

Intégration et localisation automatique de nouvelles images au sein de collections d'images structurées par contenu visuel

Emile BLETTERY, MAP- UPR 2002 CNRS / LaSTIG - IGN-Université Gustave Eiffel

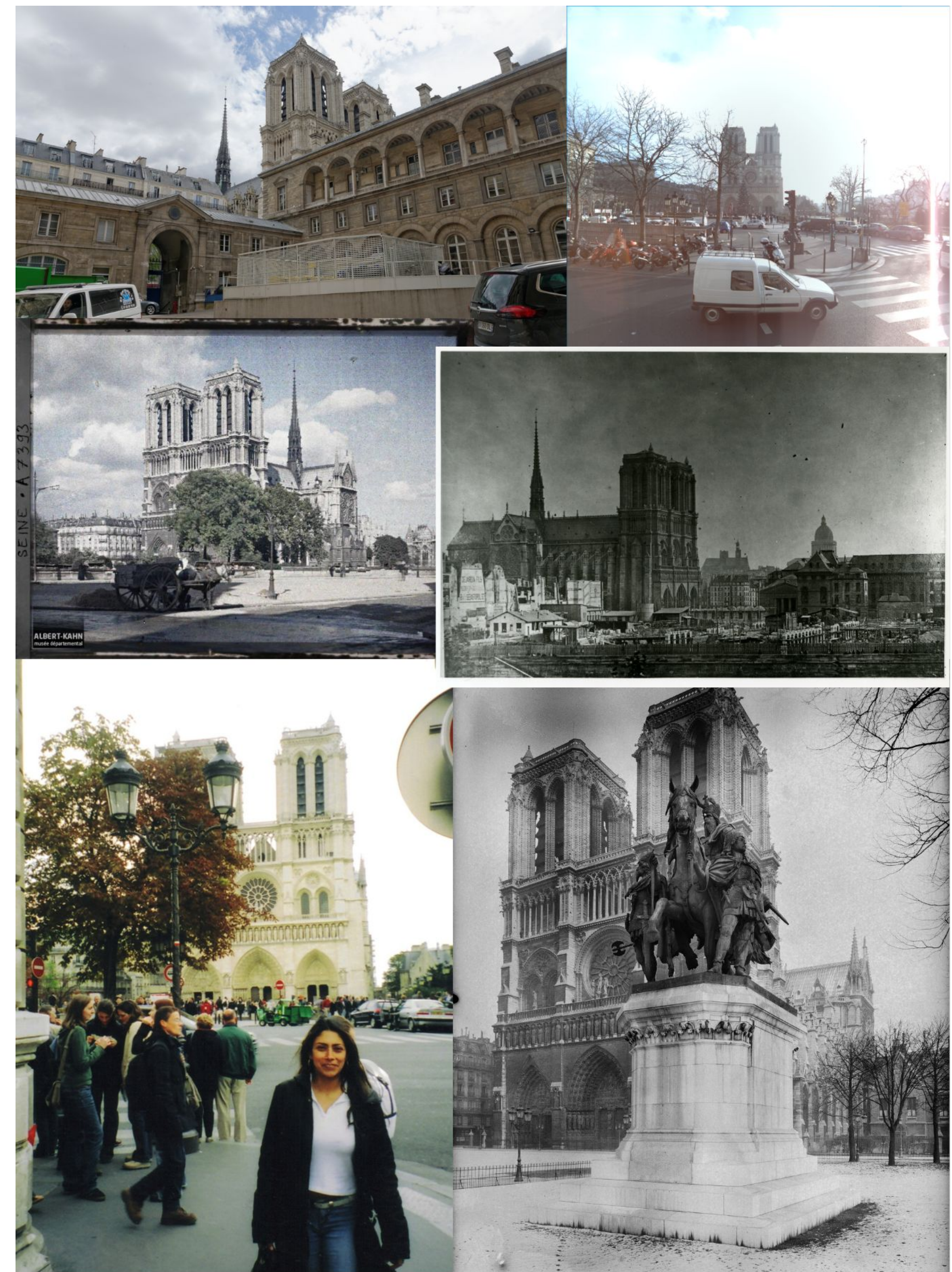


Une cathédrale de données numériques et connaissances pluridisciplinaires
Journées d'études du groupe de travail "données numériques" du chantier scientifique Notre-Dame de Paris
19 – 21 juin 2024, Campus CNRS Joseph Aiguier, Marseille

Parcours et objectifs dans le projet

- Thèse en **Janvier 2024, entre la Ville de Paris et le LaSTIG** :
 - Structuration des fonds iconographiques patrimoniaux : de l'interconnexion automatique à une validation visuelle semi-automatique.
- Fonds iconographiques patrimoniaux représentant **Paris, à l'échelle de la rue, au 20e siècle** : **8 collections différentes**.
- Trois contributions majeures :
 - **meilleure combinaison d'approches** pour lier automatiquement les contenus,
 - **nouvelles méthodes de re-ranking** pour pallier les difficultés liées à la spécificité des contenus,
 - **structuration par graphe dans une plateforme de visualisation** pour améliorer la structuration de manière semi-automatique.

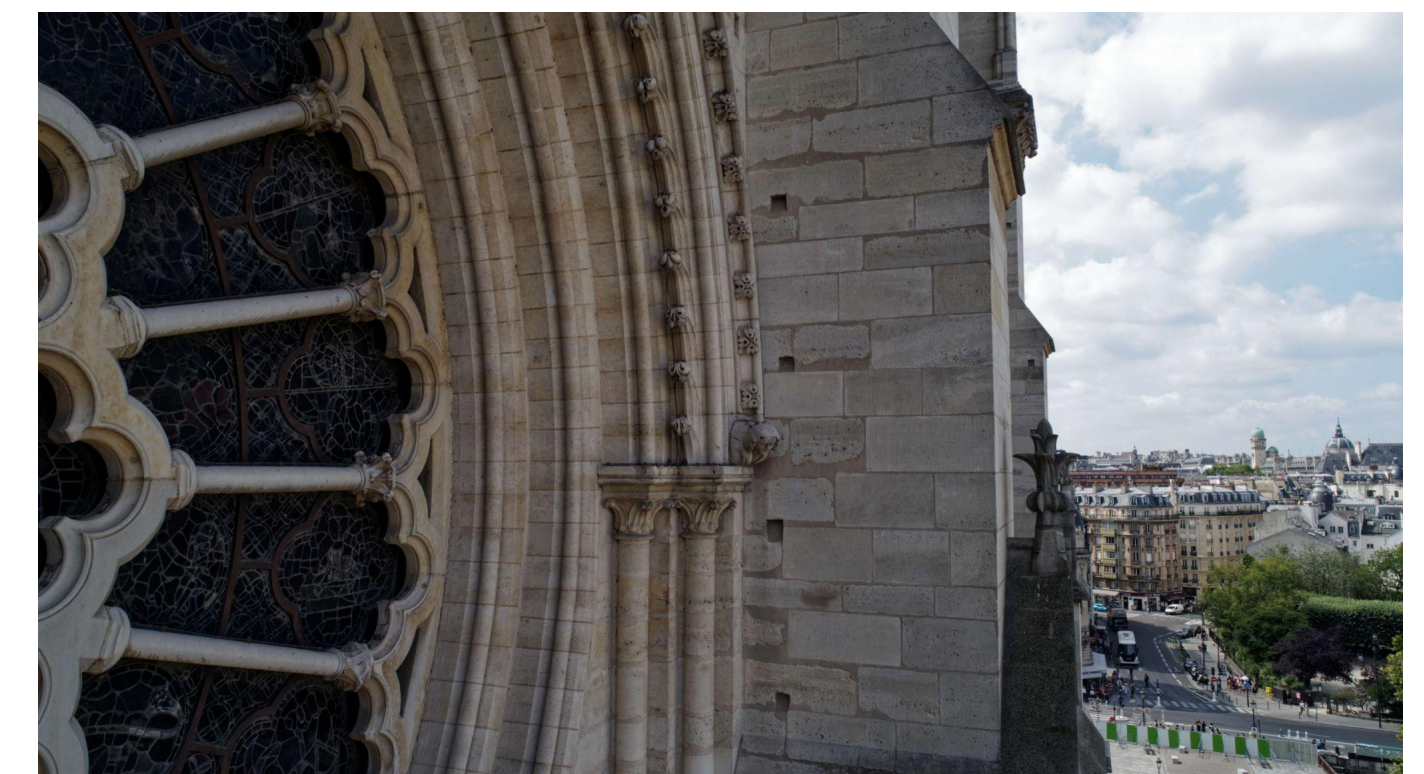
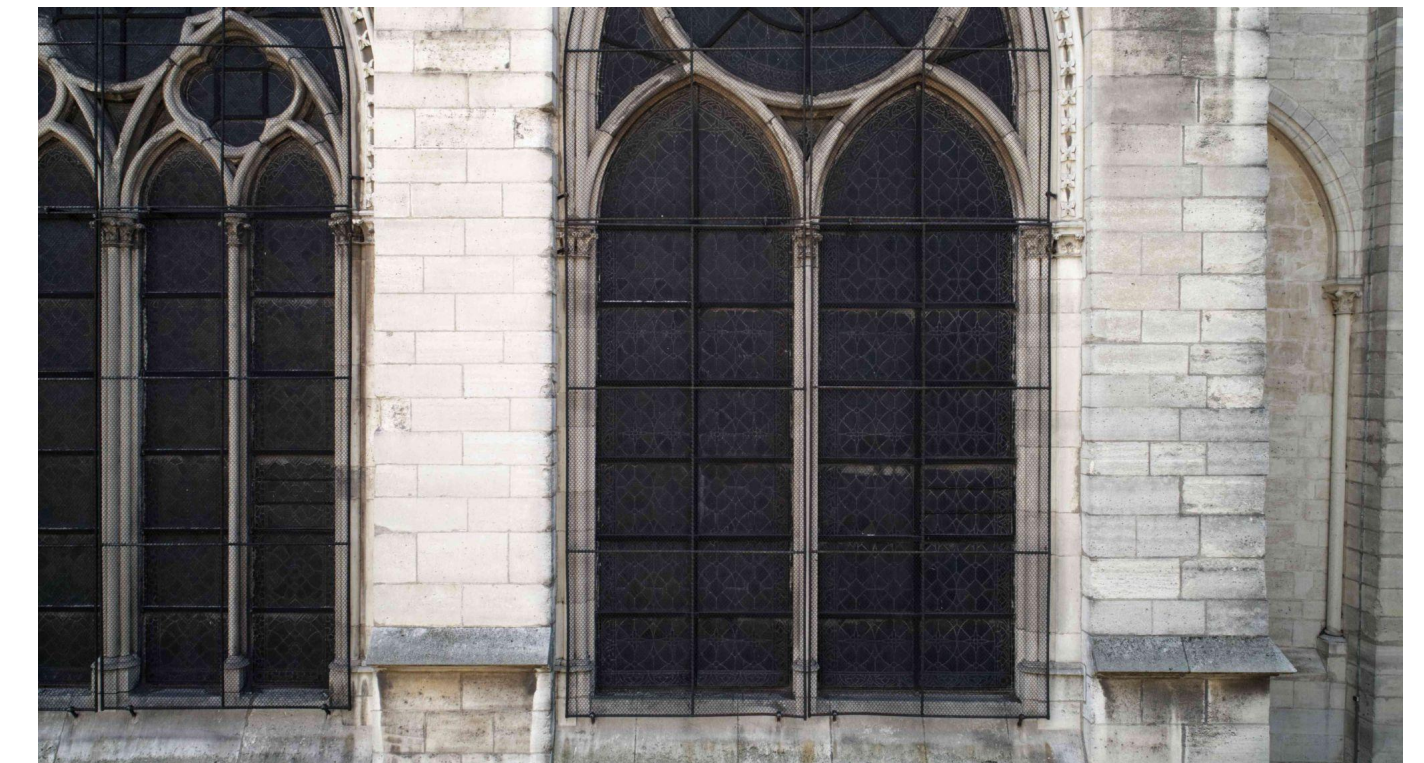
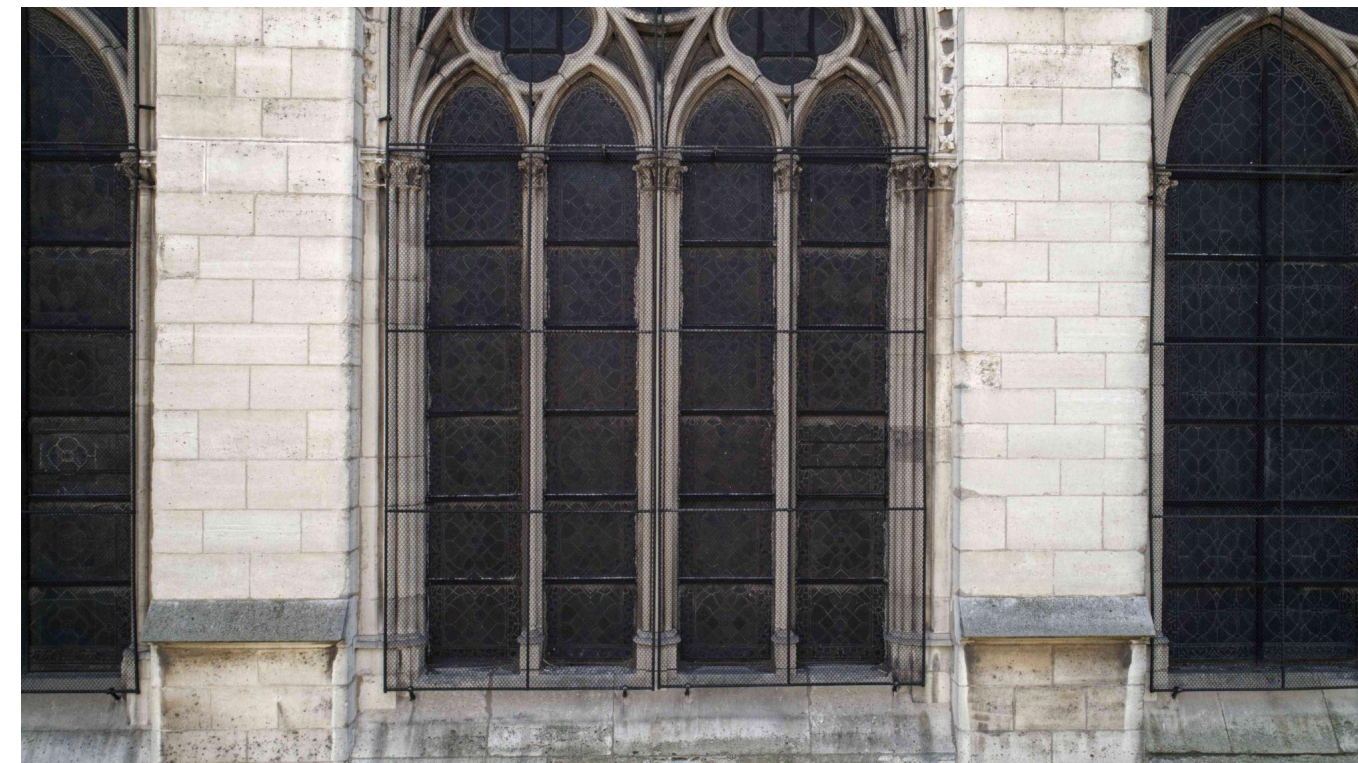
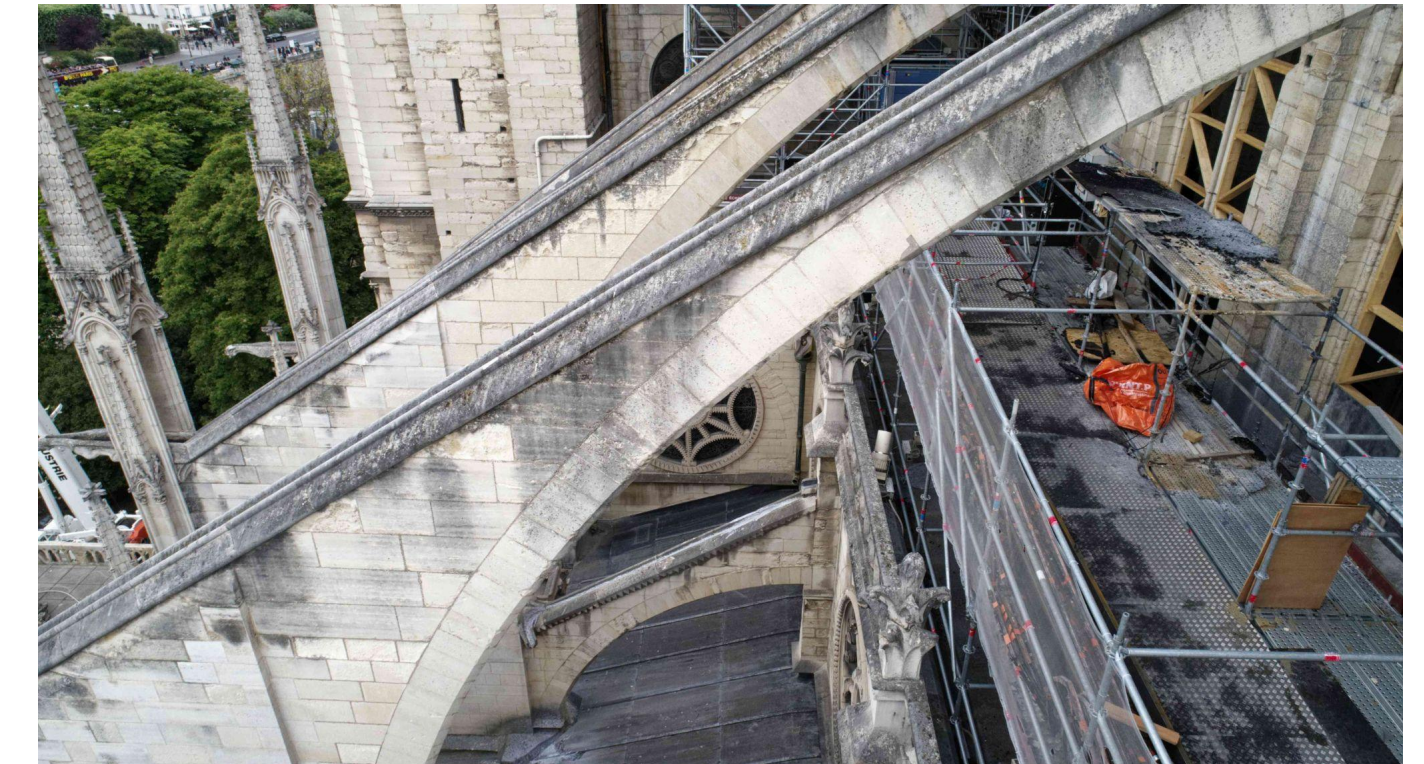
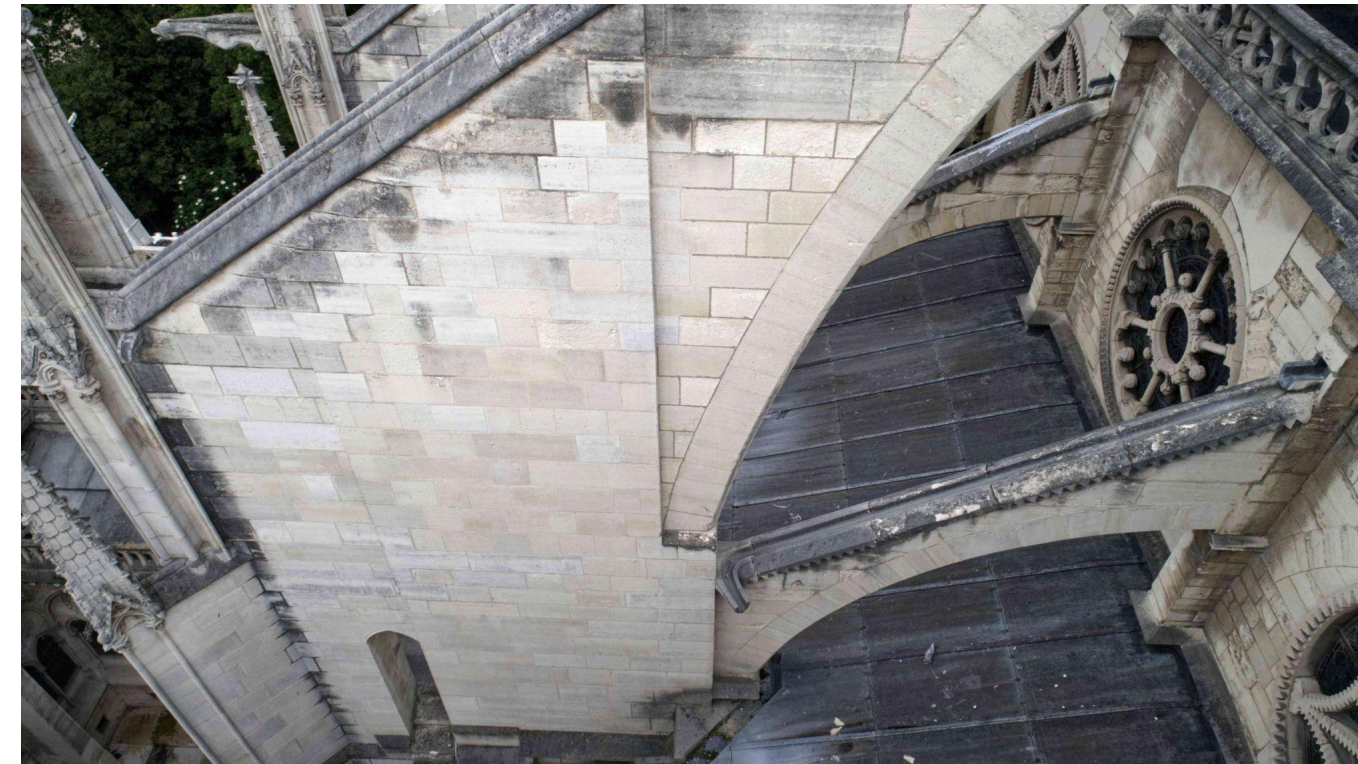
- Apport de mon expérience sur les contenus patrimoniaux et l'indexation automatique de contenus visuels au sein du projet ERC nDame_Heritage.
- Deux objectifs :
 - la **structuration** large échelle,
 - la **localisation** de nouvelles images.
- Appliquer d'abord les approches qui fonctionnaient dans la thèse, puis les **adapter aux données plus spécifiques du chantier**.



© DHAAP / CVP, © Flickr, © Musée Albert Kahn,
© MAP, © IGN, © C. Lansiaux / DHAAP / Roger-Viollet

Un jeu d'images aux spécificités différentes

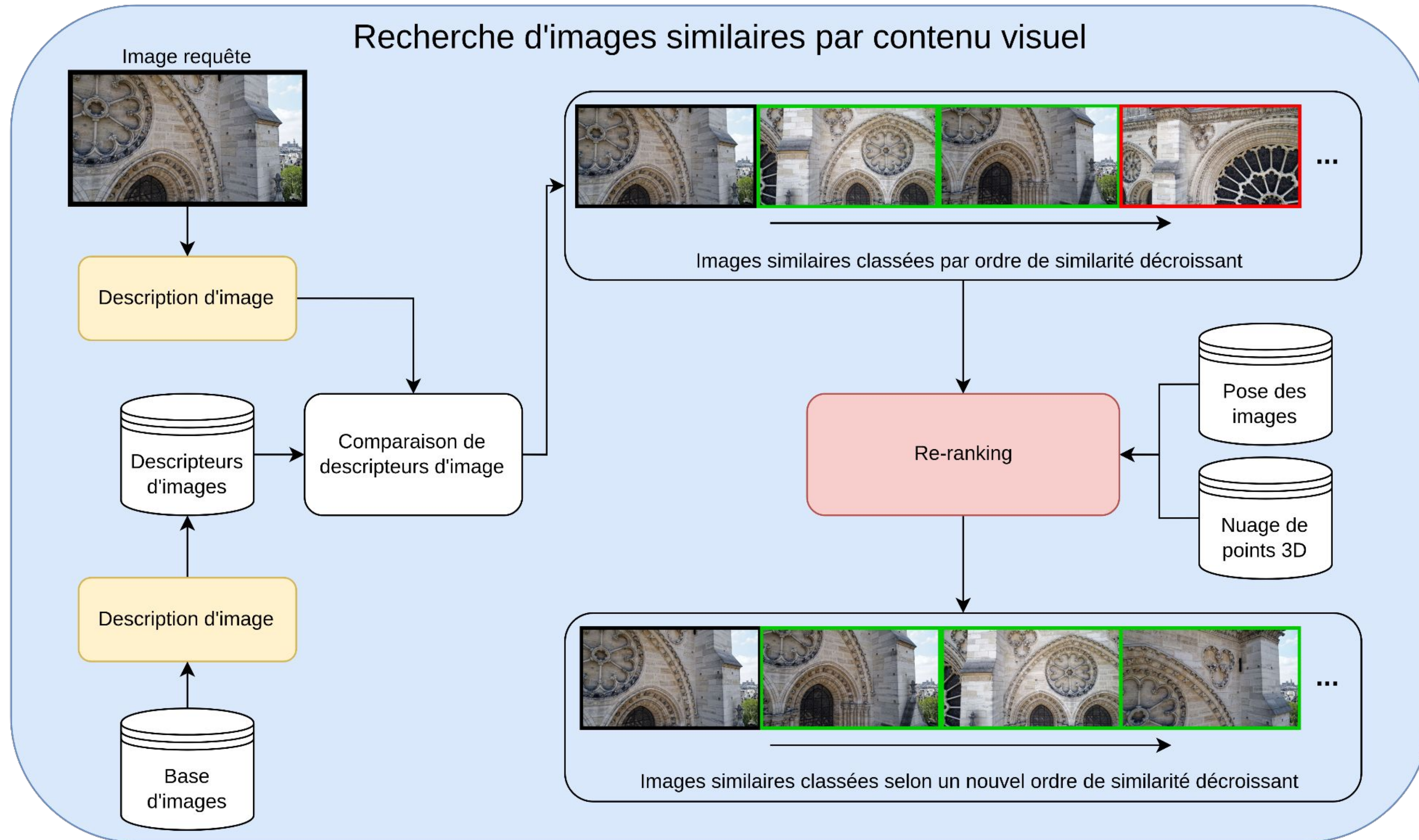
- 10 901 images :
 - représentant l'extérieur,
 - avec une large couverture du monument,
 - et une grande co-visualisation entre images.
- Différences par rapport aux jeux patrimoniaux de la thèse :
 - unique objet d'étude,
 - acquisition systématique,
 - niveau de détail plus homogène,
 - grande redondance d'éléments similaires mais différents.
- Créer une vérité terrain pour la recherche d'images similaires :
 - de manière semi-automatique,
 - pour évaluer la performance des algorithmes de recherche d'images,
 - si deux images observent le même élément.



© AGP

La recherche d'image par similarité visuelle

Principe général



© AGP

La recherche d'image par similarité visuelle

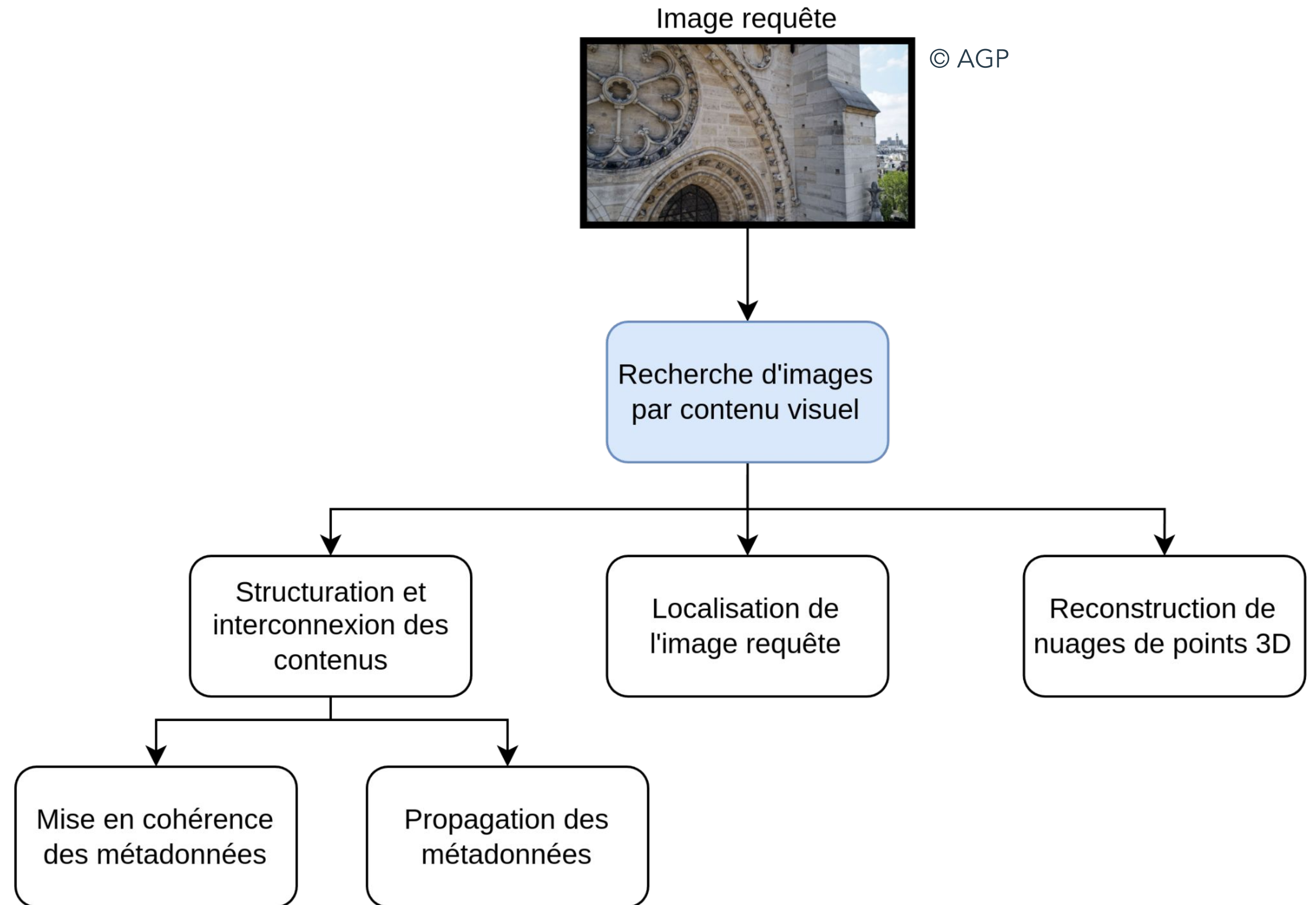
Utilisations subséquentes

- Principe : exploiter le **contenu visuel** des images pour déterminer un **score de similarité** entre images.
- **Création de liens** plus ou moins certains entre images.

- **Objectifs multiples** :
 - mise en cohérence ou propagation de métadonnées,
 - localisation d'images,
 - reconstruction 3D par Structure from Motion.

- **Avantages** :
 - **à large échelle**,
 - indépendant des métadonnées,
 - **offline ou online**.

- **Inconvénients** :
 - la **qualité du contenu visuel** des images est importante,
 - les **différences de niveau de détail** impactent la recherche d'image.

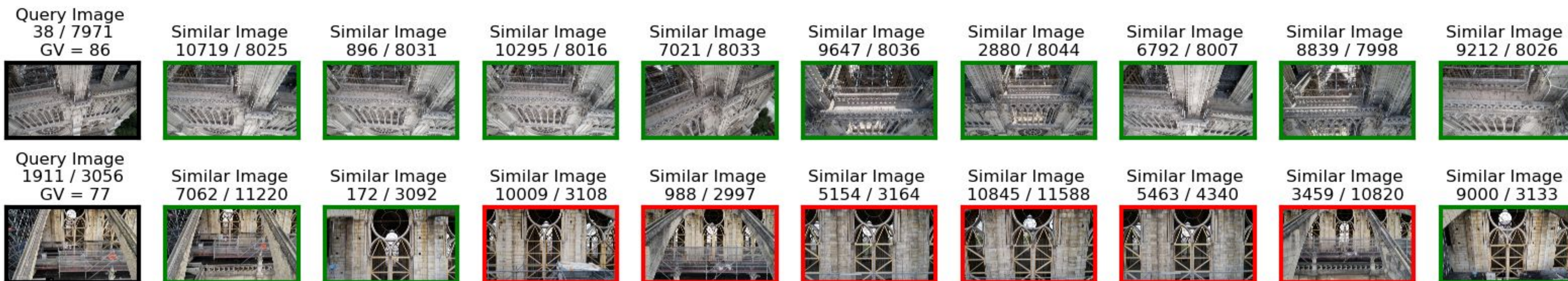


La recherche d'image par similarité visuelle

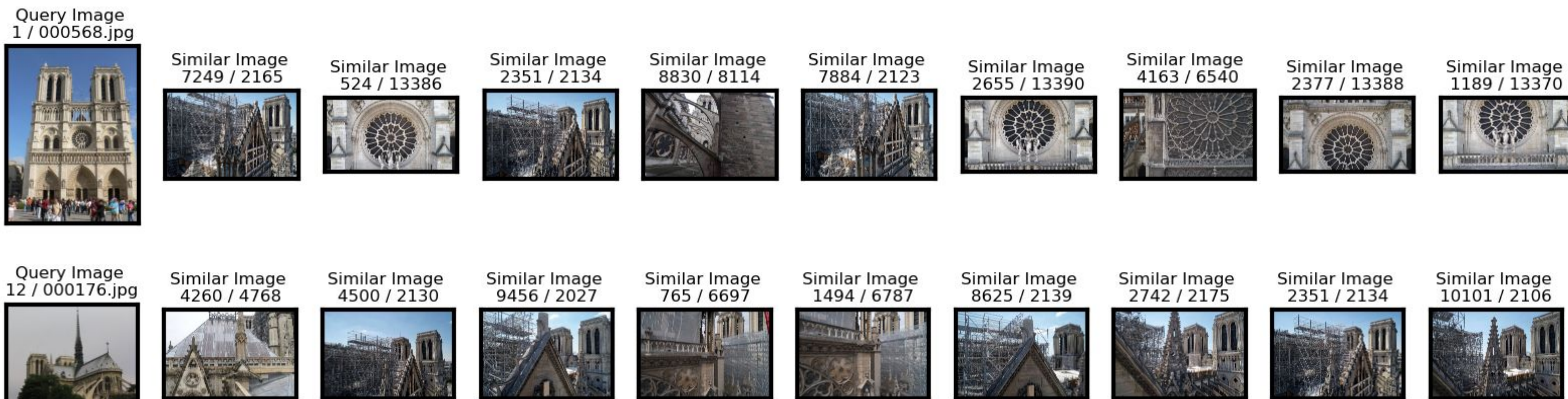
Exemples

- Descripteur local agrégé : How + ASMK (Tolias et al. 2016, 2020)
- Dictionnaire de mots visuels adapté à la cathédrale
- Recherche d'instance : **mêmes éléments** d'un objet présents sur les images, mais **pas face avant/face arrière** d'un même objet (cf reprojection d'annotations)

- Requête **intra-collection** :
 - Vert : même élément
 - Rouge : élément différent



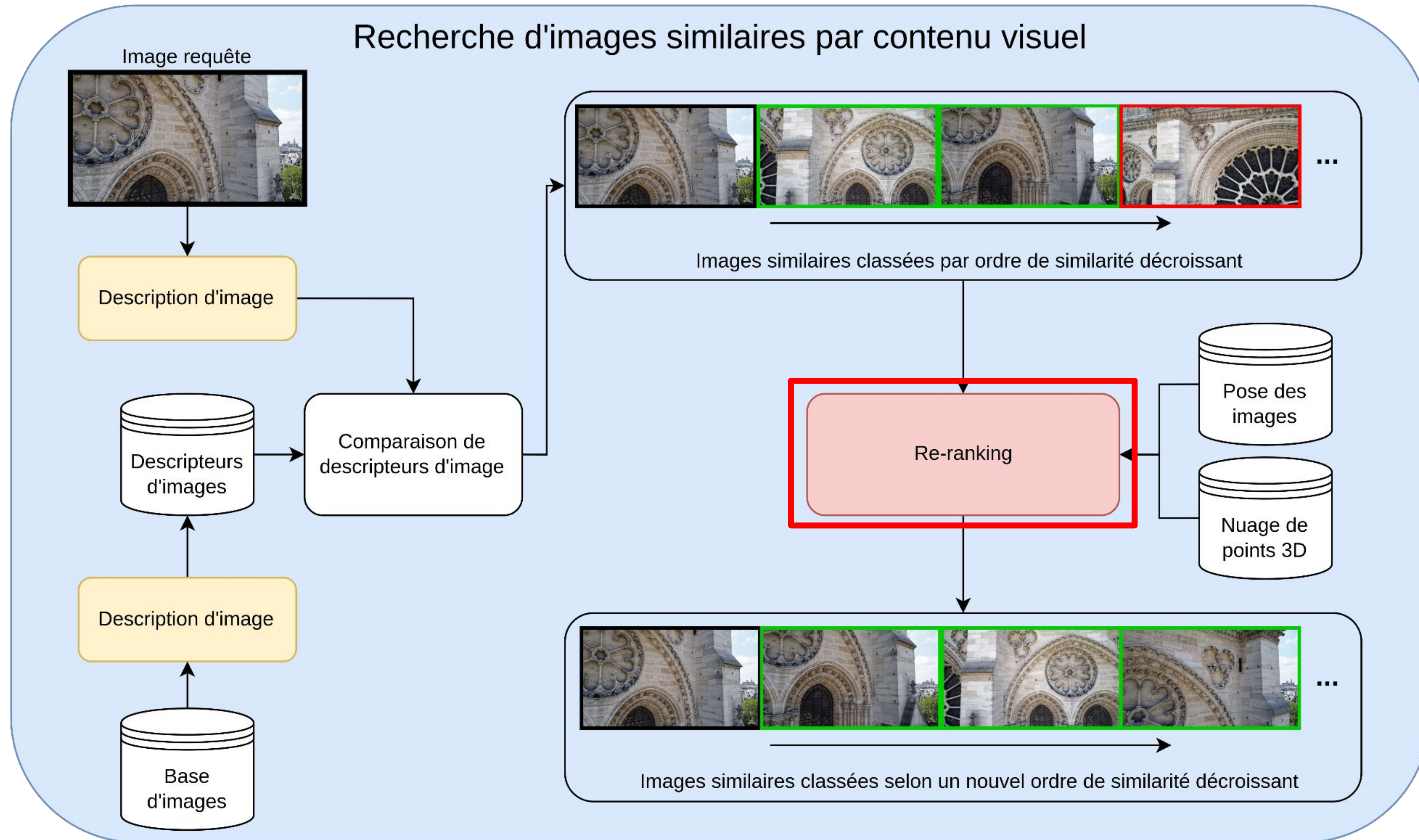
- Requête **extra-collection** :
 - Pas de vérité terrain
 - Niveau de détail différent



© AGP, © Flickr

La recherche d'image par similarité visuelle

Rappel du principe général

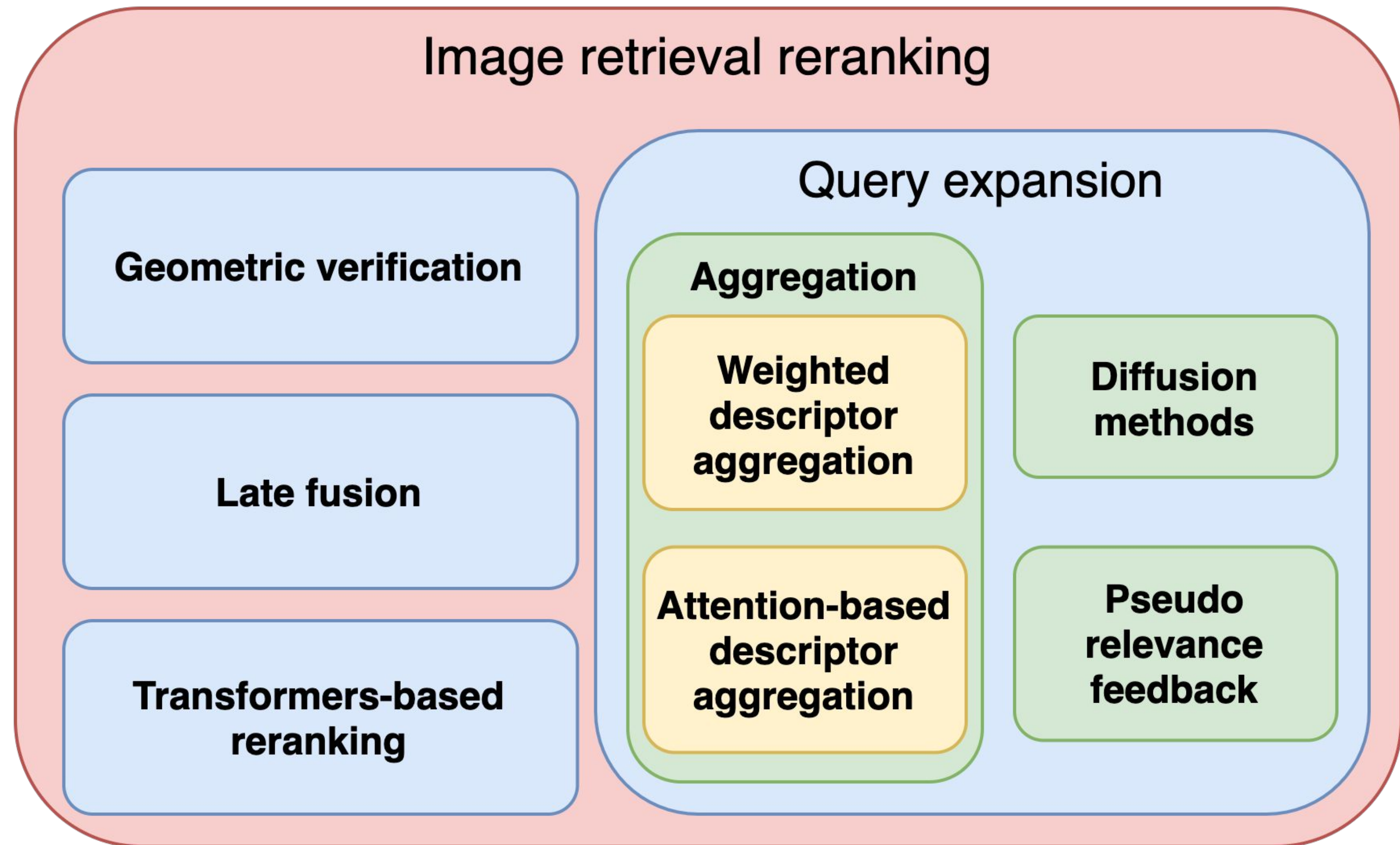


© AGP

La recherche d'image par similarité visuelle

Méthodes de re-ranking

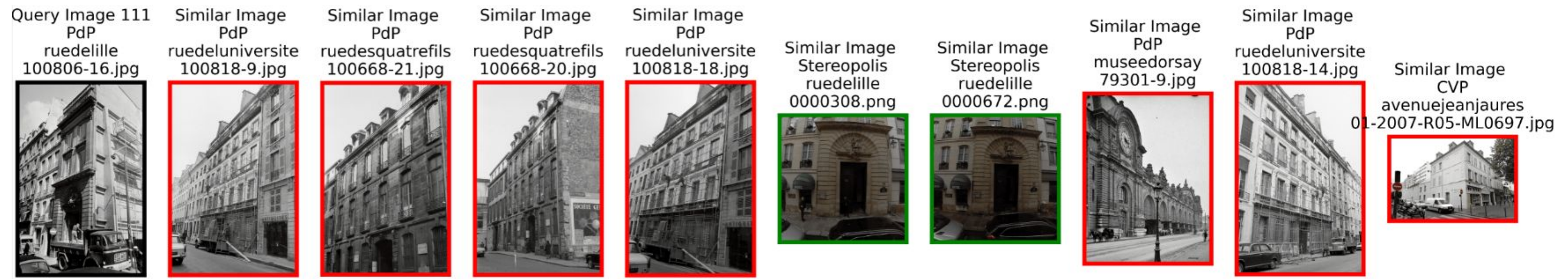
- **Contenus patrimoniaux spécifiques** :
 - mode/date/type d'acquisition,
 - contenus représentés.
- **Difficultés pour la recherche d'image**
- Importance d'une **étape de re-ranking** adaptée aux contenus
- **Nombreuses méthodes** de re-ranking :
 - plus ou moins efficaces,
 - plus ou moins coûteuses,
 - plus ou moins **combinables**.



La recherche d'image par similarité visuelle

Exemples de re-ranking

How-A



La recherche d'image par similarité visuelle

Exemples de re-ranking

How-A

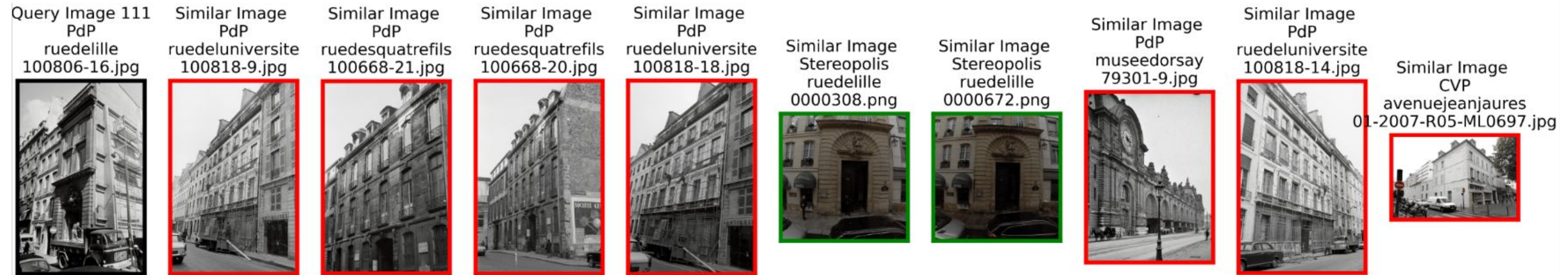
Query Image 111 PdP ruedelille 100806-16.jpg	Similar Image PdP ruedeluniversite 100818-9.jpg	Similar Image PdP ruedesquatrefils 100668-21.jpg	Similar Image PdP ruedesquatrefils 100668-20.jpg	Similar Image PdP ruedeluniversite 100818-18.jpg	Similar Image Stereopolis ruedelille 0000308.png	Similar Image Stereopolis ruedelille 0000672.png	Similar Image PdP musedorsay 79301-9.jpg	Similar Image PdP ruedeluniversite 100818-14.jpg	Similar Image CVP avenuejeanjaures 01-2007-R05-ML0697.jpg
Query Image 111 PdP ruedelille 100806-16.jpg	Similar Image PdP ruedesquatrefils 100668-25.jpg	Similar Image PdP ruedelille 144407-17.jpg	Similar Image PdP ruedelille 79300-11.jpg	Similar Image PdP ruedesquatrefils 100668-21.jpg	Similar Image PdP ruedeluniversite 100818-14.jpg	Similar Image PdP ruedeluniversite 100818-9.jpg	Similar Image PdP ruedesquatrefils 80704-22.jpg	Similar Image PdP ruedesquatrefils 100668-20.jpg	Similar Image PdP ruedesquatrefils 144413-28.jpg

How-A
+ GNN-R

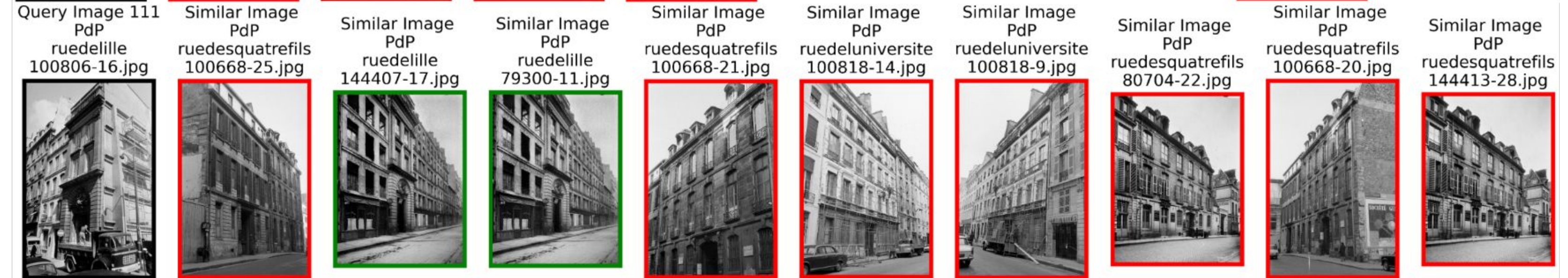
La recherche d'image par similarité visuelle

Exemples de re-ranking

How-A



How-A + GNN-R



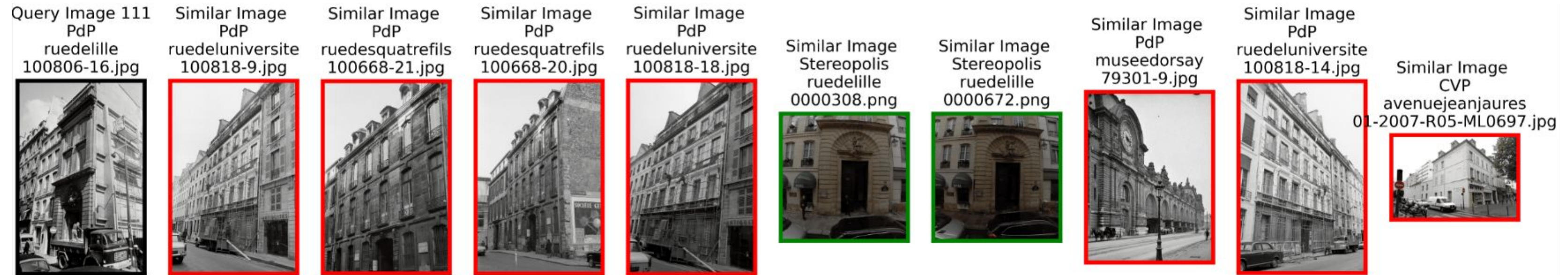
How-A + RANSAC-SG



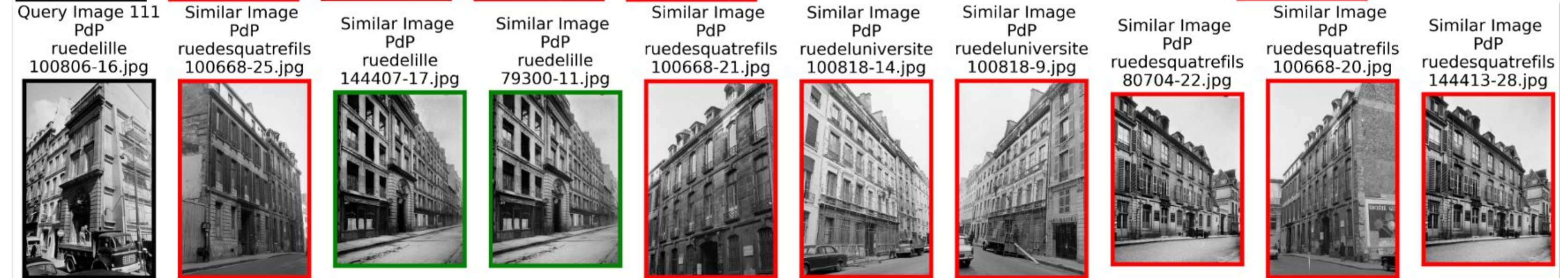
La recherche d'image par similarité visuelle

Exemples de re-ranking

How-A



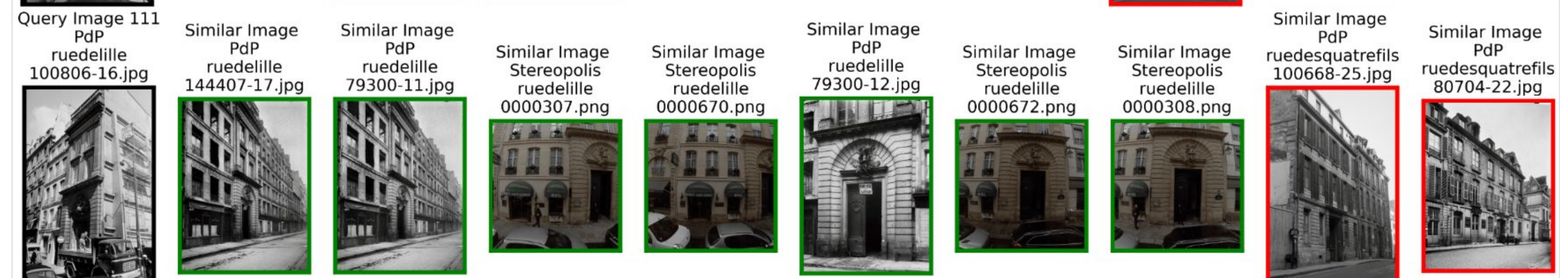
How-A + GNN-R



How-A + RANSAC-SG



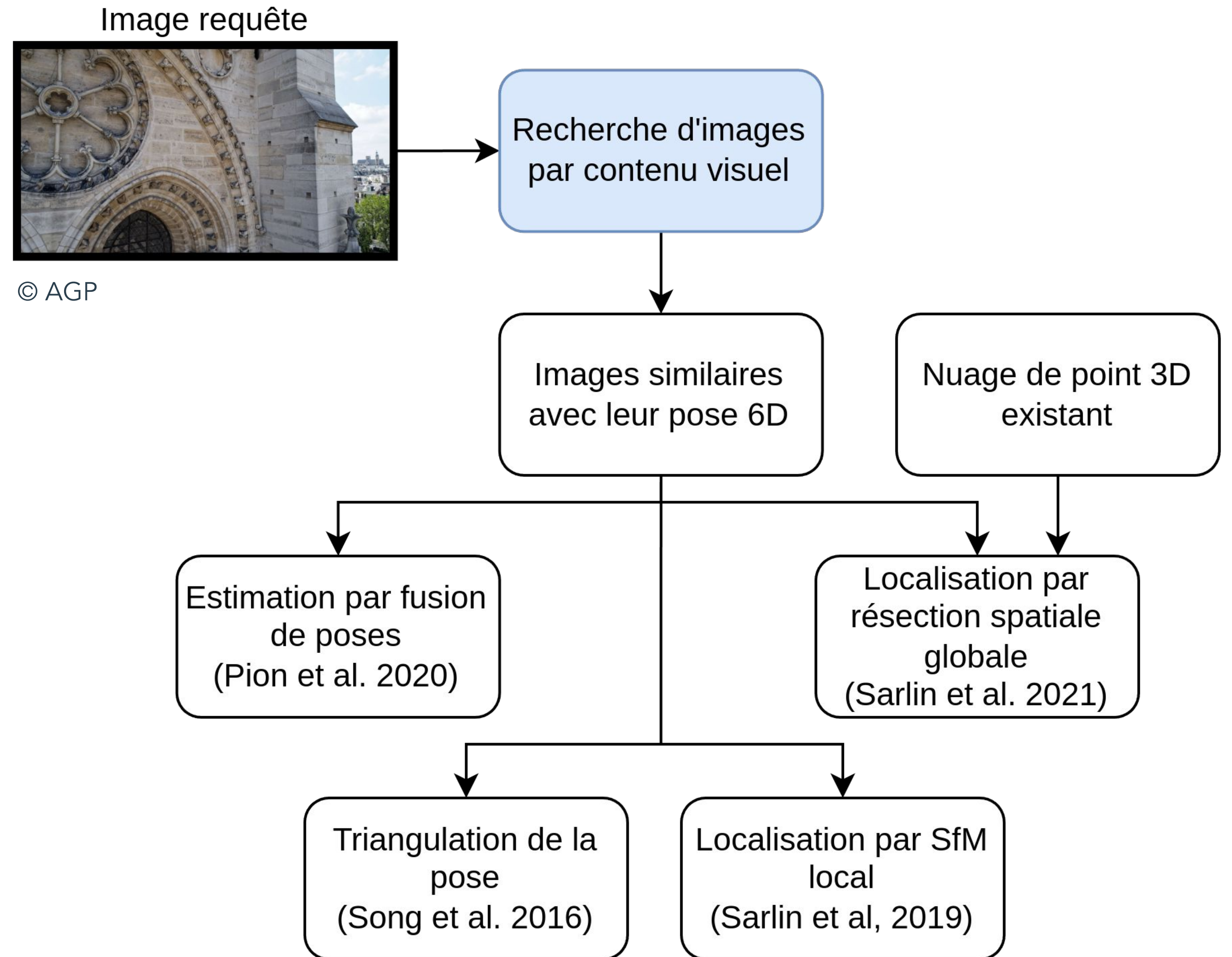
How-A + RANSAC-SG + GNN-R



Localisation d'images

Différentes approches et résultats divers

- **A partir d'images similaires localisées**, estimation de la localisation d'une image requête non localisée.
- **Différentes localisations** finales possibles :
 - localisation 2D,
 - position 3D,
 - pose 6D.
- **Différentes données** utilisées :
 - pose des images connues,
 - nuage de points 3D reconstruit à la volée,
 - nuage de points 3D préexistant.
- **Différentes méthodes** possibles :
 - exploitant différentes informations,
 - avec différents degrés d'incertitude.



Localisation d'image

Un exemple



- Localisation par triangulation de la pose:

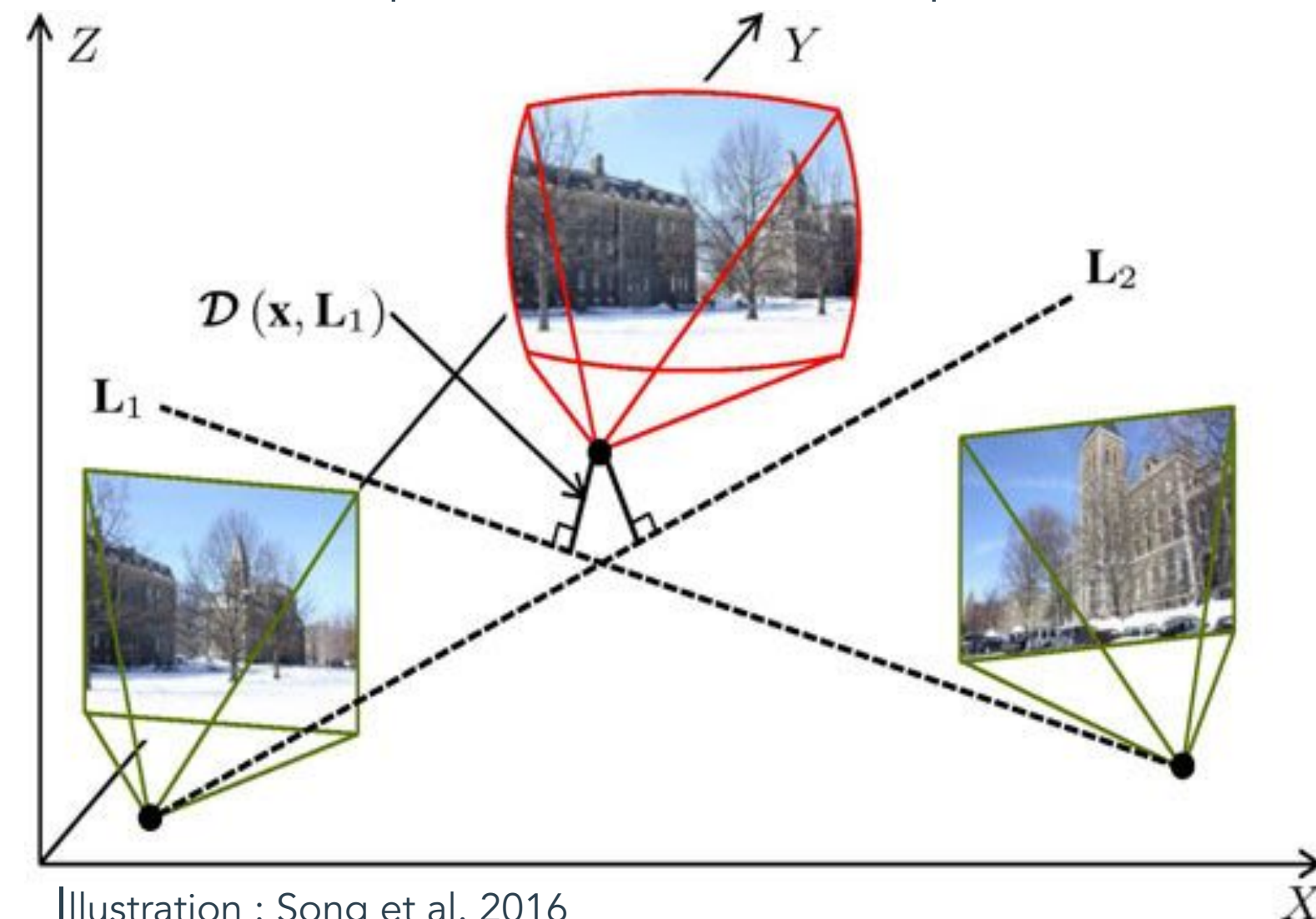
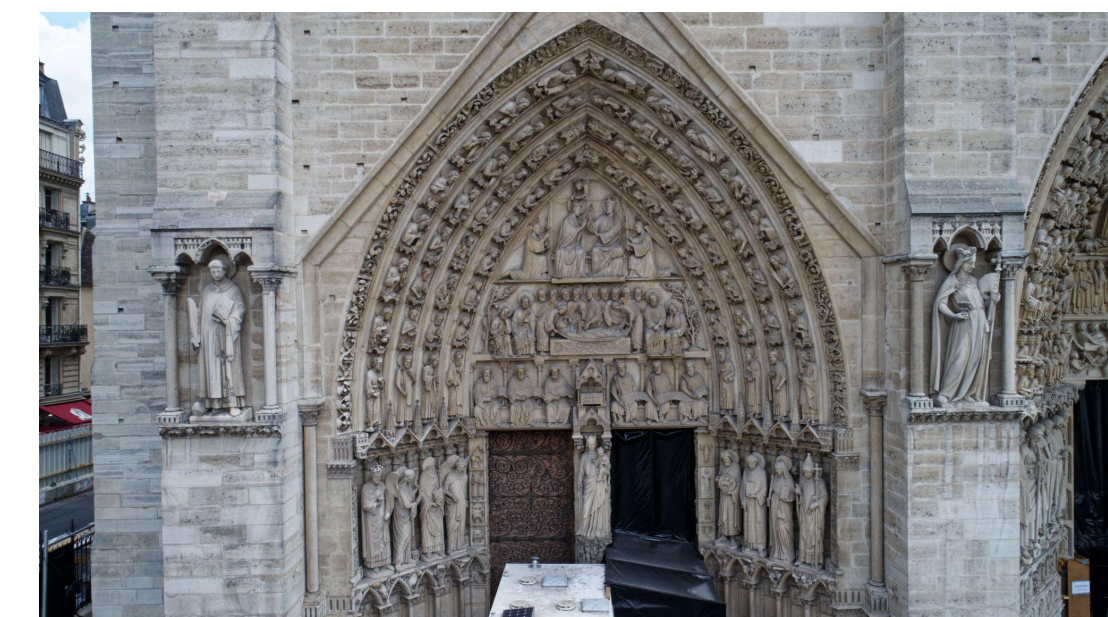
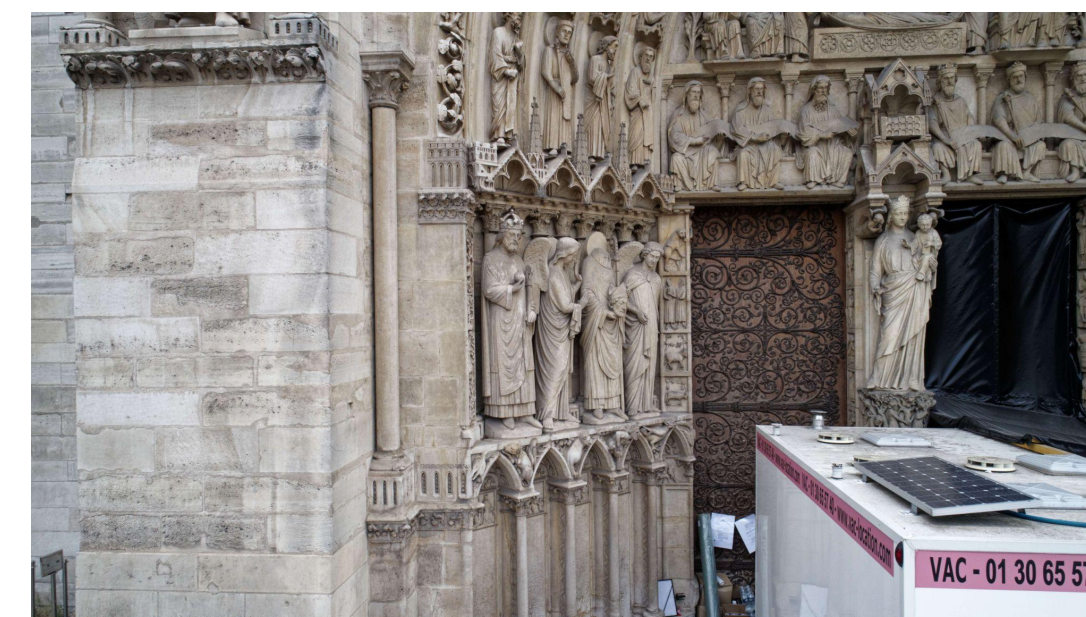


Illustration : Song et al. 2016



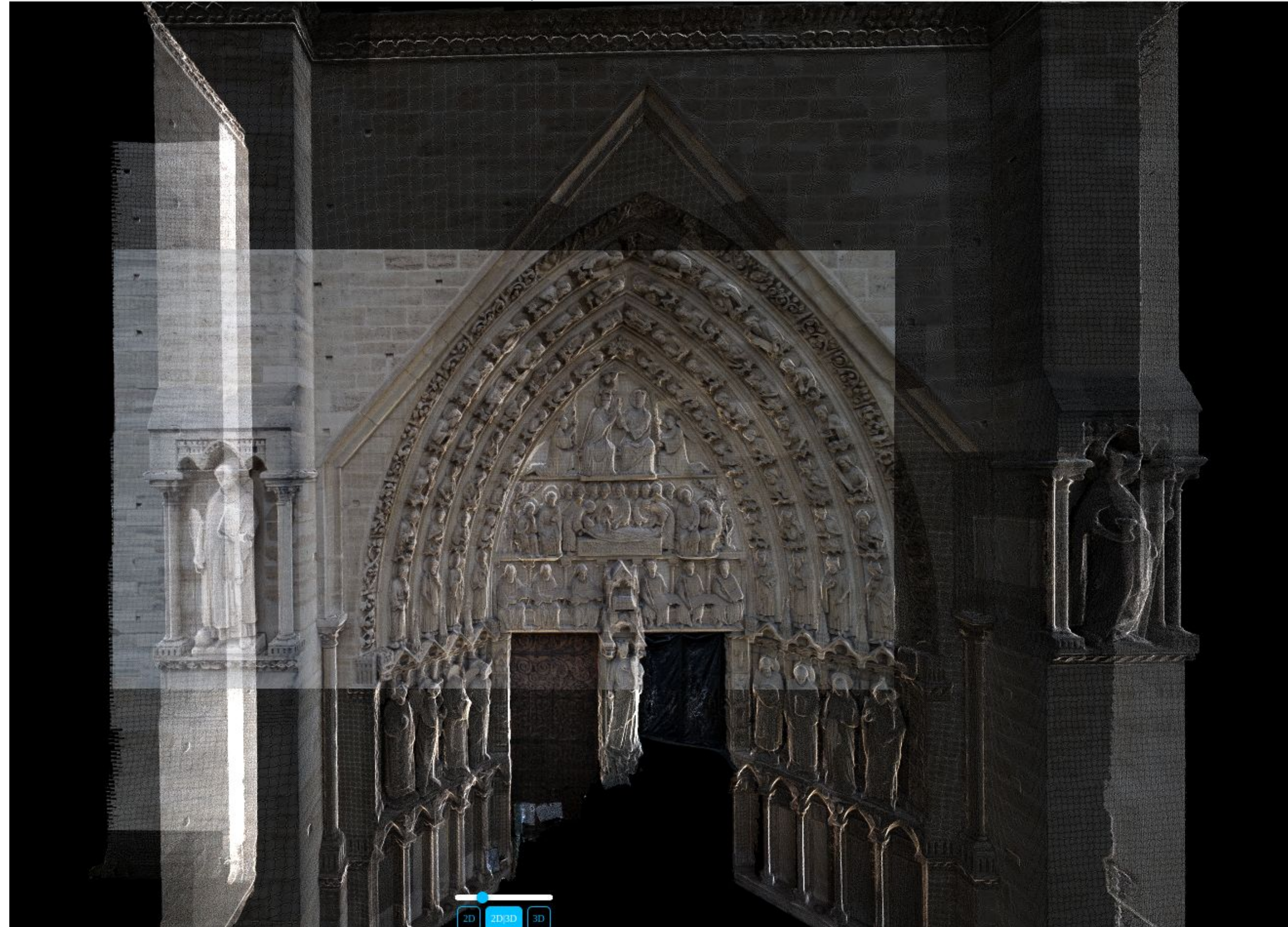
- En entrée du processus :
 - une image requête à la pose inconnue,
 - n (3 ici) images similaires avec leurs poses 6D connues.
- En sortie :
 - une pose 6D pour l'image requête.



Localisation d'images

Un exemple

La pose réelle de l'image

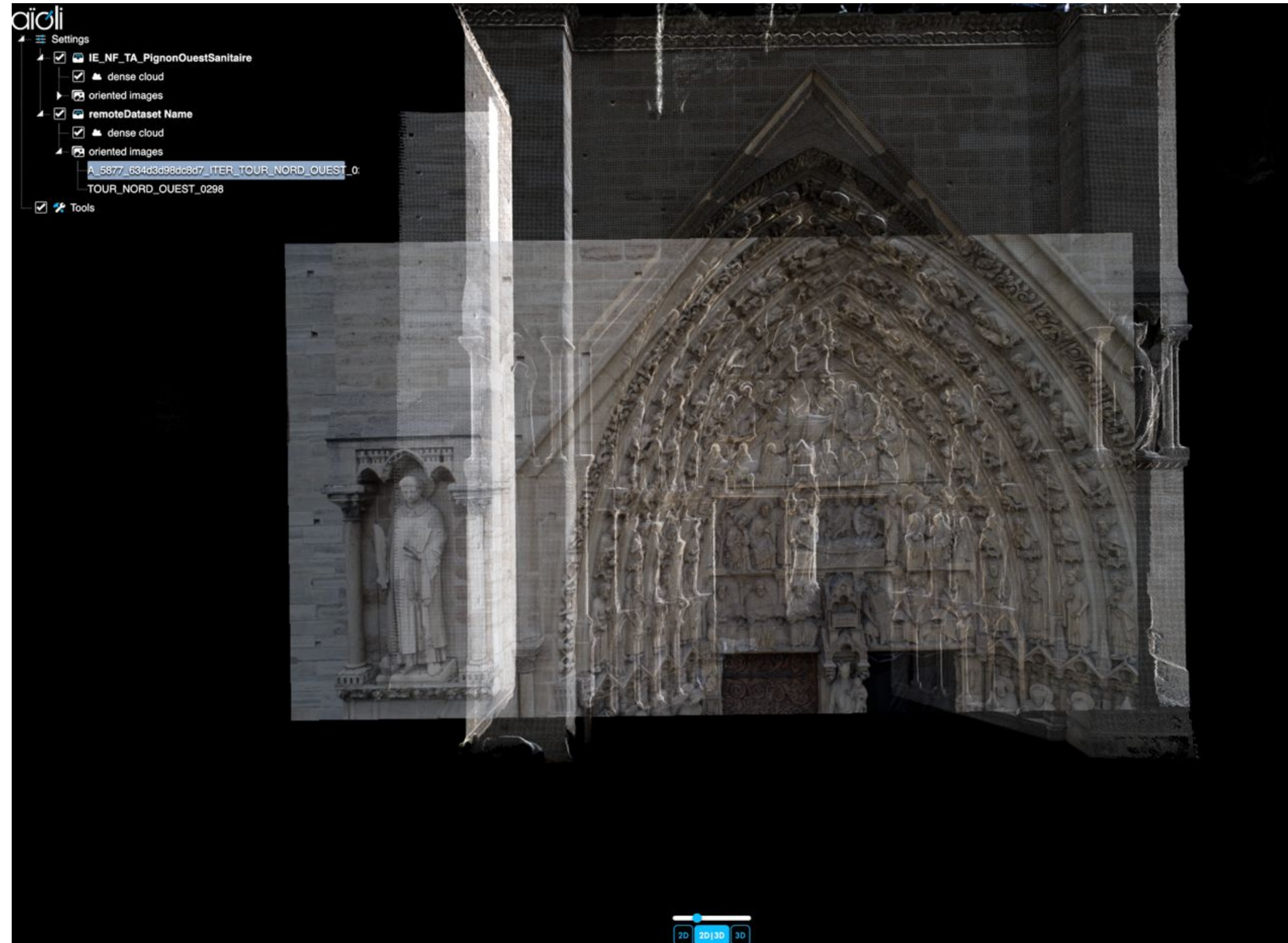


© AGP

Localisation d'images

Un exemple

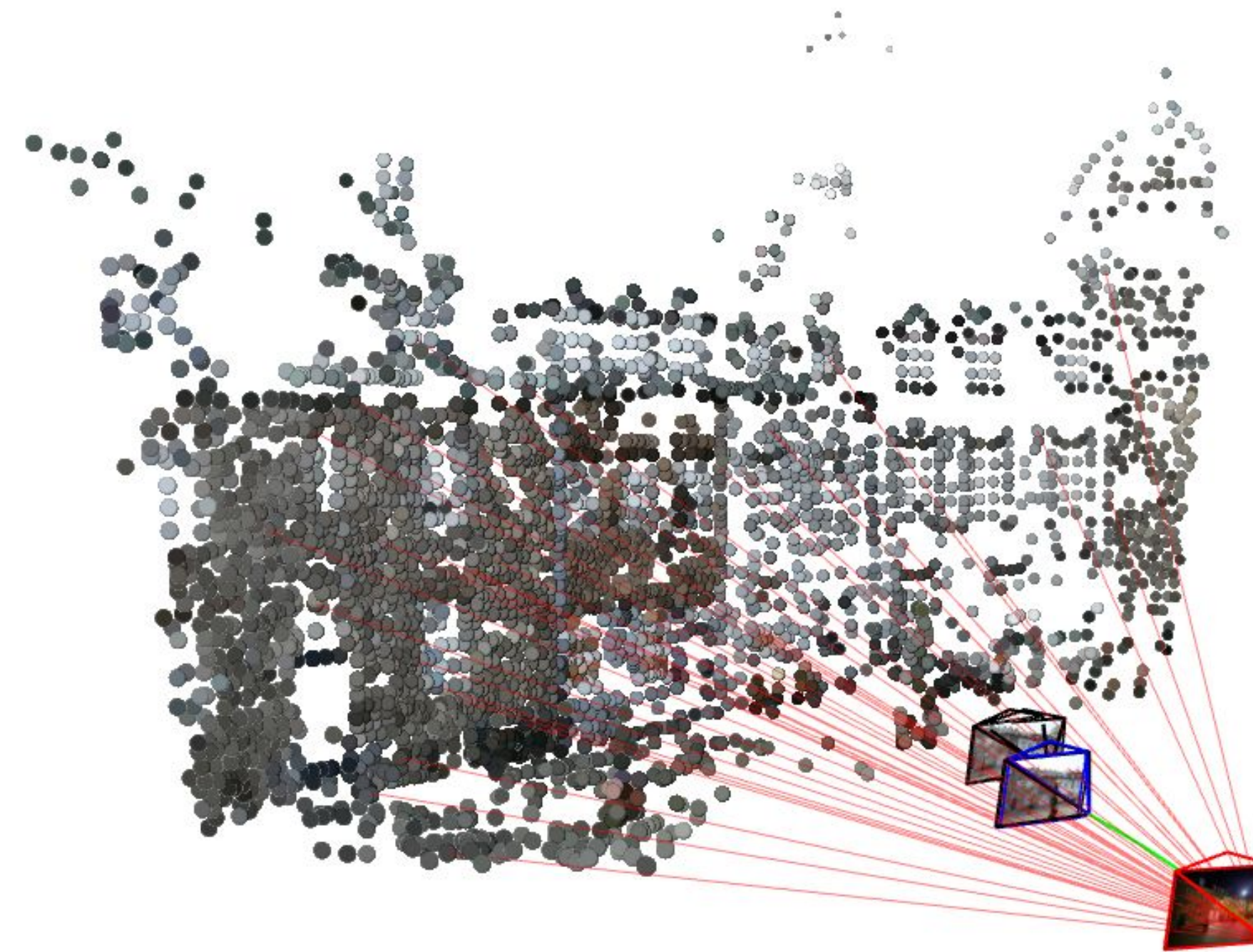
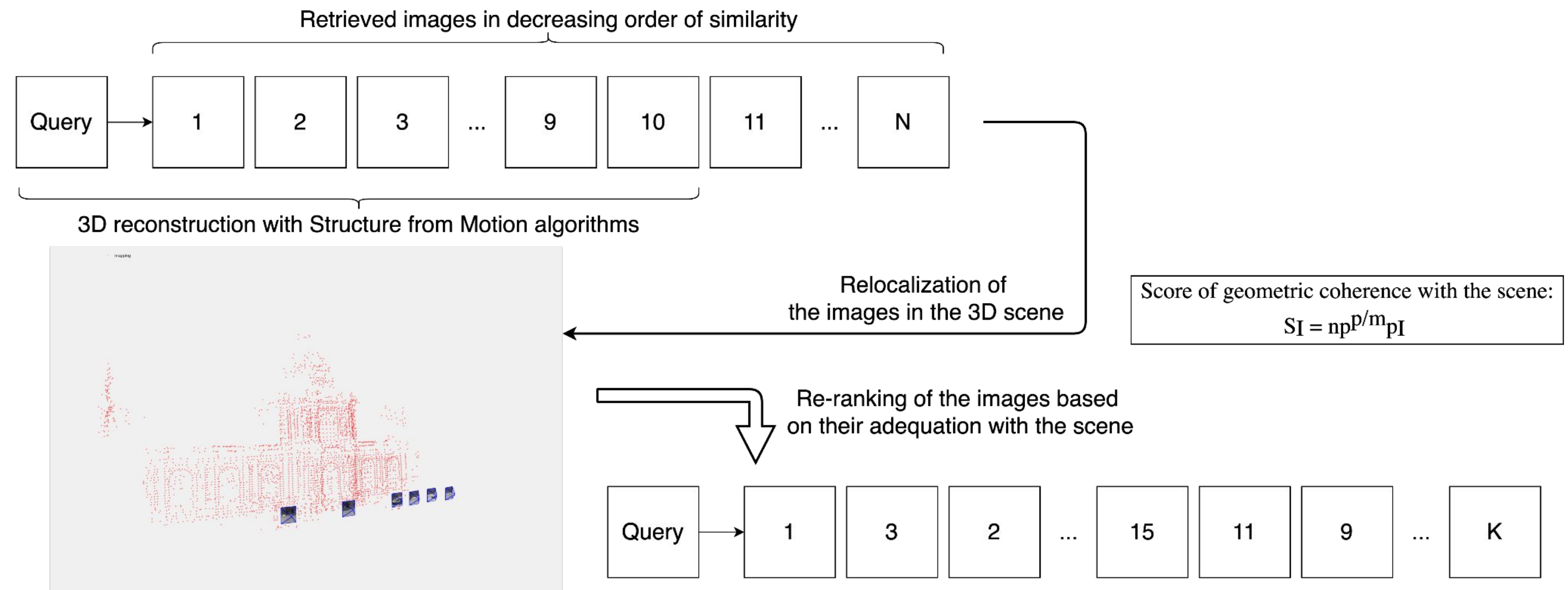
La pose estimée de l'image



Axes de recherche en cours

Exploitation de l'information 3D

- Exploitation de la **3D existante**
- Pour le **re-ranking** :
 - poursuite d'une proposition issue de la thèse,
 - objectif de valider la **cohérence d'une image dans une scène plus globale**,
 - exploiter une 3D plus "solide" qu'une 3D reconstruite à la volée à partir des premières images.
- Pour la **localisation** d'images :
 - par **résection spatiale** avec des matches 2D-3D (en rouge à droite),
 - méthode plus **robuste** et aux meilleurs résultats que les autres
 - avec une 3D plus "solide" qu'une 3D reconstruite à la volée à partir des premières images.

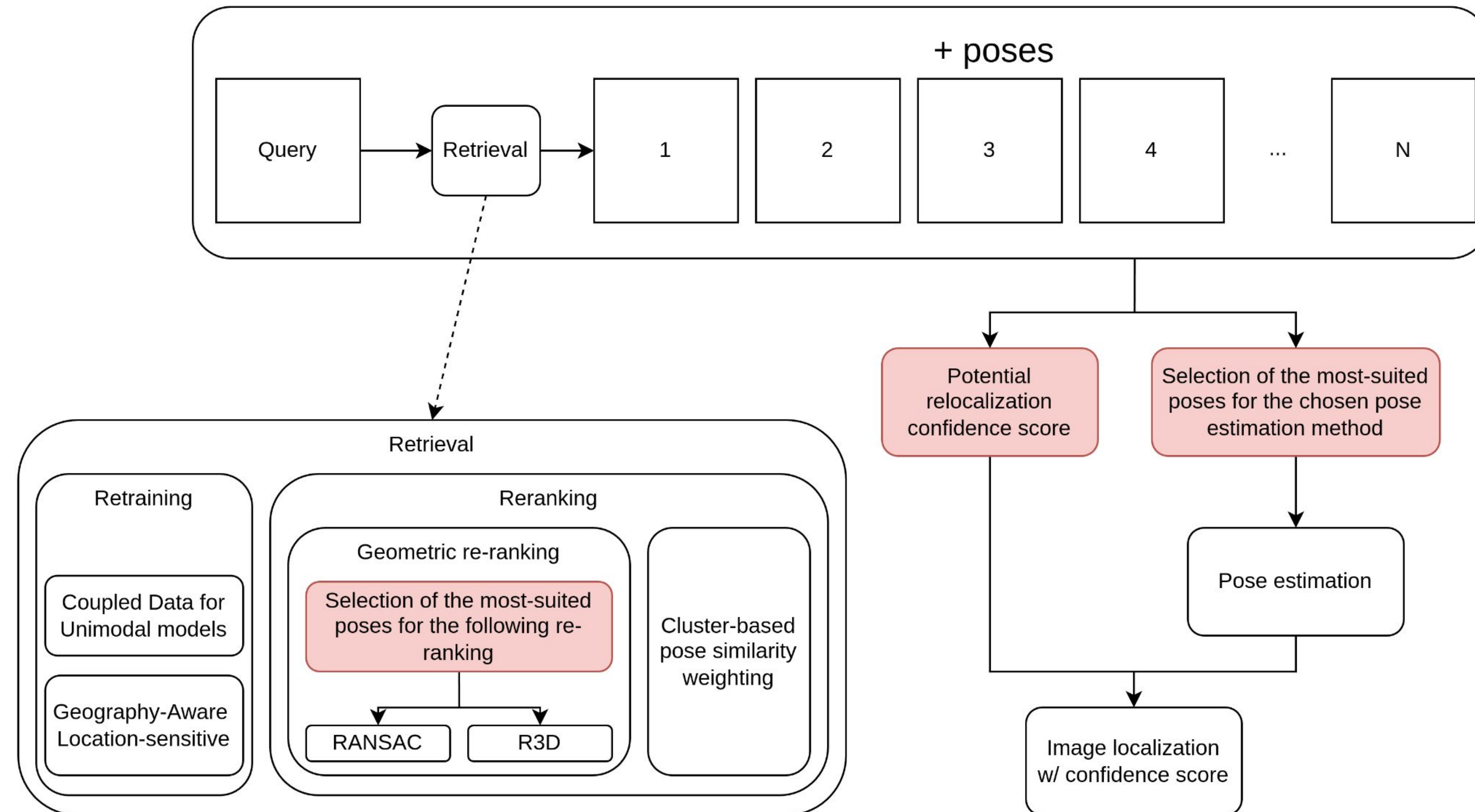


Axes de recherche en cours

Sélection de poses a priori pour optimiser la localisation

- **Localisation** automatique potentiellement **faussée** :
 - **recherche d'images** incorrecte,
 - mauvaise **initialisation des poses** similaires :
 - pas adaptées pour la triangulation de pose,
 - pas adaptées pour la reconstruction 3D à la volée pour la localisation.
- **Validation visuelle fastidieuse et inadaptée** au vu des volumes d'images exploités.
- Pour pallier ces erreurs de manière automatique :
 - estimation de la **confiance** dans la qualité de la **recherche d'images**,
 - définition **d'indicateurs** à maximiser pour **sélectionner** les images les plus adaptées,
 - spécifiques à la méthode de localisation choisie.

Retrieval suitability estimation for localization



Conclusion

- Dans la **continuité des travaux engagés lors de ma thèse** mais sur un objet d'études aux caractéristiques bien différentes.
- Mise en place de méthodes de **structuration à large échelle exploitant le contenu visuel** :
 - création d'une **vérité terrain** robuste,
 - optimisation de la recherche d'image,
 - évaluation des méthodes de **re-ranking** les plus adaptées à cette collection.
- Mise en place de premières méthodes de **localisation d'image** par triangulation de pose
- **Deux axes de recherche** principaux :
 - exploitation de l'**information 3D robuste** existante :
 - pour la structuration,
 - et la localisation,
 - mise en place d'**indicateurs** permettant d'évaluer la **pertinence de la recherche d'image** pour la localisation et de sélectionner les images les plus adaptées à chaque méthode de localisation.