



HAL
open science

ORBox: conception d'une surface intelligente embarquée pour la reconnaissance de petits objets

Frédéric Rayar, P Robillard

► To cite this version:

Frédéric Rayar, P Robillard. ORBox: conception d'une surface intelligente embarquée pour la reconnaissance de petits objets. 29ème conférence francophone sur l'Interaction Homme-Machine, AFIHM, Aug 2017, Poitiers, France. 8 p., 10.1145/3132129.3132147 . hal-01577928

HAL Id: hal-01577928

<https://hal.science/hal-01577928>

Submitted on 28 Aug 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ORBox: conception d'une surface intelligente embarquée pour la reconnaissance de petits objets

F. Rayar
Laboratoire d'Informatique
Université François-Rabelais
Tours, France
frederic.rayar@univ-tours.fr

P. Robillard
Polytech Tours
Université François-Rabelais
Tours, France
pierre.robillard@etu.univ-tours.fr

ABSTRACT

Dans le cadre de l'amélioration du confort de vie des personnes âgées ou ayant une déficience visuelle, nous présentons l'ORBox, un boîtier embarquant une surface intelligente pour la reconnaissance de petits objets. Ergonomique et facile d'utilisation pour les personnes susmentionnées, il peut être utilisé dans des scénarios quotidiens de reconnaissance de pièces de monnaie ou encore de médicaments (cachets, pilules). La faisabilité d'un tel boîtier est montrée et nous présentons des résultats prometteurs, avec un taux de reconnaissance moyen de 92,58% sur des pièces de monnaie.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

IHM'17, August 29-September 1, 2017, Poitiers, France

© 2017 Copyright is held by the owner/author(s).

ACM ISBN 978-1-4503-5109-6/17/08. <https://doi.org/10.1145/3132129.3132147>

CCS CONCEPTS

• **Human-centered computing** → **HCI design and evaluation methods** • **Computing methodologies** → **Object recognition**

KEYWORDS

Surfaces intelligentes, Reconnaissance d'objets, Personnes âgées, Déficience visuelle

ACM Reference format:

F. Rayar, and P. Robillard. 2017. ORBox: conception d'une surface intelligente embarquée pour la reconnaissance de petits objets. In *Proceedings of 29eme Conference Francophone sur L'Interaction Homme-Machine , Poitiers- Futuroscope, FR, Août 2017 (IHM 2017)*, 8 pages.

DOI: 10.1145/3132129.3132147

1 INTRODUCTION

Ces dix dernières années, la communauté d'interaction homme-machine a porté une attention particulière aux surfaces interactives (aka. smart tabletops). Avec ce type d'interfaces, les utilisateurs ont la possibilité d'interagir avec un système simplement en déposant des objets sur une surface intelligente.

Nous pouvons distinguer deux catégories d'objets : la première englobe l'ensemble des appareils électroniques (smartphones, tablettes) et la seconde correspond à des objets plus « classiques », tels des livres, des clés ou encore des aliments. Pour la première catégorie, le paradigme des objets connectés apporte une solution simple, en exploitant des protocoles de communication tel que le RFID, le NFC, le Bluetooth ou encore le Wi-Fi. Pour la seconde catégorie, on distingue 3 approches : (i) l'ajout de puces RFID [6] ou NFC sur chaque objet, (ii) l'ajout de marqueurs spécifiques collés sur les objets tels que des codes QR [5] et (iii) l'utilisation de caméras et de techniques de reconnaissance de formes pour reconnaître directement les objets [2].

Bien que l'utilisation de puces ou de marqueurs visuels facilite la reconnaissance des objets, on imagine aisément que le fait de devoir les ajouter à tous les objets devant être reconnus peut être un frein à la démocratisation des systèmes exploitant cette solution. Aussi, il apparaît pertinent de s'intéresser à la troisième approche mentionnée ci-dessus. Pour ce faire, il est possible d'utiliser une caméra externe, associée aux récentes avancées en matière de reconnaissance d'objets dans les images [10], qui permettent de résoudre le problème dans une moindre mesure. Néanmoins, le dispositif devient complexe, avec un positionnement fixe de la caméra externe ou encore l'utilisation de projecteurs. Encore une fois, ces contraintes peuvent être un frein, voire un problème d'acceptation par le grand public.

Ainsi, dans nos travaux, nous étudions la reconnaissance d'objets déposés sur une surface intelligente, non connectés, sans marqueurs attachés et sans l'utilisation de caméra externe. Plus particulièrement, nous nous intéressons aux « petits » objets, tels des pièces de monnaie ou des médicaments (pilules, cachets). Ce travail est motivé par l'utilisation d'un tel dispositif par des personnes âgées ou encore des personnes ayant une déficience visuelle. Ces derniers peuvent en effet rencontrer des difficultés à reconnaître correctement des petits objets. Cette étude entre donc dans le cadre de la « conception universelle » [13], telle que définie par la « Convention relative aux droits des personnes handicapées », adoptée le 13 décembre 2006 par l'Assemblée générale des Nations Unies. Il est à noter que bien que la reconnaissance des pièces de monnaie soit un scénario sans danger ou conséquence grave, celui de la reconnaissance de médicaments est plus critique car il engage la sécurité des personnes.

Nous présentons donc l'ORBox (Object Recognition Box), un boîtier constitué, entre autre, d'une surface intelligente, qui permet la reconnaissance de petits objets déposés sur cette surface. Plus précisément, ce boîtier possède les propriétés suivantes :

1. c'est un système embarqué et mobile ;
2. il est à la fois ergonomique et facile à utiliser par des personnes âgées ou avec une déficience visuelle ;
3. la partie logicielle a été réalisée de manière à facilement considérer de nouveaux objets à reconnaître.

2 TRAVAUX EXISTANTS

Du point de vue matériel, on peut citer l'existence des tables *reactTable* [3] et *Lumisight* [4] qui utilisent une caméra placée sous la surface d'une table, et la table *Samsung SUR40* qui utilise la technologie infrarouge pour « voir » les objets déposés sur celle-ci. Dans l'étude exploratoire de *Rayar et al.* [8], les auteurs exploitent la *SUR40*. Des objets courants (carte de visite, carte étudiante, pièces de monnaie) ont été utilisés dans les expérimentations, pour des taux de reconnaissance allant de 80% à 100%. Néanmoins, les auteurs indiquent que les résultats obtenus sont en partie dus à l'hétérogénéité des formes des objets sélectionnés. De plus, ils indiquent que leur système n'est pas robuste face à différentes pièces de monnaie.

3 CONCEPTION DE L'ORBOX

3.1. Partie matérielle

L'ORBox est un boîtier léger de 13cm x 14cm pour 11cm de hauteur. La Figure 1 présente un prototype fonctionnel de l'ORBox. Les quatre faces latérales sont faites de matière solide tandis que la face du dessus est constituée d'une plaque de plexiglas transparente. Cette plaque dispose d'un traitement anti-rayure, permettant ainsi une utilisation répétée, i.e. le dépôt de petits objets tels des pièces de monnaie ou des clés, sans risquer de détériorer la plaque.



Figure 1: Prototype de l'ORBox

A l'intérieur du boîtier est déposée une Raspberry Pi qui contrôle le boîtier. Celle-ci est surmontée d'une caméra orientée vers le haut qui permet de voir le dessous des objets posés sur le plexiglas transparent. La caméra utilisée est équipée d'un objectif fish-eye de 170° d'angle de vue afin de voir toute la surface vitrée. Du fait de l'opacité des faces latérales, l'intérieur est sombre et les photos prises par la caméra ne sont pas exploitables dans un objectif de reconnaissance d'objets. Afin de pallier à ce problème, un ensemble de diodes électroluminescentes (DEL) a été installé afin d'illuminer le boîtier. Néanmoins, ce rajout de DEL a un effet secondaire sur les images prises par la caméra : en effet, elles réfléchissent sur le plexiglas transparent et apportent du bruit visuel dans l'image. C'est pourquoi, nous avons rajouté une plaque diffusante permettant d'obtenir un éclairage constant et homogène lors de la prise des images. Ce boîtier ayant pour vocation d'être utilisé par des personnes ayant une déficience visuelle, il embarque un haut-parleur afin que l'ORBox communique le résultat de la reconnaissance à l'utilisateur.

3.2. Partie logicielle

Le scénario de l'utilisation de l'ORBox est tel que décrit ci-dessous :

1. l'utilisateur pose ses objets sur la plaque de plexiglas;
2. il demande la reconnaissance avec un appui court sur un bouton poussoir;
3. les DEL de l'ORBox s'allument et une première image est capturée;
4. les DEL s'éteignent et une deuxième image est capturée;
5. des opérations de traitements d'image et de reconnaissance d'objets sont alors effectuées par l'ORBox;
6. l'ORBox fait une annonce vocale du résultat obtenu.

Nous décrivons dans la suite un peu plus en détail l'étape 5. La première opération est de redimensionner l'image brute de 1920 x 1080 pixels en une image carrée de 1080 pixels de côté centrée sur la plaque, afin de réduire la région d'intérêt de l'image. Puis, une opération de correction des déformations géométriques causées par l'objectif est effectuée. Cette opération est réalisée en estimant les coefficients de distorsion grâce à une étape de calibration de la caméra.

Ces opérations sont effectuées sur les 2 images acquises par l'ORBox : la segmentation des objets est ainsi facilitée. En effet, en effectuant la différence absolue des deux images, on arrive à isoler les objets d'intérêt. Des opérations classiques du traitement d'images telles le seuillage, l'érosion et la dilatation sont ensuite utilisées pour récupérer les rectangles englobants les objets d'intérêt. Une opération de filtrage basée sur les dimensions des boîtes peut être effectuée pour filtrer certains faux objets. Une fois les régions d'intérêt obtenues, nous utilisons une méthodologie classique de l'apprentissage supervisé pour reconnaître les objets. La Figure 2 illustre la méthodologie employée. Dans un premier temps, un certain nombre d'images des objets sont récupérées, décrites par des caractéristiques et annotées par leur classe. Fort de cette base d'apprentissage, lorsqu'un nouvel objet doit être reconnu, on le décrit à son tour avec les caractéristiques, puis utilise un classifieur qui va prédire la classe du nouvel objet. Afin de réaliser cette opération d'apprentissage, on utilise une interface d'administration dédiée. Il s'agit d'une interface web qui pourra être utilisée soit

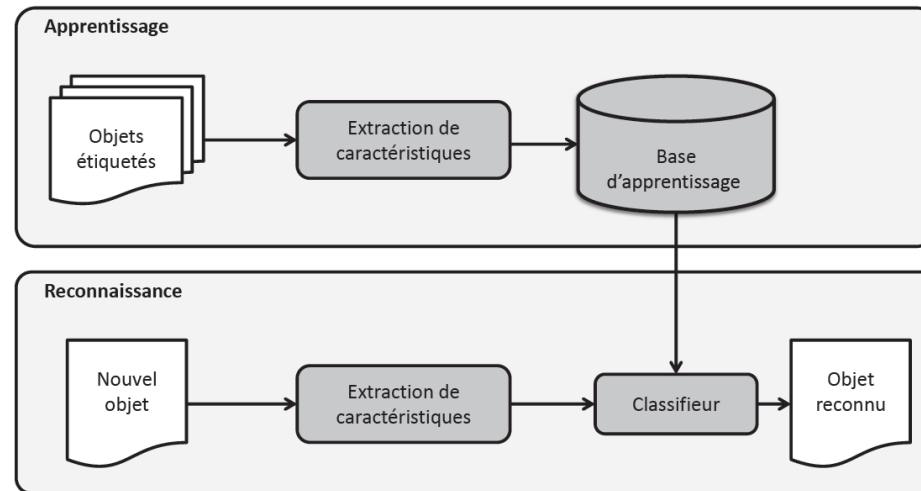


Figure 2. Méthodologie classique de l'apprentissage supervisé

depuis un smartphone, soit depuis un ordinateur. La communication avec la Raspberry de l'ORBox est réalisée via le Wi-Fi. Les différentes actions réalisables sont implémentées sous la forme de web services REST. Nous ne faisons pas apparaître dans cet article les détails techniques liés à cette partie.

4 EXPÉRIMENTATIONS

Nous présentons ici les résultats obtenus sur le scénario de la reconnaissance de pièces de monnaie. Le Table 1 résume l'ensemble des données d'expérimentation considérées. Ces images ont été acquises directement à l'aide de l'ORBox.

Des combinaisons de deux types descripteurs ont été considérées dans notre étude :

1. Les descripteurs de couleurs : un histogramme normalisé du canal teinte du modèle Teinte Saturation Valeur a été utilisé. En effet, nous avons remarqué deux profils d'histogrammes mono-mode : les pièces rouges ont un mode centré autour de 15, tandis que les pièces jaunes et celles de un et deux euros ont un mode centré autour de 20.
2. Les descripteurs de formes : le périmètre du contour, l'aire, la largeur et la hauteur du rectangle englobant.

Deux classifieurs ont été utilisés, nommément la méthode des plus proches voisins [11] et les machines à vecteurs de support (SVM) [12]. Chaque méthode de classification a ensuite été testée sur l'ensemble des huit classes de pièces, de paire avec la validation croisée en « k-fold cross validation ». Cette méthode consiste à diviser les échantillons originaux en k sous-ensembles, puis

Table 1: Données d'expérimentation considérée

<i>Objet</i>	<i>Nombre d'images</i>
1 centime	146
2 centimes	185
5 centimes	100
10 centimes	91
20 centimes	131
50 centimes	95
1 euro	107
2 euros	66

à sélectionner l'un des k sous-ensembles d'échantillons comme ensemble de validation et les $(k - 1)$ autres sous-ensembles constitueront l'ensemble d'apprentissage. On calcule l'erreur de prédiction pour l'ensemble de validation, soit le nombre d'échantillons dont la classe a été incorrectement prédite sur le nombre total d'échantillon de l'ensemble de validation. Puis on répète k fois l'opération en sélectionnant un autre ensemble de validation parmi les $(k-1)$ sous-ensembles qui n'ont pas encore été utilisés. Ainsi chaque échantillon a été utilisé exactement une fois comme ensemble de validation. L'erreur de prédiction totale pour cette méthode est la moyenne des k taux de prédiction. La Table 2 présente les différents taux d'erreur obtenus avec les différentes combinaisons de descripteurs et de classifieurs testés.

Table 2: Résultats obtenus pour 921 échantillons répartis en 8 classes

<i>Descripteurs utilisés</i>	<i>Classifieur</i>	<i>Erreur moyenne</i>
Aire et 60 barres d'histogramme	1-NN 150-fold	20,17%
Aire et 60 barres d'histogramme	9-NN 150-fold	22,14%
Aire, périmètre, hauteur et largeur du rectangle, 15/42 barres d'histogramme	SVM RBF 20-fold	12,61%
Aire, périmètre, hauteur et largeur du rectangle, 15/42 barres d'histogramme	SVM Sigmoid 20-fold	78,23%
Aire, périmètre, hauteur et largeur du rectangle, 15/42 barres d'histogramme	SVM CHI2 20-fold	7,42%

5 DISCUSSION

Comme on peut le voir dans la Table 2, il nous a été possible d'obtenir un taux d'erreur moyen de 7,42% (i.e., un taux de reconnaissance moyen de 92,58%) en utilisant une combinaison des descripteurs de forme et de couleur, associée à un SVM. Ces résultats sont satisfaisants et prometteurs pour le scénario de la reconnaissance de pièces de monnaie. Néanmoins, pour le scénario de reconnaissance de médicaments, il nous faut tendre vers le 0% de taux d'erreur, la sécurité des personnes étant mise en jeu. Pour ce faire, nous pourrions nous inspirer de travaux existants dans la littérature. En effet, on peut noter que des travaux ont été menés pour reconnaître ce type d'objets, mais en utilisant des photos « propres » de pièces [9] ou les médicaments [7]. La première étude indique obtenir un taux de reconnaissance de 97% en utilisant d'autres descripteurs, notamment les transformées de Fourier. Néanmoins, il est à noter que les auteurs utilisent les images des deux faces d'une pièce pour la reconnaître. Le second travail quant à lui, utilise des descripteurs avancés (SIFT, MLBP) pour obtenir un taux de reconnaissance de 73,17%. Nous avons nous-mêmes mené des expérimentations dans ce sens (non reporté dans le

présent article), en utilisant des descripteurs binaires. Cependant, nous avons obtenu des taux d'erreur équivalents d'environ 30% en moyenne. De

plus, le temps de calcul et de comparaison de ces descripteurs est plus élevé, et peut induire un temps de réponse plus long de la part de notre système. Une réflexion sur le choix des descripteurs et du classifieur pour obtenir un compromis entre un taux de reconnaissance élevé et un temps de réponse acceptable de l'ORBox sera donc envisagée dans les perspectives de nos travaux.

On peut se demander, et à juste titre, pourquoi concevoir une surface intelligente pour reconnaître les petits objets étudiés au lieu d'exploiter directement les travaux existants. Néanmoins, il faut noter que lesdits travaux sous-entendent l'utilisation d'un smartphone pour l'acquisition. Ce premier point peut déjà être un premier frein à l'utilisation d'un tel système par des personnes âgées. De plus, les dispositifs mis en jeu sont plus complexes que notre proposition. On peut citer par exemple Cunha et al. [1] qui utilisent un dispositif « complexe » pour utiliser HelpmePills, leur application mobile de reconnaissance de médicaments pour les personnes âgées. Ils proposent d'utiliser plusieurs éléments: (i) une surface plane (une table par exemple), (ii) un smartphone, (iii) un support pour ce smartphone (livres, tasse), (iv) un marqueur à détecter et enfin (v) la pilule ou le cachet déposé sur le marqueur.

6 CONCLUSIONS ET PERSPECTIVES

Dans cet article, nous avons donc montré avec l'ORBox, la faisabilité de la reconnaissance de petits objets sur une surface intelligente. Celle-ci est embarquée dans un boîtier ergonomique et facile à utiliser pour des personnes âgées ou avec une déficience visuelle. L'ORBox a été présentée lors de la Journée Portes Ouvertes au sein de notre établissement en mars 2017, et le retour du public a été positif, validant l'intérêt et l'utilisabilité d'un tel boîtier par le grand public.

Les travaux futurs porteront principalement sur l'amélioration du taux de reconnaissance de notre système, tout en maintenant un temps de réponse acceptable. Nous nous attaquerons aussi au scénario de reconnaissance de médicaments, en œuvrant à tendre vers le 0% de taux d'erreur. Enfin, nous souhaitons prêter l'ORBox à une personne âgée ou avec une déficience visuelle afin d'évaluer la pertinence de notre boîtier dans un cas d'utilisation réel et in-situ, et avoir un retour de la part du public ciblé.

REMERCIEMENTS

Les auteurs tiennent à remercier A. Augay, A. Fournier, V. Delys, D. Lecomte et P. Robillard pour leurs travaux qui à mené à la réalisation d'un premier prototype. Les auteurs tiennent aussi à remercier Polytech'Tours (Département Informatique Industrielle) pour avoir financé l'achat des différents composants des prototypes de l'ORBox.

RÉFÉRENCES

- [1] António Cunha, Telmo Adão, and Paula Trigueiros. 2014. HelpmePills : A Mobile Pill Recognition Tool for Elderly Persons. *Procedia Technology* 16 (2014), 1523-1532.
- [2] Shahram Izadi, Ankur Agarwal, Antonio Criminisi, John Winn, Andrew Blake, and Andrew Fitzgibbon. 2007. CSlate: A Multi-Touch and Object Recognition System for Remote Collaboration using Horizontal Surfaces. In *Proceedings of the Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems (Tabletop 2007)*. IEEE, 3–10.
- [3] Sergi Jordà, Martin Kaltenbrunner, Günter Geiger, and Ross Bencina. 2005. The reacTable. In *Proceedings of the 2005 International Computer Music Conference, ICMC 2005, Barcelona, Spain, September 4-10, 2005*. Michigan Publishing.
- [4] Y. Kakehi, M. Iida, T. Naemura, Y. Shirai, M. Matsushita, and T. Ohguro. 2005. Lumisight Table : an interactive view-dependent tabletop display. *Computer Graphics and Applications, IEEE* 25, 1 (2005), 48–53.
- [5] Martin Kaltenbrunner and Ross Bencina. 2007. reac- TIVision : A Computer-vision Framework for Table based Tangible Interaction. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction (TEI '07)*. ACM, 69–74.
- [6] Sébastien Kubicki, Sophie Lepreux, Yoann Lebrun, Philippe Dos Santos, Christophe Kolski, and Jean Caelen. 2009. New Human-Computer Interactions Using Tangible Objects : Application on a Digital Tabletop with RFID echnology. In *Human-Computer Interaction. Ambient, Ubiquitous and Intelligent Interaction, 13th International Conference, HCI International 2009, San Diego, CA, USA, July 19-24, 2009, Proceedings, Part III*. Springer Berlin Heidelberg, 446–455.
- [7] Young-Beom Lee, Unsang Park, Anil K. Jain, and Seong-Whan Lee. 2012. Pill-ID: Matching and retrieval of drug pill images. *Pattern Recognition Letters* 33, 7 (2012), 904 – 910.
- [8] Frédéric Rayar and Armand Renaudeau. 2015. Marker-free Object Recognition on Tabletops: An Exploratory Study. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces (ITS '15)*. ACM, 373–378.
- [9] Marco Reisert, Olaf Ronneberger, and Hans Burkhardt. 2007. A Fast and Reliable Coin Recognition System. *ACM*, 415–424.
- [10] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej arpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. 2015. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)* 115, 3 (2015), 211–252.
- [11] Gregory Shakhnarovich, Trevor Darrell, and Piotr Indyk. 2006. *Nearest-Neighbor Methods in Learning and Vision : Theory and Practice (Neural Information Processing)*. The MIT Press.
- [12] Ingo Steinwart and Andreas Christmann. 2008. *Support Vector Machines*. Springer Publishing Company, Incorporated.
- [13] Nation unies. 2006. Convention relative aux droits des personnes handicapées. 2006. <http://www.un.org/french/disabilities/default.asp?id=1413>.