

# Extraction automatique de pendules dans des images de caténaire

Caroline PETITJEAN<sup>1</sup>, Laurent HEUTTE<sup>1</sup>, Vincent DELCOURT<sup>2</sup>, Régis KOUADIO<sup>1,2</sup>

<sup>1</sup>Université de Rouen, LITIS EA 4108  
BP 12, 76801 Saint-Etienne-du-Rouvray, France

<sup>2</sup>Direction de l'Innovation et de la Recherche SNCF,  
45 rue de Londres, 75379 Paris Cedex 8, France

Caroline.Petitjean@univ-rouen.fr, Laurent.Heutte@univ-rouen.fr,  
Vincent.Delcourt@sncf.fr

**Résumé** – Cet article présente un système original permettant d'extraire automatiquement des images de caténaire, des éléments caractéristiques de celles-ci, appelés pendules, dans le but futur d'en détecter les défauts. Le système d'extraction des pendules est basé sur une approche descendante, dans laquelle la connaissance *a priori* concernant la position des pendules et leur type à l'intérieur des portées caténaire est intégrée en amont du processus de reconnaissance, permettant ainsi d'en fiabiliser les performances. Les expériences menées sur une base conséquente d'images acquises en condition réelle (temps variable, présence de tunnel) montrent des résultats très prometteurs, pour cette application de vision par ordinateur sans équivalent : le rappel et la précision sont tous deux de 99.06% pour la base considérée, et le taux de bonne classification moyen sur les deux types de pendules existants est de 99.57%.

**Abstract** – The aim of this paper is to present an automatic, image-based system for catenary maintenance, a novel application of machine vision which has no equivalent today. This study focuses on the detection of droppers in catenary staves. The system takes benefit from the fact that dropper location inside catenary staves follows mounting rules, an information that is integrated into a top-down approach in order to speed up and make reliable the extraction of droppers. Experiments conducted on a large database of real images allow to obtain very satisfying recall and precision rates (both 99.06%) and correct classification rate (99.57%). These results enable for continuing investigations for dropper fault detection.

## 1 Introduction

La caténaire est la ligne aérienne de traction électrique des trains. Elle est composée d'un fil de contact, permettant de faire circuler l'électricité, et d'un fil porteur soutenant le fil de contact. Entre ces deux câbles sont positionnés trois éléments verticaux, réalisant principalement le soutien des câbles : le bras de rappel (BR), le pendule simple (PS) et le pendule connexion (PC) (Figure 1). La portion de caténaire délimitée par deux bras de rappel est appelée portée.

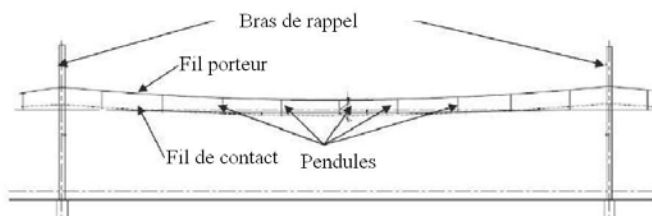


FIGURE 1 – Vue perpendiculaire de la caténaire

La combinaison des pendules à l'intérieur d'une portée suit des règles dites de montage, qui en spécifient, par exemple, le nombre et le positionnement. Il existe ainsi différents types de portées, de différentes longueurs, avec un nombre variable de PS et de PC. La Figure 2 montre l'exemple d'une portée de longueur 63 m comportant 7 pendules simples.

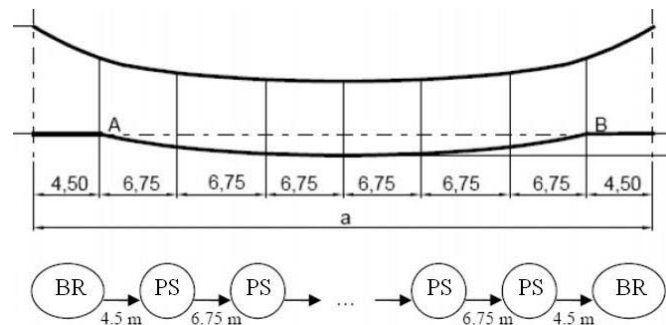


FIGURE 2 – Règles de montage pour une portée de 63 m composée de 7 PS

La maintenance préventive de la caténaire consiste à vérifier la présence et l'intégrité de chaque composant. Actuellement réalisée par inspection visuelle, cette tâche longue et fastidieuse sera bientôt automatisée. La Direction de la Recherche et de l'Innovation de la SNCF a ainsi mis en place un système d'acquisition dédié, embarqué sur une rame TGV roulant à 320 km/h [3], permettant d'obtenir des images représentant des segments adjacents de portées caténaires (Figures 3 et 4). Cette vitesse élevée impose une fréquence d'acquisition très rapide (53kHz). Grâce à un asservissement de la fréquence d'obturation à la vitesse de la rame, le système permet l'obtention d'images nettes. Par l'utilisation de filtres, il est également ro-

buste aux conditions climatiques.



FIGURE 3 – Système d'acquisition embarqué sur une rame TGV (acquisition perpendiculaire à la caténaire)

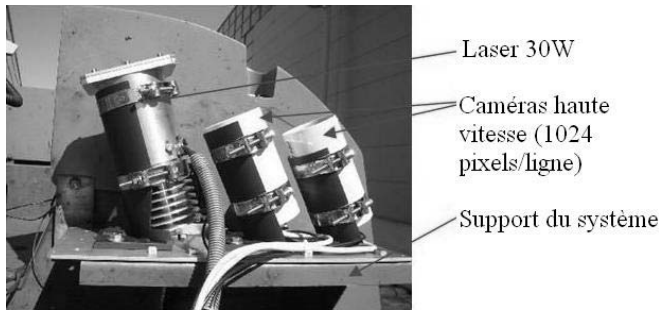


FIGURE 4 – Système d'acquisition composé du laser et des caméras

Les images obtenues sont présentées en Figure 5. Chaque image a pour taille 1024x768 pixels, ce qui représente une sous-partie de la caténaire de 1.38m de long, la résolution horizontale étant de 1.8 mm par pixel à vitesse constante. Une portée caténaire se compose ainsi d'une quarantaine d'images environ. Un système de détection des bras de rappel intégré au système d'acquisition permet d'avoir accès aux ensembles d'images composant une portée.

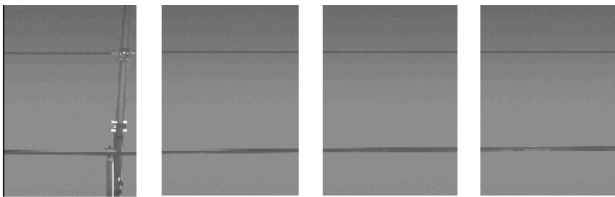


FIGURE 5 – Images de la caténaire fournies par le système d'acquisition

L'objectif de cette étude est de développer un système automatique de traitement des images obtenues, permettant l'identification des pendules constituant la caténaire. Le but ultime

sera de détecter les défauts sur ces pendules (typiquement, des pendules cassés ou une mauvaise connection électrique...). Dans ce contexte, l'utilisation des règles de montage, informant sur la localisation et le type de pendule en fonction du type de portée, peut être particulièrement utile pour notre système. A ce jour, il n'existe pas de travaux équivalents concernant l'analyse automatique de la caténaire, par d'autres compagnies ferroviaires dans le monde. A notre connaissance, seul [6] propose un système de vision embarqué, dans lequel l'identification des éléments de la caténaire est manuelle.

Dans la suite de l'article, nous présentons l'approche adoptée dans la Section 2, et les résultats sont fournis en Section 3. La Section 4 conclut et donne les perspectives de ce travail.

## 2 Approche descendante pour la localisation automatique de pendules

L'analyse de scène classique peut être abordée par deux approches, qui diffèrent selon que la connaissance a priori est incorporée tôt (approche descendante) ou tard (approche ascendante) dans le processus de reconnaissance [1]. Dans le cadre de notre application, la connaissance *a priori* repose sur les différents types de portées existants et les règles de montage, qui indiquent, pour chaque type de portée, le nombre et l'ordonnement des PS et des PC, ainsi que les distances les séparant. Ces informations peuvent contraindre le modèle et rendre ainsi la détection de pendules plus fiable. Cependant, la résolution horizontale des images variant durant les phases d'accélération et de décélération, les informations de distance inter-pendules ne peuvent être utilisées. Nous avons donc conçu une première approche ascendante, dans laquelle les composants verticaux et horizontaux étaient segmentés sans connaissance *a priori*, mais dont la séquence était vérifiée par un alignement sur tous les modèles de portées existants [4]. Même si les résultats étaient bons en terme de précision et de rappel, cette approche était trop coûteuse car l'analyse était conduite sur une image entière d'une portée caténaire. Une approche descendante, même si elle est limitée aux images de résolution fixe, permet de guider le processus de reconnaissance grâce à la connaissance *a priori* et ainsi d'accélérer les temps de calcul. Les résultats obtenus avec cette nouvelle approche descendante seront comparés à ceux obtenus avec l'approche ascendante dans la Section 3.

L'approche proposée dans cet article est résumée à la Figure 6. L'entrée du système est un ensemble d'images représentant une portée caténaire. La concaténation d'une quarantaine d'images fournit une image de 1024 par environ 30000 pixels. Quatre étapes composent le système :

1. une localisation grossière des pendules utilisant la connaissance *a priori* (modèle géométrique),
2. une détection des points d'ancrage des pendules aux câbles, permettant de segmenter les pendules (discrimination PS/PC),
3. la classification des hypothèses de pendules,
4. l'alignement sur les modèles de portée caténaire pour vérification.

Ces étapes sont détaillées dans la suite.

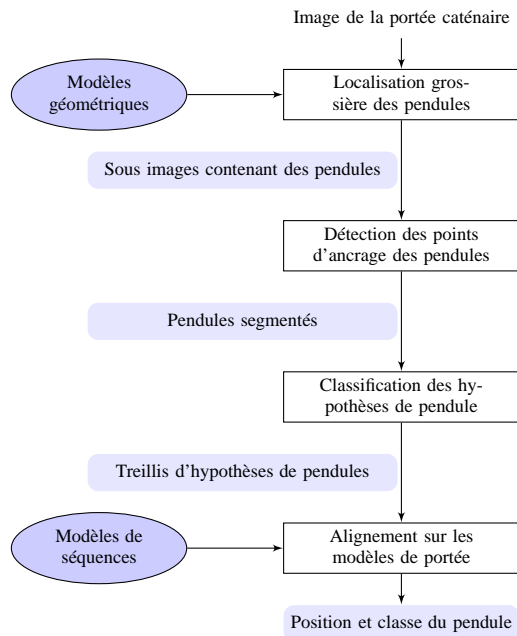


FIGURE 6 – Présentation générale du système proposé

## 2.1 Localisation grossière des pendules

La connaissance *a priori* est ici utilisée pour localiser les pendules et éviter ainsi une analyse aveugle de l'image. L'information utilisée est un modèle géométrique, qui contient les différents types de portées, leur longueur et la répartition des PS et des PC en leur sein. Il existe 74 types de portées, du modèle simple à 4 pendules, au plus complexe fait d'une combinaison de 15 PS et PC. La distance entre deux bras de rappel, i.e. la longueur de la portée, est tout d'abord calculée [3, 4]. Cette longueur est une caractéristique discriminante qui permet de filtrer les modèles et donc va permettre de prédire grossièrement les différentes positions possibles des pendules, à partir de tous les modèles de portée possédant cette longueur. En sortie sont fournies les sous-images qui contiennent des pendules, dont la largeur est choisie assez grande (1500 pixels) pour prendre en compte les variations possibles de la résolution horizontale, à cause de légères accélérations et décélérations pendant l'acquisition.

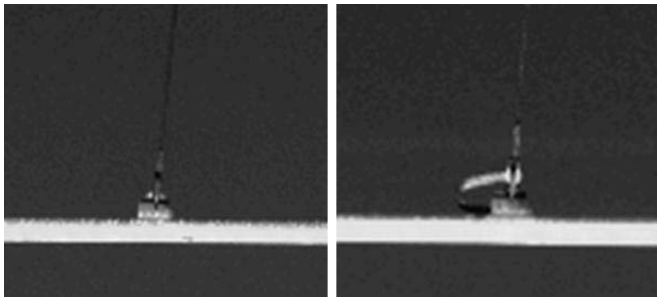


FIGURE 7 – Différence entre PS et PC.

## 2.2 Détection des points d'ancrage des pendules

Il s'agit maintenant de détecter, dans les sous-images censées contenir des pendules, les points d'ancrage, i.e. les inter-

sections entre les pendules et les câbles horizontaux. Notons que ces images, contrairement aux apparences, ne sont pas si faciles à segmenter. En effet, le fond est bruité et non homogène, et le contraste fond/pendule est faible. Les pendules sont également très fins, leur largeur étant égale à seulement 2 à 3 pixels. D'autre part, il y a dans l'image des éléments plus clairs ou plus foncés que le fond (Figure 8). Soulignons que pour améliorer la visualisation du contenu des images, le contraste de toutes les images présentées dans ce papier a été manuellement modifié.

Ainsi, pour simplifier la détection des pendules, l'image est binarisée, en utilisant les opérateurs *TopHat* et *BotHat*, qui permettent de segmenter des éléments qui sont respectivement plus clairs et plus foncés que le fond [5].

$$\begin{aligned} TopHat &= Image - Ouverture(Image) \\ BotHat &= Fermeture(Image) - Image \end{aligned} \quad (1)$$

Les deux éléments structurants utilisés sont construits en fonction des éléments à détecter. L'un est une ligne horizontale de 10 pixels de long, et l'autre est une ligne verticale de même longueur. Un exemple de résultat de binarisation est présenté en Figure 9.

La segmentation est ensuite effectuée par projection de l'image binarisée, verticalement pour détecter les éléments verticaux (PS, PC, bras de rappel), et horizontalement pour détecter les fils porteur et de contact. L'intersection entre éléments verticaux et horizontaux permet de localiser les deux points d'ancrage dans l'image binarisée.

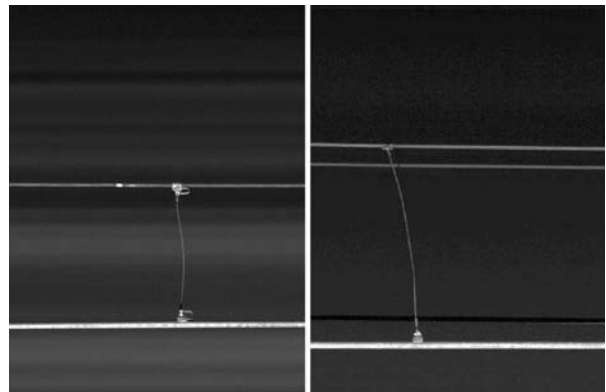


FIGURE 8 – Objets foncés et clairs dans la caténaire

## 2.3 Classification des pendules

La discrimination entre PS et PC s'effectue au niveau des points d'ancrage. En effet, le PC présente des connexions supplémentaires aux câbles (Figure 7). La densité de pixels (nombre de pixels blancs divisé par le nombre total de pixels) est calculée en ces points d'ancrage, dans une fenêtre de taille 10 x 10 pixels. Un simple vecteur contenant les 2 valeurs de densité est assez discriminant pour classer les pendules. La classification est ensuite effectuée par un réseau de neurones de type perceptron multi-couches (MLP, Multi-Layer Perceptron) 2-5-2, choisi pour sa vitesse en phase de décision et ses bonnes propriétés de généralisation [2]. Il fournit en sortie des probabili-

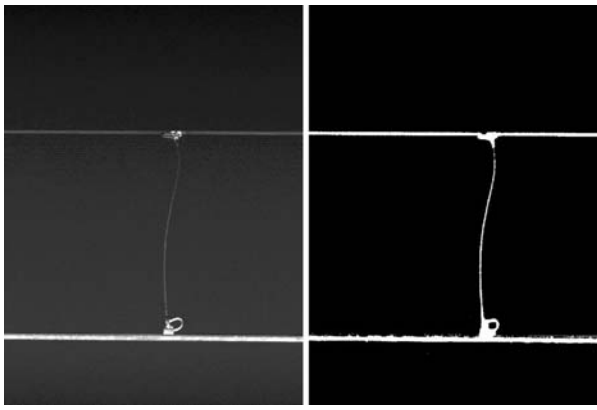


FIGURE 9 – (a) Image originale, (b) Image binarisée

tés *a posteriori* pour l'alignement des modèles de séquences. L'étape suivante consiste à corriger les possibles erreurs de classification en alignant les observations sur l'image avec divers modèles de séquences de pendules, i.e. modèles de portée.

## 2.4 Alignement sur les modèles de portées caténaire

Dans une approche descendante, la reconnaissance est principalement une étape de vérification, qui cherche à aligner des modèles sur les séquences d'observation, en sortie de l'étape de classification. L'alignement consiste à chercher le chemin optimal dans le treillis d'hypothèses. Plus précisément, on dispose, pour chaque sous-image de pendule, des probabilités que le pendule soit un PS ou un PC, fournies par le classifieur MLP. D'autre part, on dispose de tous les modèles de portée correspondant à la longueur de la portée. Par exemple, une portée de 63 m peut être composée de 7 pendules simples, ou d'une succession de 2 PS, 3 PC, et 2 PS (Figure 10). Dans une approche descendante, les probabilités de chaque pendule peuvent donc être multipliées directement à travers les chemins du treillis correspondant à ces deux modèles. Le modèle gagnant est celui du chemin de plus forte probabilité.

	Obs1	Obs2	Obs3	Obs4	Obs5	Obs6	Obs7	
p(PS)	0.76	0.83	0.87	0.97	0.038	0.78	0.98	p(M1)=0.0154
p(PC)	0.24	0.04	0.14	0.05	0.83	0.14	0.11	p(M2)=0.0002

FIGURE 10 – Alignement sur le résultat du MLP des deux modèles de la portée caténaire composée de 7 pendules, l'un avec 7 pendules simples (M1), l'autre avec 4 pendules simples et 3 pendules connexion (M2).

## 3 Résultats

Les expériences ont été menées sur une base de 689 images réelles de portées caténaire, représentant 5161 pendules (PS et PC), étiquetés manuellement par un rectangle englobant pour constituer la vérité terrain. Deux tiers des éléments sont utilisés pour l'apprentissage et le dernier tiers pour le test, comme

indiqué dans le Tableau 1.

TABLE 1 – Nombre d'images

	Base d'apprentissage	Base de test	Total
Portées	474	215	689
PS	2752	1262	4014
PC	803	344	1147
Total pendules	3555	1606	5161

La première étape de localisation grossière des pendules permettant d'extraire des sous-images atteint un taux de segmentation de 100%. Le filtrage des modèles par la longueur calculée entre deux bras de rappel est donc optimal. Le taux de bonne reconnaissance des pendules de l'étape de classification est donné dans le Tableau 2. Si le taux de bonne classification est bon pour les PS avant alignement (MLP uniquement), il nécessite en revanche une amélioration pour les PC. Cela pourrait provenir de la binarisation des sous-images, qui augmentent la densité de pixels autour des points d'ancrage. L'alignement sur les modèles de portées permet d'atteindre de meilleurs résultats, notamment en permettant de retrouver quasiment 4% des pendules connexion grâce à l'alignement des modèles.

TABLE 2 – Taux de bonne classification

	Avant alignement	Après alignement
PS	99.36%	99.68%
PC	95.93%	99.13%
Moyenne sur les pendules	98.62%	99.56%

Le rappel et la précision, dont les définitions sont rappelées ci-dessous, sont deux mesures permettant d'estimer la qualité des résultats obtenus. Le rappel et la précision du système sont tous deux égaux à 99.06%. Ces excellents résultats permettent une poursuite du travail centrée sur la détection des défauts.

$$\text{Rappel} = \frac{\text{Nombre de pendules correctement reconnus}}{\text{Nombre de pendules à reconnaître}}$$

$$\text{Précision} = \frac{\text{Nombre de pendules correctement reconnus}}{\text{Nombre de pendules proposés par le système}}$$

TABLE 3 – Performance globale de l'approche descendante

Rappel	99.06%
Précision	99.06%
Mesure F1	99.06%

Des tests effectués avec l'approche ascendante, décrite en Section 2 et implémentée dans [4], sur la même base d'images que celle décrite précédemment, ont permis d'obtenir les taux de rappel et de précision donnés dans le Tableau 4. Ces résultats

restent satisfaisants, quoique moins bons que ceux obtenus par la nouvelle approche. La nécessité de traiter toute une image de caténaire simultanément (de taille environ 1024 x 30000 pixels) rend également le temps de calcul très élevé.

TABLE 4 – Performance globale de l’approche ascendante

Rappel	95.7%
Précision	98.1%
Mesure F1	96.9%

## 4 Conclusion et perspectives

Nous avons présenté une approche descendante pour l’extraction automatique de pendules dans des images de caténaire et avons montré que la connaissance *a priori* peut être utilisée pour effectuer une détection fiable des pendules. Les résultats expérimentaux obtenus sur une base d’images conséquente montrent l’intérêt de l’approche, qui est une application originale en vision par ordinateur n’ayant aucun équivalent aujourd’hui.

Si la méthode développée est performante sur des images de caténaire dite simple (avec un seul fil de contact et un fil porteur), les modèles utilisés devront être adaptés aux caténaires doubles, voire triples. Il existe en effet des portions de caténaires ayant plusieurs paires fil de contact/fil porteur, pour lesquelles des modèles spécifiques doivent être mis en œuvre (Figure 11).

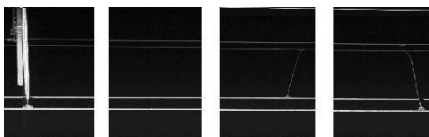


FIGURE 11 – Images d’une caténaire double

Les perspectives de ce travail concernent également la détection de défauts, en particulier la détection temps réel de l’usure du fil de contact, de la présence d’objets non identifiés sur la caténaire, des pendules cassés, d’une mauvaise connexion électrique sur les pendules.

## Références

- [1] J. Battle, A. Casals, J. Freixenet and J. Martí. *A review on strategies for recognizing natural in colour images of outdoor scenes*. Image and Vision Computing, vol. 18, 515–530, 2000.
- [2] C.M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.
- [3] R. Kouadio, V. Delcourt, L. Heutte and C. Petitjean. *Video based catenary for preventive maintenance on Iris 320*. Proc. WCRP 2008, Seoul, Korea, 2008.
- [4] F. Montreuil, R. Kouadio, C. Petitjean, L. Heutte and V. Delcourt. *Automatic extraction of information for catenary scene analysis*. Proc. EUSIPCO 2008, Lausanne, Switzerland, 2008.
- [5] C. Lacoste, X. Descombes, J. Zerubia and N. Baghdadi. *Unsupervised line network extraction from remotely sensed images by polyline process*. Proc. EUSIPCO 2004, Vienna, Austria, 2004.
- [6] U. Richter and R. Schneider. *Automatische optische Inspektion von Oberleitungen*. Elektrische Bahnen, vol. 99, 94–100, 2007.