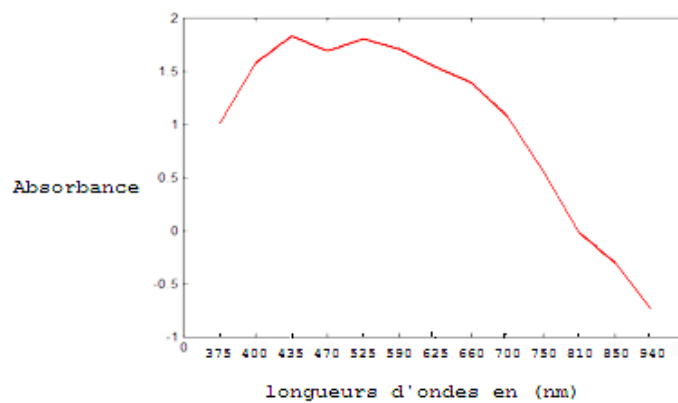


Figure 5 : spectre d'absorption du groupe 2

Le spectre moyen du groupe 2 rend compte de la présence d'une nouvelle substance dans les cellules: l'amodiaquine. Un complexe entre l'hémoglobine et l'amodiaquine est créé lors de l'interaction entre le parasite et le médicament, justifiant l'hypothèse selon laquelle le médicament se fixe préférentiellement dans la cellule infectée. L'application de l'ACP sur les intensités moyennes en fonction des longueurs d'onde du groupe 2 nous a permis de comparer les composantes principales CP_i (C_i) entre elle. CP_i (C_i) représente la composante principale i pour la concentration C_i (**Figure 6**) et d'isoler la composante 4. (**Figure 7**)

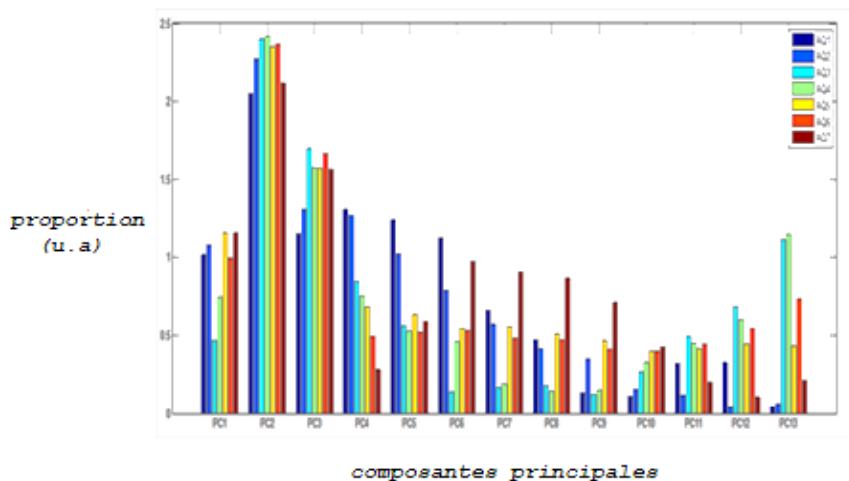


Figure 6: évolution de l'intensité des pixels en fonction de la concentration de l'Amodiaquine

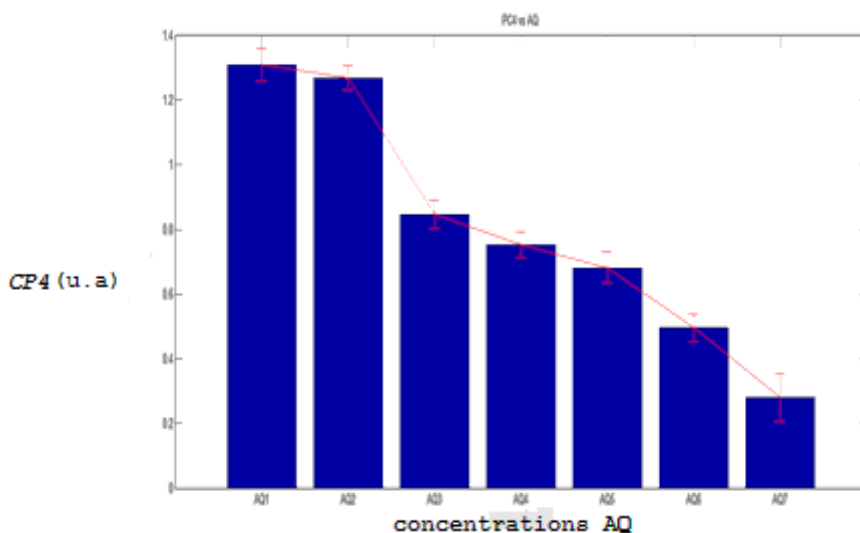


Figure 7 : Evolution de la valeur de la composante principale 4 (PC4) en fonction de la concentration

4. Discussion

Les deux méthodes de classification que nous avons utilisées, ont, de façon indépendantes l'une de l'autre, permis de regrouper les cellules en deux groupes similaires en fonction de leurs empreintes spectrales; le premier groupe de spectres correspond bien à celui des érythrocytes non-infectées. Les spectres du second groupe sont conformes à la signature d'un complexe formé de l'amodiaquine et de l'hémoglobine. Ceci confirme bien la présence de l'amodiaquine à l'intérieur des globules rouges; de plus, ces résultats associés à ceux de l'ACP permettent de retrouver la variation des concentrations du médicament dans les cellules que nous avons effectuée. **La Figure 6** nous montre une décroissance progressive de l'intensité des pixels pour la CP4 à la différence des autres. En isolant cette composante particulière, nous obtenons la **Figure 7** qui montre clairement que des doses croissantes de médicaments peuvent être observées dans la cellule à partir de la quatrième composante principale des spectres. Ceci s'explique par les proportions des composantes des cellules infectées auxquelles les médicaments ont été appliqués. Le constituant le plus important attendu dans l'érythrocyte infecté, est l'hémoglobine; il doit être suivi par le parasite lui-même et/ou les déchets de sa digestion que sont l'hémozoïne. Il est donc prévu que le complexe formé par le médicament et l'hème puisse se retrouver en quatrième position en terme de proportion, ceci explique donc le fait que ce soit la quatrième composante principale qui rende compte des concentrations relatives appliquées aux cellules.

5. Conclusion

L'utilisation de l'ACP, de la CHA et de la méthode des k-moyennes, donne lieu à des résultats dont l'analyse qualitative et quantitative nous a permis de retrouver les différentes concentrations appliquées aux cellules sanguines analysées. Il a été donc possible de cartographier les distributions de concentration relative des médicaments dans la zone d'intérêt. En outre, le procédé que nous avons conçu est conceptuellement simple, et suppose que les différentes substances sont réparties, dans la zone d'intérêt, indépendamment les unes des autres.

Références

- [1] - M. PETREMAND, M. LOUYS, Ch. COLLET ET F. FLITTI, Représentation couleur d'images multi-spectrales astronomiques. *Traitement du signal*, 21 (2004), 499-516.
- [2] - François HUSSON. Analyse en composantes principales (ACP) Théorie et pratique. (17 janvier 2014)
- [3] - J. CLATWORTHY, D. BUICK, M. HANKINS, J. WEINMAN ET R. HOME. The use and reporting of Cluster Analysis in health psychology. *Journal of Health Psychology*, 10 (2005), 329-358.
- [4] - Nick Coleman. Cluster Analysis Tutorial, (2002) 1-15
- [5] - HASTIE, TREVOR, TIBAHIRANI., ROBERT., FRIEDMAN ET JEROME. Hierarchical clustering. *The Elements of Statistical learning*. New York: Springer. 2 (2009), 520-528.
- [6] - R. M. Cormack, A review of classification. *Journal of the Royal Statistical Society*, A 134 (1971):321-367.
- [7] - ERIC W. WEISSTEIN, "K-means Clustering Algorithm" from Wolfram Mathworld. (16 octobre 2014)

- [8] - J. T. ZOUEU, S. OUATTARA, A. TOURE, S. SAFI ET S. T. ZAN. Spectroscopic approach of multi-spectral imagin of plasmodium falciparum infected human erythrocytes, IEEE proceedings, ICTON Mediterranean Winter Conference., 23 (2009), 1-7
- [9] - J. ABOMA MERDASA; MIKKEL BRYDEGAARD; SUNE SVANBERG; JERÉMIE T. ZOUEU. Journal of Biomedical Optics. 18(03), 036002 (March 2013)
- [10] - M. BRYDEGAARD, GUAN Z. AND S. SVANBERG. Broad-band multi-spectral microscope for imaging transmission spectroscopy employing an array of light-emitting diodes. American Journal of Physics, 77 (2009), 104.
- [11] - L. KAUFMAN AND P. J. ROUSSEEUW. Finding Groups in Data: An Introduction to Cluster Analysis. Hoboken,NJ: John Wiley & Sons, Inc., (27 may 2008).