



Appel à candidature - contrat doctoral

Titre : Classification automatique d'images photographiques pour le suivi spatio-temporel des chantiers de restauration

Cadre : Fondation des Sciences du Patrimoine

Abstract

Au cours de la dernière décennie, un grand nombre de projets de documentation numérique ont démontré le potentiel de la modélisation à partir d'images (photomodélisation, photogrammétrie, ...) d'objets patrimoniaux dans le contexte de la documentation, de la conservation et de la restauration. L'inscription de ces méthodes émergentes dans le suivi quotidien des activités d'un chantier de restauration du patrimoine (contexte dans lequel des centaines de photographies par jour peuvent être acquises par des multiples acteurs et en fonction de plusieurs besoins d'observation et analyse) fait émerger des nouvelles questions à l'intersection de la gestion des masses de données, de leur enrichissement sémantique ainsi que de leur classification automatique à de fins de recherche par similarité de contenu. Cette thèse a l'ambition d'introduire une approche de rupture pour le traitement massif de grandes collections de photographies, distribuées dans l'espace et dans le temps. Ce projet s'inscrit au sein du chantier scientifique du CNRS et du Ministère de la Culture pour la restauration de Notre-Dame de Paris et, en particulier, dans le cadre d'un groupe de travail en charge de la construction d'un "écosystème numérique" capable d'évoluer au fur et à mesure que les études avancent, en centralisant progressivement les ressources collectées, produites, analysées et interprétées par les scientifiques et les professionnels impliqués dans le chantier de restauration. Par le biais d'interactions avec les groupes de travail impliqués dans ce chantier scientifique (bois/charpente, pierre, vitraux, métaux, structure, acoustique, émotions patrimoniales, etc), on s'intéressera au problème de l'indexation automatique de masses de photographies en combinant plusieurs méthodes d'enrichissement sémantique 2D/3D complémentaires.

Mots-clés : heritage science, computer vision, content-based retrieval, multimodal data analytics, data semantics.

Etat de l'art

Au cours de la dernière décennie, l'émergence d'un grand nombre de projets de documentation numérique a démontré le potentiel de la numérisation des artefacts patrimoniaux à différentes échelles. En particulier, l'utilisation de technologies de numérisation 3D (et de l'imagerie numérique au sens large) dans le cadre de programmes de documentation sur le patrimoine culturel a permis l'apparition d'une nouvelle génération de supports graphiques, utiles à des objectifs multiples: analyse archéologique, suivi des phénomènes de détérioration, diffusion et représentation des connaissances historiques, etc. La numérisation 3D (jeux de méthodes s'appuyant sur des solutions d'enregistrement du réel) est aujourd'hui considérée comme un champ de recherche technologique basé sur l'utilisation de capteurs [Remondino et al. 2010]. Dans ce cadre, grâce aux progrès de la photogrammétrie et de la vision par ordinateur, les dix dernières années ont été caractérisées par une impressionnante croissance des approches de reconstruction 3D basées sur l'image [Vu et al., 2009 ; Wenzel et al., 2012 ; Rupnik et al., 2017].

L'inscription de ces nouvelles méthodes de documentation numérique dans le suivi quotidien des activités d'un chantier de restauration du patrimoine (contexte dans lequel des centaines de photographies par jour peuvent être acquises par des multiples acteurs et en fonction de plusieurs besoins d'observation et analyse) fait émerger des nouvelles questions à l'intersection entre la gestion des masses de données, leur enrichissement sémantique ainsi que leur classification automatique à des fins de recherche par similarité de contenu. En effet, toute observation et analyse d'un matériau, d'une forme ou d'un espace dans le champ de la restauration du patrimoine implique fréquemment l'identification de métriques de similarité à mobiliser à des fins de comparaison, indexation, interprétation et classification des formes comme des phénomènes à étudier. La confrontation de ce besoin quotidien de mémorisation, analyse et catégorisation, avec les avancées les plus récentes des sciences du numérique appliquées à la documentation du patrimoine, permet d'identifier trois enjeux complémentaires.

Tout d'abord, la mesure de la similarité entre divers 'objets numériques', qui est depuis plusieurs années une zone de recherche active, depuis l'introduction des techniques d'analyse d'image par la vision par ordinateur jusqu'aux approches plus récentes d'analyse de modèles 3D par la géométrie algorithmique. Ces méthodes fournissent aujourd'hui de bons résultats pour les applications d'indexation et reconnaissance d'objets multimédia [Baltrušaitis et al., 2019] [Ma et al, 2021], et leur application au domaine du patrimoine culturel [Biasotti et al., 2015].

Ensuite, la gestion de contenus numériques hétérogènes, s'appuyant principalement sur des structures formelles (thésaurii, ontologies, etc.) capables de décrire de façon univoque des éléments d'ordre conceptuel implicites et explicites (ainsi que leurs interrelations) [Doerr, 2002] mobilisés dans la documentation des artefacts patrimoniaux, y compris dans le domaine spécifique de la conservation-restauration [Bannour et al., 2018].

Enfin, les méthodes permettant d'enrichir des représentations des formes et des espaces par des attributs descriptifs : des méthodes pour lier des tags sémantiques aux représentations 3D [Havemann et al., 2008], jusqu'aux approches pour organiser des jeux de données hétérogènes autour d'une description morphologique [Manferdini et al., 2008]. De plus, l'analyse conjointe des données spatiales et temporelles est devenue d'une importance particulière pour la l'étude des transformations ou évolution des objets [Krauß and Tian, 2020], pour la distribution temporelle des événements catégorisés, ainsi que pour la répartition des données spatio-temporelles [Belussi et al., 2018].

Plusieurs travaux du laboratoire MAP (CNRS/MC) se sont concentrés sur l'intersection de ces trois enjeux en introduisant des approches pour la structuration sémantique de jeux de représentations multidimensionnelles pour la documentation et la conservation du patrimoine [De Luca et al., 2011, Manuel et al., 2014, Messaoudi et al., 2017].

Objectifs

Le degré de diffusion actuel des appareils photographiques numériques personnels, uni à l'acquisition croissante de compétences en photogrammétrie au sein de la communauté des scientifiques et des professionnels du patrimoine, permet d'envisager aujourd'hui la construction collaborative (et participative) de corpus d'images et de modèles numériques capables de documenter (et aussi cartographier) au quotidien les activités d'un chantier de restauration.

Cette thèse a l'ambition d'introduire une approche méthodologique de rupture pour le traitement massif de grandes collections de photographies via l'intégration de méthodes récentes de vision par ordinateur, d'indexation d'images et de structuration sémantique de données hétérogènes. En particulier, on abordera les questions de la distribution spatiale et temporelle des photographies de façon conjointe à celle de l'indexation par le contenu. Tout d'abord en exploitant la relation projective entre les images et les scènes tridimensionnelles reconstruites à partir de ces images [Manuel A. 2016], comme voie de propagation de jeux d'attributs sémantiques recueillis par des méthodes d'étiquetage

et d'annotation. Ensuite, en évaluant le potentiel des méthodes d'appariement et d'indexation monomodale [Gominski et al., 2019][Ma et al., 2021] et multimodale [Piasco et al., 2018 ; Piasco et al., 2021] d'images par apprentissage dans des scénarii d'analyse des transformations complexes à l'échelle de l'architecture monumentale. Par le croisement de ces méthodes on cherchera à identifier des pistes pour l'enrichissement sémantique automatique d'images spatialisée et distribuées dans le temps, mais aussi des nuages de points résultant des calculs de corrélation photogrammétrique, notamment au travers d'approches de classification [Poux, 2019] susceptibles de prendre en compte plusieurs jeux de descripteurs (morphologiques, visuels, temporels, ...).

Sur la dimension description sémantique et indexation de données hétérogènes, nous explorerons en particulier les approches de type graph embedding [Trsedya et al. 2019] pour introduire des mesures de similarité modulables dans les graphes sémantiques RDF hétérogènes. Ces graphes combinent les propriétés et relations de niveau conceptuel des objets d'étude, incluant les caractéristiques en lien avec la restauration-conservation [Bannour et al. 2018], avec les relations spatiotemporelles et visuelles extraites avec les méthodes mentionnées ci-dessus.

En ce qui concerne le prototypage informatique, ce travail de thèse pourra s'appuyer sur l'intégration et/ou l'extension de plusieurs briques logicielles (développées ces dernières années par les trois laboratoires impliqués dans le suivi scientifique) concernant l'appariement d'images et la génération de nuages de points 3D par photogrammétrie, l'annotation 2D/3D basée sur des attributs descriptifs et de niveaux de lecture organisés autour de structures terminologiques, la visualisation interactive de représentations multi-dimensionnelles au sein d'interfaces web.

Cadre d'expérimentation

Tout en étant orienté vers l'introduction de méthodes reproductibles et d'outils utilisables dans le domaine de la documentation, conservation et restauration du patrimoine au sens large, ce projet de thèse s'inscrit au sein du chantier scientifique du CNRS et du Ministère de la culture pour la restauration de Notre-Dame de Paris et, en particulier, dans le cadre du groupe de travail sur les "données numériques" (www.notre-dame.science)

Dans le cadre de ce projet, plusieurs dizaines de milliers d'images photographiques, nuages de points et modèles 3D documentant la cathédrale dans plusieurs états temporels (avant, après incendie et pendant la restauration) sont à disposition du chantier scientifique et seront exploitables dans le cadre de cette thèse.

Profil du candidat

S'agissant d'un sujet de thèse très interdisciplinaire, des candidatures sont attendues de plusieurs secteurs disciplinaires relevant de Sciences et Technologies de l'Information et de la Communication avec, si possible, des expériences dans les applications numériques à la documentation du patrimoine.

Equipe d'encadrement :

Dan Vodislav, ETIS, CY Cergy Paris Université: <http://depinfo.cyu.fr/~vodislav>

Livio De Luca, MAP, CNRS/Ministère de la Culture, <http://www.map.cnrs.fr/ldl>

Valerie Gouet-Brunet, LASTIG, IGN-ENSG, Univ. Gustave Eiffel : <https://www.umr-lastig.fr/vgouet/>

Ecole doctorale :

EM2PSI, CY Cergy Paris Université

Lieu de travail

Cette thèse bénéficie d'un encadrement réparti sur trois équipes d'accueil distribuées entre la région Parisienne (ETIS, LASTIG) et Marseille (MAP). Le lieu de travail sera établi en fonction du profil du candidat.

Rémunération : 35k euros brut/an

Pour candidater :

Pour candidater, envoyer un document PDF (un seul fichier) incluant:

- + Curriculum Vitae détaillé
- + lettre de motivation
- + Diplôme (donnant accès à une inscription en thèse de doctorat) et notes du master (M2)
- + lettre(s) de recommandation

à l'équipe d'encadrement :

dan.vodislav@cyu.fr

valerie.gouet@ign.fr

livio.deluca@map.cnrs.fr

Date limite de candidature : 15 avril 2021

Date de début de la thèse: à définir entre le printemps et l'automne 2021

Références bibliographiques

- [Bannour et al., 2018] Bannour I., Marinica C., Bouiller L., Pillay R., Darrieumerlou C. et al. CRMcr - a CIDOC-CRM extension for supporting semantic interoperability in the conservation and restoration domain, Digital Heritage 2018, Oct 2018, San Francisco, United States.
- [Baltrušaitis et al., 2019] T. Baltrušaitis, C. Ahuja and L. Morency, "Multimodal Machine Learning: A Survey and Taxonomy," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 41, no. 2, pp. 423-443, 1 Feb. 2019.
- [Belussi et al., 2018] Belussi A., Billen R., Hallot P., Migliorini S. Computing Techniques for Spatio-Temporal Data in Archaeology and Cultural Heritage—Introduction. In: Fogliaroni P., Ballatore A., Clementini E. (eds) Proceedings of Workshops and Posters at the 13th International Conference on Spatial Information Theory (COSIT 2017). COSIT 2017. Lecture Notes in Geoinformation and Cartography. Springer, Cham. https://doi.org/10.1007/978-3-319-63946-8_47
- [Biasotti et al., 2015] Biasotti S., Cerri A., Falcidieno B., Spagnuolo M. 3D Artifacts Similarity Based on the Concurrent Evaluation of Heterogeneous Properties. Journal of Computing and Cultural Heritage, 8-4. ACM, 2015.
- [De Luca et al., 2011] De Luca L., Busarayyat C., Stefani C., Veron P., Florenzano M. A semantic-based platform for the digital analysis of the architectural heritage. Computers & Graphics. Volume 35, Issue 2, April 2011, Elsevier. Pages 227-241.
- [Doerr, 2002] Doerr Martin. The CIDOC CRM—an ontological approach to semantic interoperability of metadata. AI Magazine, Special Issue on Ontologies, November 2002.
- [Gominski et al., 2019] D. Gominski, M. Poreba V. Gouet-Brunet, L. Chen, Challenging deep image descriptors for retrieval in heterogeneous iconographic collections, SUMAC'19 Workshop @ ACM Multimedia 2019, Oct 2019, Nice, France.
- [Havemann et al., 2008] Havemann S, Settgast V, Berndt R, Eide Ø, Fellner D. The Arrigo showcase reloaded— towards a sustainable link between 3D and semantics. Proceedings of VAST, 2008.
- [Ma et al, 2021] Jiayi Ma, Xingyu Jiang, Aoxiang Fan, Junjun Jiang and Junchi Yan, Image matching from handcrafted to deep features: a survey, International Journal of Computer Vision, volume 129, pages 23-79, 2021.
- [Krauß and Tian, 2020] Thomas Krauß and Jiaojiao Tian, Automatic Change Detection from High-Resolution Satellite Imagery, Remote Sensing for Archaeology and Cultural Landscapes, pp 47-58, 2020.
- [Manferdini et al., 2008] Manferdini A., Remondino F., Baldissini S., Gaiani M., Benedetti B. 3D modeling and semantic classification of archaeological finds for management and visualization in 3D archaeological databases. Proceedings of the 14th international conference on virtual systems and multiMedia (VSMM), Cyprus, 2008.
- [Manuel A., 2016] Adeline Manuel. Annotation sémantique 2D/3D d'images spatialisées pour la documentation et l'analyse d'objets patrimoniaux. Thèse de Doctorat, Arts et Métiers ParisTech, 2016.

- [Manuel et al., 2014] Manuel A., Stefani C., De Luca L., Véron P. A hybrid approach for the semantic annotation of spatially oriented images, *IJHDE (International Journal of Heritage in the Digital Era)*, Volume 3 Number 2, pp 305- 320. Multi Science Publishing, 2014.
- [Messaoudi et al., 2017]Messaoudi T., Véron P., Halin G., De Luca L. An ontological model for the reality-based 3D annotation of heritage building conservation state. *Journal of Cultural Heritage*, Volume 29, Elsevier 2018. <https://doi.org/10.1016/j.culher.2017.05.017>
- [Piasco et al., 2018] N. Piasco, D. Sidibé, C. Demonceaux, V. Gouet-Brunet, A survey on Visual-Based Localization: On the benefit of heterogeneous data, *Pattern Recognition*, Elsevier, 2018, 74, pp.90 - 109.
- [Piasco et al., 2021] N. Piasco, D. Sidibé, V. Gouet-Brunet, C. Demonceaux, Improving Image Description with Auxiliary Modality for Visual Localization in Challenging Conditions, *International Journal of Computer Vision*, Springer Verlag, 129 (1), pp.185-202, 2021.
- [Poux et al., 2019] POUX, Florent, NEUVILLE, Romain, HALLOT, Pierre et BILLEN, Roland. Smart point cloud: Definition and remaining challenges. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Copernicus Publications, 2016, Vol. 4, no W1, p. 119–127
- [Remondino and al., 2010] Remondino, F., Rizzi, A. Reality-based 3D documentation of natural and cultural heritage sites—techniques, problems, and examples. *Appl Geomat* 2, 85–100 (2010).
- [Rupnik et al., 2017] Rupnik E., Daakir M., Pierrot-Deseilligny M. MicMac—a free, open-source solution for photogrammetry. *Open Geospatial Data*, 1 Software and Standards, 2017.
- [Tangelder et al., 2008] Tangelder J-W., Veltkamp R. A survey of content based 3D shape retrieval methods. *Multimedia Tools Applications*, 39-3. 2008
- [Trsedya et al., 2019] B. D. Trsedya, J. Qi, and R. Zhang. Entity alignment between knowledge graphs using attribute embeddings, *AAAI* 2019
- [Vu et al., 2009] Vu, H-H. and Keriven, R. and Labatut, P. and Pons, J.-P. Towards high-resolution large-scale multi-view stereo. *Conference on Computer Vision and Pattern Recognition (CVPR)*, Miami, Jun 2009.
- [Wenzel et al., 2012] Wenzel, K., Abdel-Wahab, M., Cefalu, A., and Fritsch, D. High-Resolution Surface Reconstruction from Imagery for Close Range Cultural Heritage Applications. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXIX-B5, 2012