

# Une application de classification vidéo en mécanique des fluides diphasiques oscillatoires

Donatello Conte\*, Dominique Li\*, Yuqi Huang\*\*, Haoyi Niu\*\*

\*Laboratoire LIFAT, Université de Tours, France

\*\*Université du Zhejiang, Chine

donatello.conte@univ-tours.fr, dominique.li@univ-tours.fr

huangyuqi@zju.edu.cn, niuhaoyi@zju.edu.cn

**Résumé.** Nous présentons une application industrielle en mécanique des fluides diphasiques des techniques de *machine learning* et d'analyse de vidéos. Dans le contexte de l'identification des écoulements diphasiques oscillatoires à grande vitesse dans la galerie de refroidissement de pistons de moteurs automobiles, la complexité soulevée par l'utilisation des matériaux opaques et la continuité du mouvement rend difficile l'observation et l'analyse des écoulements du liquide de refroidissement dans la galerie. Nous proposons dans cet article une méthode visuelle pour la caractérisation des écoulements diphasiques, qui a été validée par la classification de vidéos. A notre connaissance, il s'agit du premier article à introduire l'étude visualisée des écoulements diphasiques à l'aide de techniques d'apprentissage supervisé.

## 1 Introduction

Baucoup de problèmes industriels impliquent des écoulements diphasiques, par exemple, l'écoulement gaz-liquide dans le transport pétrochimique, ou l'écoulement liquide-solide dans le traitement des eaux usées. Parmi ces écoulements diphasiques, en relation avec le transfert de chaleur et de masse, l'écoulement gaz-liquide attire particulièrement l'attention. Cependant, l'état d'écoulement gaz-liquide est effectivement compliqué en raison des bulles et des gouttelettes qui sont générées dans les interactions des deux phases. De nombreux travaux de recherche industrielle ont été consacrés à l'utilisation de capteurs pour étudier et classifier les écoulements gaz-liquide, dont la plupart portent sur l'utilisation des paramètres d'écoulements diphasiques. En effet, toutes ces études dépendent de la sensibilité et de la précision des capteurs, de sorte que les résultats soient limités.

Dans notre recherche, nous nous intéressons à l'étude des motifs d'écoulements diphasiques à l'intérieur de la galerie de refroidissement des pistons du moteur automobile. La galerie est modélisée par une chambre contenant du gaz et du liquide, activée par un moteur électrique, avec des quantités de gouttelettes de liquide qui y sont dégagées. Un tel modèle pourrait simuler le processus de dissipation de la chaleur dans les pistons de refroidissement interne d'un moteur, et il pourrait servir pour mesurer l'efficacité de refroidissement du piston et de l'ensemble du moteur. Dans les systèmes d'écoulements diphasiques gaz-liquide, l'interface entre ces deux phases change constamment, ce qui rend difficile la capture et l'analyse

instantanée des écoulements. En outre, il est facile de voir que les motifs d'écoulements sont étroitement liés à des facteurs tels que, par exemple, la taille, la forme et l'état de chauffage du conteneur, la tension superficielle du milieu, etc. Il est donc difficile de construire une équation mathématique pour décrire un modèle d'écoulement à deux phases tenant compte des variables ci-dessus.

Nous nous tournons, dans cet article, vers des techniques de *machine learning* et d'analyse de vidéos pour identifier les motifs d'écoulements diphasiques. Les écoulements sont capturés par une caméra à grande vitesse, qui encadre un modèle simplifié et transparent de galerie de refroidissement sous oscillation forcée à grande vitesse tout en simulant le mouvement du piston entraîné par un moteur. Dans notre approche, étant donnée une vidéo, les mouvements gaz-liquide de la chambre de refroidissement sont représentés par des séquences de déplacements des flux optiques calculés sur la vidéo. Ces séquences de déplacements peuvent être converties en séquences symboliques par la technique de *bag-of-words* qui est largement utilisée en vision par ordinateur et en traitement de langues naturelles. Ces séquences symboliques, représentant des motifs visuels du mouvement, sont utilisées comme caractéristiques d'écoulements diphasiques permettant d'estimer la vitesse du moteur à l'aide des principales méthodes de classification existantes en reconnaissance de formes. Notre évaluation expérimentale montre que toutes les séquences vidéo peuvent être correctement classées par rapport à la vitesse du moteur, ce qui indique en outre qu'il est possible d'apprendre des motifs d'écoulements diphasiques. Au meilleur de notre connaissance, il s'agit du premier article à introduire une telle étude des motifs visuels pour la caractérisation de l'écoulement diphasique avec des techniques d'apprentissage et d'analyse vidéo.

Le reste de l'article est organisé comme suit. Dans la section 2, nous présentons le contexte de notre travaux sur des écoulements diphasiques et le système d'acquisition. L'approche proposée est illustrée dans la section 3 et validée dans la section 4. Nous donnons des conclusions et perspectives dans la dernière section.

## 2 Contexte et installation expérimentale

L'industrie automobile a actuellement tendance à augmenter la densité de puissance des moteurs à combustion interne. En tant que composant essentiel de ces moteurs, le piston subit une charge thermique plus élevée que les autres. Il existe différentes façons d'augmenter la charge thermique, la plus efficace étant d'améliorer la capacité de refroidissement du piston, où le refroidissement par pulvérisation, le jet d'huile et la galerie de refroidissement sont les types les plus utilisés.

Comme montré dans la Fig. 1(a), la majeure partie de la chaleur (jusqu'à 70%–90%) se dissipe à travers les courroies annulaires et l'arrière de la couronne avec des structures refroidies par pulvérisation et à jet d'huile, à travers une galerie de refroidissement (60%–70% de la chaleur dissipée [Kajiwara *et al.*, 2003]). Puisque la zone de la tête de piston subit le plus de charges thermiques et est l'endroit ayant la température la plus élevée, la galerie de refroidissement est en effet le dispositif de refroidissement le plus efficace. Puisque l'état d'écoulement d'huile-gaz oscillé dans les pistons a une grande influence sur l'efficacité du refroidissement, il est important d'étudier les principes d'écoulement diphasique dans de tels dispositifs.

Dans nos expérimentations, la conception des équipements est illustrée à la Fig. 1(b), qui comprend deux parties : (1) *la console d'expérimentation* qui pourrait définir une vitesse de

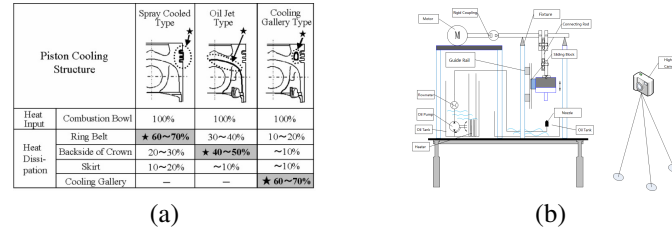


FIG. 1 – (a) Structure de refroidissement du piston et distribution de chaleur dissipée. (b) Conception des équipements expérimentaux.

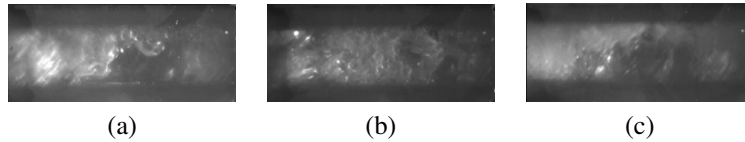


FIG. 2 – Motifs d'écoulements diphasiques à différentes vitesses du moteur : (a) 200 tr/min; (b) 400 tr/min; (c) 600 tr/min.

moteur, et (2) l'appareil de visualisation, pour lequel nous utilisons une caméra à haute vitesse pour obtenir les vidéos des motifs d'écoulements diphasiques dans le piston. Par exemple, la Fig. 2 montre des tels motifs à l'intérieur du modèle de galerie de refroidissement aux vitesses du moteur 200 tr/min, 400 tr/min et 600 tr/min. En effet, 9 séquences vidéos<sup>1</sup> de 30 fps sont enregistrés dans nos expérimentations, correspondant aux vitesses du moteur à 200 tr/min, 250 tr/min, 300 tr/min jusqu'à 600 tr/min, où chaque séquence dure environ 10 minutes.

### 3 Approche de classification proposée

Dans cette section, nous présentons notre approche de classification, qui consiste en l'extraction du flux optique (section 3.1) dans un intervalle de temps estimé (section 3.2). La vidéo est divisé en plusieurs partie (section 3.3) et les différentes sequences de flux optiques sont classifiées par des techniques de *machine learning* (section 3.4).

#### 3.1 Extraction du flux optique

Pour détecter les motifs du flux par rapport à des écoulements diphasiques, nous utilisons la technique proposée par [Shi and Tomasi, 1994] pour la détection des points caractéristiques dans une vidéos. Cette technique consiste à détecter les pixels de l'image qui représentent des coins. Certains de ces pixels sont supprimés car il ne sont pas des maximum locaux en termes de luminosité. Les coins restants sont triés, par une mesure de qualité, dans l'ordre décroissant, et les coins moins important en termes de qualité sont aussi supprimés.

1. <https://github.com/flow-patterns/cooling-gallery>

Ensuite ces points caractéristiques sont suivis dans la séquence vidéo pour calculer la valeur du flux optique en correspondance de ces points. Le flux optique est une estimation du mouvement apparent des pixels entre deux images consécutives d'une vidéo. Le flux optique reflète intuitivement le modèle de mouvement du fluide dans notre étude.

Pour le calcul du flux optique nous utilisons la méthode de Lucas-Kanade [Lucas and Kanade, 1981] qui fournit une estimation du mouvement des points caractéristiques sous la contrainte de deux hypothèses : (1) deux images consécutives sont séparé par un petit intervalle de temps  $\Delta t$  (voir section 3.2), de telle sorte que les objets ne se soient pas déplacés de manière significative ; (2) la texture des objets, en termes de niveau de gris, change lentement. La méthode considère un masque  $3 \times 3$  autour d'un point identifié  $(x, y)$ . Ensuite les 9 dérivées (le long de l'axe vertical, horizontal et le long de la dimension d'intensité lumineuse)  $(f_x, f_y, f_t)$  sont calculées, de sorte que le problème revient à la résolution de 9 équations avec deux variables inconnues qui est un problème surdéterminé, pour lequel une solution est obtenue par la méthode du moindre carré. En d'autres termes,

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \sum_i f_{x_i}^2 & \sum_i f_{x_i} f_{y_i} \\ \sum_i f_{x_i} f_{y_i} & \sum_i f_{y_i}^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum_i f_{x_i} f_{t_i} \\ -\sum_i f_{y_i} f_{t_i} \end{bmatrix},$$

où  $(x_i, y_i)$  sont les points appartenant au voisinage de  $(x, y)$ . La solution  $u$  et  $v$  de l'équation ci-dessus est le vecteur du flux optique calculé dans le point  $(x, y)$ , qu'on appelle aussi *vecteur de déplacement*.

Pour chaque point de flux optique tracé le long de la séquence vidéo, la liste temporelle de tous les vecteurs de déplacement est une *séquence de déplacement*. Nous pouvons donc représenter une vidéo comme un ensemble de séquences de déplacement.

### 3.2 Estimation de l'intervalle de temps de l'algorithme LK

Dans cette section nous discutons comment estimer l'intervalle de temps  $\Delta t$  plus approprié pour l'application de l'algorithme de Lucas-Kanade (LK) pour le calcul du flux optique. Cet intervalle sera exprimé en terme de nombre de frames  $N$  consécutives à considérer par l'algorithme LK.

Par rapport au nombre total d'images  $F$  et à la durée  $T$  de la vidéo  $T$ , l'intervalle de temps entre deux images consécutives est clairement  $\Delta f = T/F$ . La vitesse du piston de refroidissement dans nos vidéos suit le comportement des courbes sinusoïdales montrées dans la Fig. 3 [Yu *et al.*, 2019], donc nous pouvons estimer la valeur de  $N$  par  $N = n/i = (F/T)(1/i)$ , où  $n = 1/\Delta f$  et  $i$  est une constante ( $2/T \leq i \leq F/T$ ).

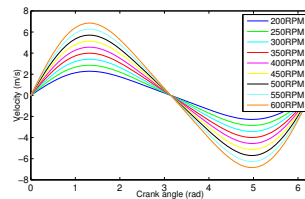


FIG. 3 – Vitesse du piston de refroidissement par rapport à la vitesse de rotation du moteur.

Dans nos expérimentations, le débit de la vidéo est d'environ 30 images par seconde, donc la valeur de l'intervalle calculé de  $N$  est  $N = (F/T)(1/i) = 30(1/i)$ , où  $i = 2$  ( $i = 6$  dans le cas de segmentation des vidéos, voir la section 3.3) est finalement sélectionné en considérant la longueur moyenne des séquences extraite. Le nombre d'images prise en considération pour l'application de l'algorithme est donc de 15 images (5 images pour les vidéos segmentées).

### 3.3 Segmentation des vidéos

Au cours de nos expérimentations, certaines vidéos présentent des occultations du liquide par le tube qui est attaché au conteneur du liquide (voir la Fig. 4) où les rectangles rouge et bleue représentent respectivement le tube et le conteneur. Pour palier à ce problème, nous segmentons les séquences vidéos en quatre parties correspondantes à quand le conteneur est en haut (*top*), quand il descend (*downward*), quand il est en bas (*bottom*) et quand il remonte (*upward*).

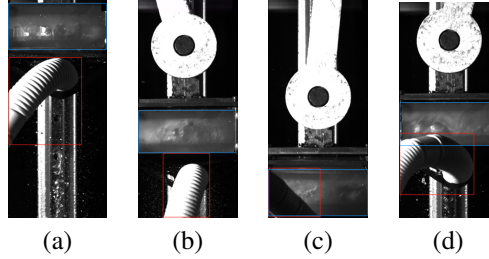


FIG. 4 – Segments de vidéos : (a) *top*; (b) *downward*; (c) *bottom*; (d) *upward*.

Soit  $\lambda_1^t$  et  $\lambda_2^t$  les instants de temps de la position physique supérieure du conteneur entre deux cycles consécutifs, la durée du mouvement par cycle est  $\sigma = \lambda_2^t - \lambda_1^t$ . Le nombre total  $k$  de cycles contenus dans une vidéo est donc  $k = F/(30\sigma)$ . De manière analogue, soit  $\lambda_m^t$ ,  $\lambda_{m+1}^t$ ,  $\lambda_m^b$ , et  $\lambda_{m+1}^b$  les instant où le conteneur est en haut et en bas entre deux cycles consécutifs. En considérant que le conteneur reste en haut et en bas pendant un temps  $\delta$ , les instants de temps par lesquelles une vidéo peut donc être divisé en 4 segments sont les suivants :

$$\begin{aligned} T_{top} &= [\lambda_m^t - \delta, \lambda_m^t + \delta] & T_{downward} &= [\lambda_m^t + \delta, \lambda_m^b - \delta] \\ T_{bottom} &= [\lambda_m^b - \delta, \lambda_m^b + \delta] & T_{upward} &= [\lambda_m^b + \delta, \lambda_{m+1}^b - \delta] \end{aligned}$$

Notre expérimentation montrent que les segments *top* et *downward* peuvent caractériser le flux mieux qu'en considérant la séquence vidéo entière.

### 3.4 Classification des séquences vidéos

Soit  $\mathcal{V}$  un ensemble de vidéos dont chacune correspond à une vitesse du moteur et  $\mathcal{C}$  un ensemble d'étiquettes de classe correspondant à des vitesses du moteur, chaque vidéo  $V_i \in \mathcal{V}$  est associée à une étiquette  $C_i \in \mathcal{C}$ . Dans notre approche, les vidéos seront représentées par un ensemble  $\mathcal{S}$  de séquences de déplacement de points caractéristiques, notons  $\mathcal{S} \models \mathcal{V}$ . Par conséquent, étant donné une vidéo  $V$ , le problème mentionné dans cet article peut être décrit

comme apprendre des caractéristiques qui permettent de décrire les écoulements diphasiques causés par le mouvement de la galerie.

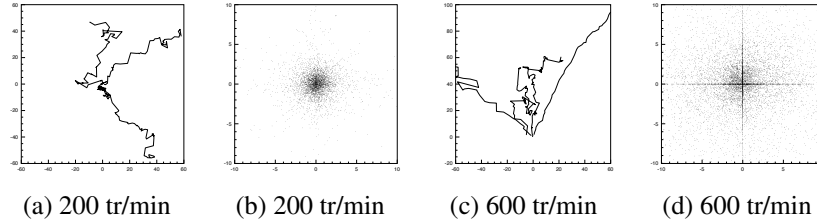


FIG. 5 – Exemple de séquences de flux optiques à différentes vitesses de moteur (a, c) et leur représentation en terme de déplacement de points caractéristiques (b, d).

Soit  $\mathcal{C}$  l'ensemble de toutes les étiquettes de classe,  $S$  un ensemble de séquences,  $h$  un classifieur tel que pour chaque séquence  $s \in S$ ,  $h(s) = C$  où  $C \in \mathcal{C}$  est une étiquette de classe. Étant donné un classifieur  $h$ , nous définissons la confiance de la classe  $C$  par rapport à un ensemble de séquences  $S$ , désigné par  $\sigma_h(C, S)$ , comme le nombre de toutes les séquences  $s \in S$  classées en  $C$ , c'est-à-dire,  $\sigma_h(C, S) = |\{s \in S \mid h(s) = C\}|/|S|$ . Une classe étiquetée  $C$  est appelée la classe dominante d'un ensemble de séquences  $S$  par rapport à un classifieur  $h$  si pour toutes les autres classes  $C' \in \mathcal{C}$  nous avons  $\sigma_h(C, S) > \sigma_h(C', S)$ ; sinon nous disons qu'il n'y a pas de classe dominante et que le groupe  $S$  n'est pas classifié par  $h$ .

Pour la phase de classification, nous appliquons le modèle *bag of words* [Csurka *et al.*, 2004], pour représenter les mouvements des points de flux optiques tracés, les distributions des vecteurs de déplacement étant assez différentes par rapport aux vitesses des moteurs (Fig. 5). Une étape dans le modèle *bag of words* est celle de la discrétisation. Ici nous utilisons le clustering  $k$ -means pour convertir les vecteurs de déplacement en attributs symboliques,  $k$  représentant les  $k$  centroïdes de la partition. Le modèle de *bag of words* les plus utilisés sont basés sur des unigrammes. Pour représenter la dynamique temporelle des séquences, nous utilisons ici des multigrammes.

## 4 Évaluation expérimentale

Dans nos évaluations nous avons testé le réseau bayésien (BN)<sup>2</sup>. Les résultats montrés dans notre évaluation expérimentale sont des valeurs moyennes obtenues par une validation croisée de 5-folds. Sur chaque ensemble nous appliquons le clustering  $k$ -means indépendamment des autres ensembles, de manière que les ensembles de test sont absolument exclus du processus d'entraînement.

La Tab. 1 donne quelques informations statistiques de base sur tous les ensembles de séquences, où nous indiquons avec  $|S|$  le nombre de séquences,  $||S||$  le nombre de symboles de déplacements distincts,  $|s|$  le nombre moyen de symboles de déplacements par séquence, et  $\{t, d, b, u\}$  les différents segments de la vidéo. Notez que les résultats expérimentaux sur la vidéo complète sans segmentation ne sont pas inclus dans cet article, car les résultats de classification sont bien pires que les résultats pour les vidéos segmentées.

2. Nous avons également testé le SVM mais les résultats étaient beaucoup plus mauvais que ceux obtenus par BN.

Video Clip	$ S _t$	$\ S\ _t$	$ s _t$	$ S _d$	$\ S\ _d$	$ s _d$	$ S _b$	$\ S\ _b$	$ s _b$	$ S _u$	$\ S\ _u$	$ s _u$
200 rpm	382	22920	60	1158	69480	60	78	4692	60	896	48436	54
250 rpm	133	6558	49	492	25950	52	148	3724	25	589	30476	51
300 rpm	311	18191	58	560	30800	55	182	11843	65	630	41774	66
350 rpm	198	12276	62	744	46128	62	187	11619	62	654	40548	62
400 rpm	134	7877	58	340	16866	49	87	6333	72	612	39154	63
450 rpm	127	8340	65	366	21152	57	78	4172	53	328	21561	65
500 rpm	264	25872	98	537	52626	98	233	22373	96	506	49278	97
550 rpm	177	10975	62	303	14055	46	107	8377	78	398	33693	84
600 rpm	178	11228	63	376	22694	60	128	9365	73	424	35612	83
TEST	312	21347	68	769	49526	64	206	14312	69	878	64154	73
TRAIN	1232	84014	68	3051	196336	64	796	54961	69	3498	255081	72

TAB. 1 – Informations statistiques sur la base de donnée, top ( $t$ ), downward ( $d$ ), bottom ( $b$ ), and upward ( $u$ ). En bas : valeurs moyennes des ensembles de test et training.

La Tab. 2 montre la performance du classificateur BN. Nous pouvons conclure que les segments *up* et *downward* peuvent mieux caractériser les modèles de flux à deux phases dans notre étude.

	$ F $	200rpm	250rpm	300rpm	350rpm	400rpm	450rpm	500rpm	550rpm	600rpm	+
$t$	180	<b>100.00+</b>	80.77+	92.31+	<b>100.00+</b>	61.54+	44.00+	97.44+	51.43+	40.00+	<b>9</b>
$d$	70	96.91+	70.10+	<b>98.97+</b>	<b>100.00+</b>	73.53+	36.99+	<b>100.00+</b>	48.33+	<b>90.67+</b>	<b>9</b>
$b$	2000	66.67+	<b>100.00+</b>	50.00+	<b>100.00+</b>	<b>94.12+</b>	20.00-	73.08+	23.81-	32.00+	7
$u$	400	73.15+	91.67+	71.30+	<b>100.00+</b>	55.56+	15.38-	97.03+	51.90+	64.29+	8

TAB. 2 – Précision de classification des vidéos avec le classificateur BN avec le modèle à unigramme. Pour chaque vitesse de moteur, +/– indiquent si la vidéo est classifié correctement/incorrectement.

Nous avons testé également les modèles avec bigrammes et multigrammes avec valeur de  $k = 80$ . Les résultats sont présentés dans la Tab. 3.

	$ F $	200rpm	250rpm	300rpm	350rpm	400rpm	450rpm	500rpm	550rpm	600rpm
$UG_t$	80	92.31+	80.77+	84.62+	92.31+	61.54+	36.00+	94.87+	60.00+	40.00+
$BG_t$	4097	97.44+	73.08+	84.62+	89.74+	30.77-	48.00+	89.74+	54.29+	37.14-
$MG_t$	632	97.44+	80.77+	89.74+	92.31+	53.85+	48.00+	100.00+	57.14+	62.86+
$UG_d$	80	96.91+	65.98+	97.94+	100.00+	63.24+	35.62+	100.00+	53.33+	92.00+
$BG_d$	4133	97.94+	69.07+	93.81+	98.97+	72.06+	30.14+	94.85+	56.67+	81.33+
$MG_d$	827	100.00+	72.16+	93.81+	100.00+	67.65+	32.88+	96.91+	56.67+	88.00+
$UG_b$	80	40.00+	57.69+	3.85-	19.23-	5.88-	0.00-	38.46+	42.86+	28.00-
$BG_b$	3980	80.00+	73.08+	50.00+	88.46+	23.53-	6.67-	76.92+	71.43+	16.00-
$MG_b$	891	66.67+	69.23+	61.54+	84.62+	29.41-	53.33+	88.46+	61.90+	28.00+
$UG_u$	80	73.15+	94.44+	67.59+	98.15+	53.70+	20.00-	96.04+	46.84+	66.67+
$BG_u$	4956	72.22+	95.37+	68.52+	99.07+	54.63+	29.23+	94.06+	54.43+	66.67+
$MG_u$	1764	75.00+	92.59+	67.59+	100.00+	55.56+	30.77+	96.04+	50.63+	69.05+

TAB. 3 – Comparaison de la classification des séquence par rapport aux modèles avec unigrammes/bigrammes/multigrammes ( $UG/BG/MG$ ) et le classificateur BN avec  $k = 80$  pour le clustering  $k$ -means.

Nos résultats montrent que les modèles basés sur les multigrammes, qui considèrent la dynamique temporelle, peuvent améliorer la performance globale de la classification des séquences.

## 5 Conclusion

Dans cet article, nous avons présenté une application d'analyse de vidéos pour la classification des écoulements à deux phases à l'intérieur d'une galerie de refroidissement des pistons des moteurs de véhicules. Notre approche et notre évaluation expérimentale montrent que l'écoulement diphasique, dans le contexte d'une oscillation forcée à grande vitesse, peut être caractérisé par des motifs visuels. Dans notre future orientation de recherche, nous sommes intéressés à utiliser des techniques d'apprentissage profond pour étudier ce problème. Nous nous intéressons également à la compréhension des motifs d'écoulements diphasiques, par exemple, pour estimer, par des techniques de régression, des vitesses de moteur différentes de celles contenues dans les vidéos d'entraînement.

## Références

- Gabriella Csurka, Christopher Dance, Lixin Fan, Jutta Willamowski, and Cédric Bray. Visual categorization with bags of keypoints. In *Workshop on statistical learning in computer vision, ECCV*, pages 1–2, 2004.
- Hidehiko Kajiwara, Yukihiro Fujioka, and Hideo Negishi. Prediction of temperatures on pistons with cooling gallery in diesel engines using cfd tool. *SAE Technical Paper 2003-01-0986*, page 9, 2003.
- Bruce D. Lucas and Takeo Kanade. An iterative image registration technique with an application to stereo vision. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence, IJCAI 1981*, pages 674–679, 1981.
- Jianbo Shi and Carlo Tomasi. Good features to track. In *Conference on Computer Vision and Pattern Recognition, CVPR 1994*, pages 593–600, 1994.
- Xiaoli Yu, Dong Yi, Yuqi Huang, Yiji Lu, and Anthony Paul Roskilly. Experimental investigation of two-phase flow and heat transfer performance in a cooling gallery under forced oscillation. *International Journal of Heat and Mass Transfer*, 132 :1306–1318, 2019.

## Summary

This paper presents an industrial application of machine learning techniques in fluid mechanics to the two-phase flow identification problem, under the context of forced high-speed oscillation with reference to the cooling gallery inside automobile engine pistons. We propose a video classification method to characterize two-phase flow patterns within a simplified cooling gallery model. Different experimental evaluations show that our approach is effective to identify two-phase flow patterns. To our knowledge, this is the first paper to introduce visualized study of two-phase flows with machine learning techniques in related engineering domains.