

Navigation immersive dans des photographies historiques

Mathieu Brédif

Scarleth Castro, Evelyn Paiz, Alexandre Devaux, Sidonie Christophe

LASTIG, Univ. Gustave Eiffel, ENSG, IGN, F-94160 Saint-Mandé, France

Photographies historiques

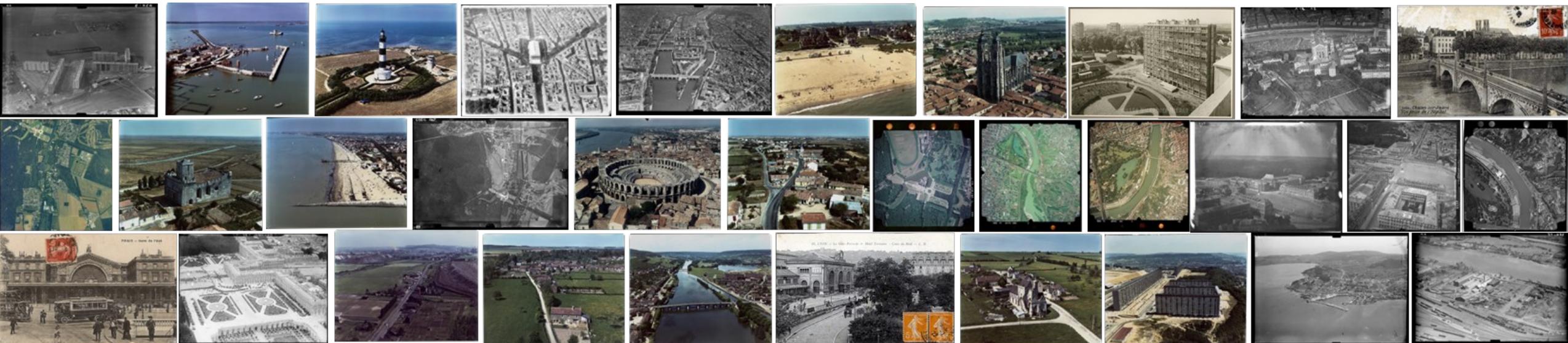
- Alegoria (ANR) : Valorisation de collections d'images historiques décrivant le territoire français à différentes époques.



Images vues du sol

Photographies historiques

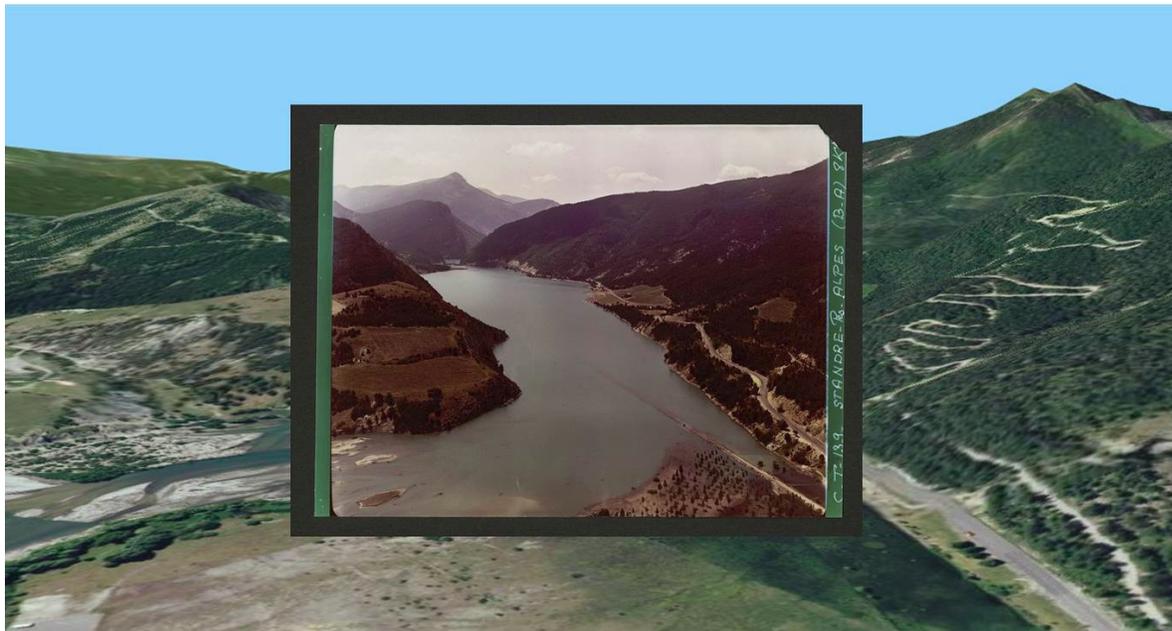
- Alegoria (ANR) : Valorisation de collections d'images historiques décrivant le territoire français à différentes époques.



Images aériennes obliques

Navigation immersive

- Visualisation interactive de ces images dans un contexte 3D
- Desktop = navigateur web (pas encore en RA/RV)



Motivation

Replonger dans le passé pour mieux le comprendre et l'étudier

A spatio-temporal web-application for the understanding of the formation of the Parisian metropolis [Blettery 2020]

La méthode de la photo-interview à partir de la photographie aérienne. Le cas d'un bidonville à Nanterre dans les années 1960 [Delavoipière 2020]

Restituer les bidonvilles de Nanterre - Retour d'expérience d'un outil de visualisation 3D des données spatiales à l'usage des sciences sociales [Lecat 2021]



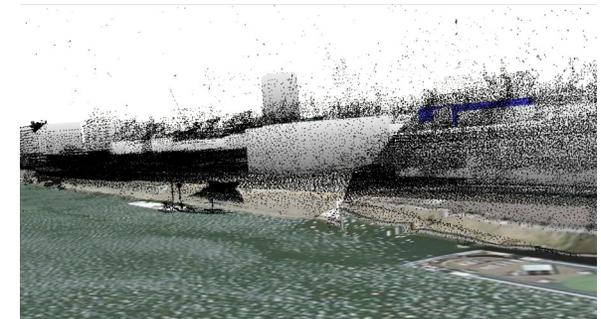
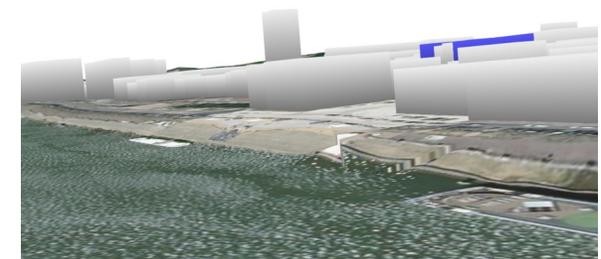
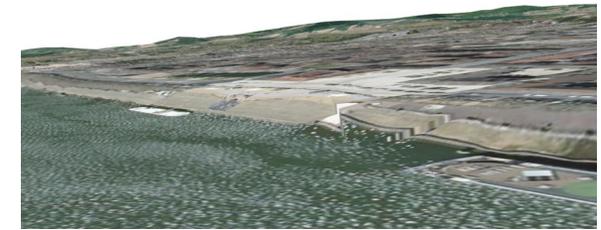
Données utilisées

Données historiques (1900-2000)

- Photographies (10-100Mpix)
Entre 100 et 50000 photos par fonds
Datation approchée (intervalle de 1 jour à 10 ans)
Géoréférencement à la commune ou à la rue

Données actuelles (2021)

- Modèle 3D sol : BD Alti IGN (~1m)
- Texture : BD Ortho IGN (~10cm)
- Modèle 3D bâti : BD Topo IGN (~1m)
- Nuages de points 3D lidar (véhicule terrestre)
milliards de points
géoréférencement (~1cm)



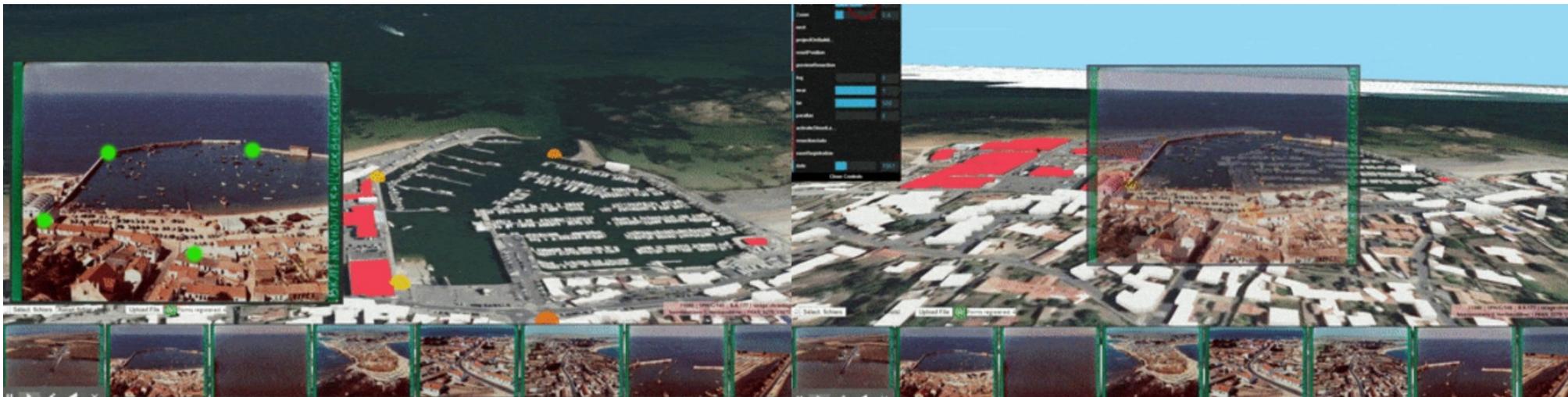
Géoréférencement des photographies

Building Rome in a day (Agarwal et al, ICCV2009)



→ Nombre insuffisant d'images historiques pour un géoréférencement (pose + calibration interne) et une reconstruction 3D des scènes historiques.

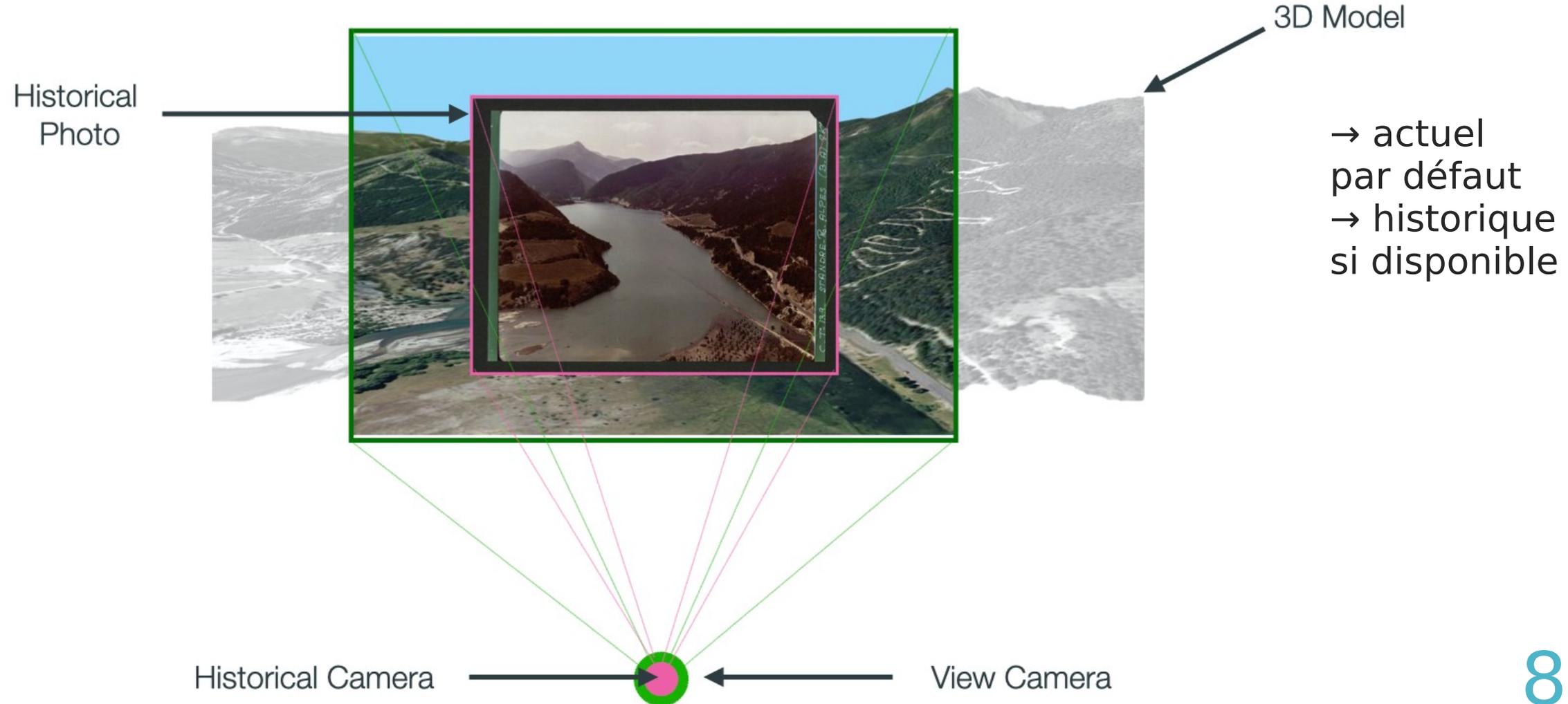
Mais géoréférencement (Semi-)automatique **approché** possible des cameras utilisant les données actuelles [Blanc 2018] [Harrach 2019] [Blettery 2021]



Que faire sans modèle 3D historique ?

Du point de vue (dézoomé) d'une image

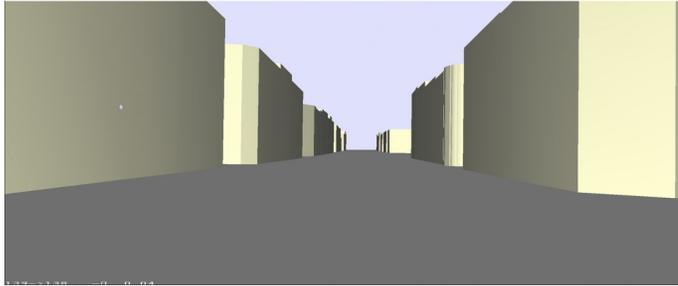
Visualization Result



Comment se déplacer ?

Rendu Basé Image (IBR)

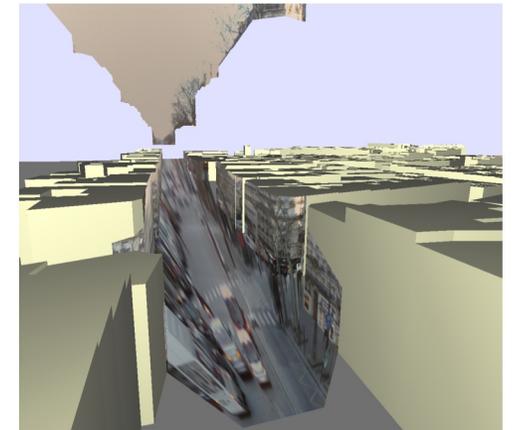
→ Mélanger les reprojctions des images sur un modèle 3D approché



Le même point de vue



D'autres points de vue



*Image-Based Rendering of LOD1 3D City Models
for traffic-augmented Immersive Street-view Navigation, [Brédif 2013]*

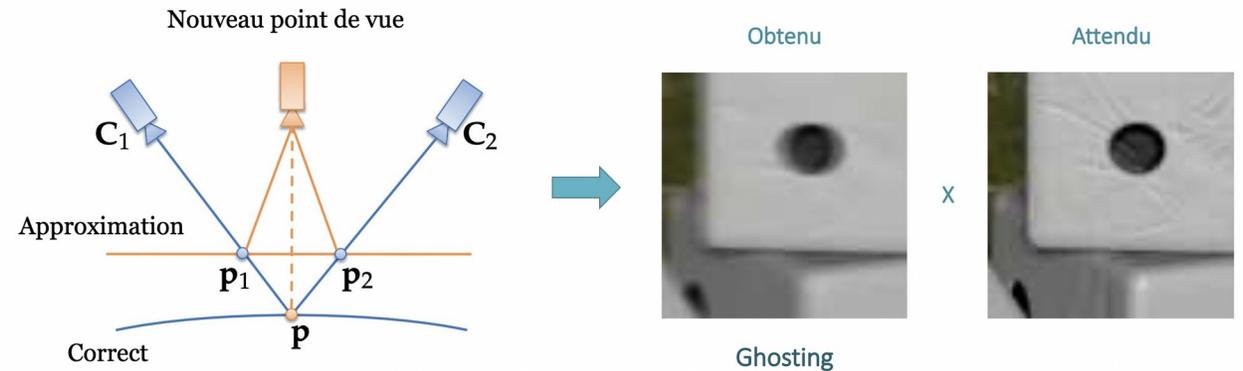
IBR avec incertitudes

Ambient Point Clouds for View Interpolation (Goesele, 2010)



Error-concealed Image-based Rendering (Eisemann 2011)

Bayesian View Synthesis and Image-Based Rendering Principles (Pujades 2014)



Projective Texturing Uncertain Geometry: silhouette-aware box-filtered blending using integral radial images (Brédif, 2014)



Cadre et contraintes

Données massives

- Accessibles en flux tuilés sur le web (WMTS, WFS, IIIF...)
- Limiter les prétraitements
- Mise en base et indexation des géoréférences/datations d'image (PostGIS)
 - Accès : Couche WFS de points 3D (=position des caméras historiques)

Accès web + prise en main facilitée

- Client léger desktop WebGL 

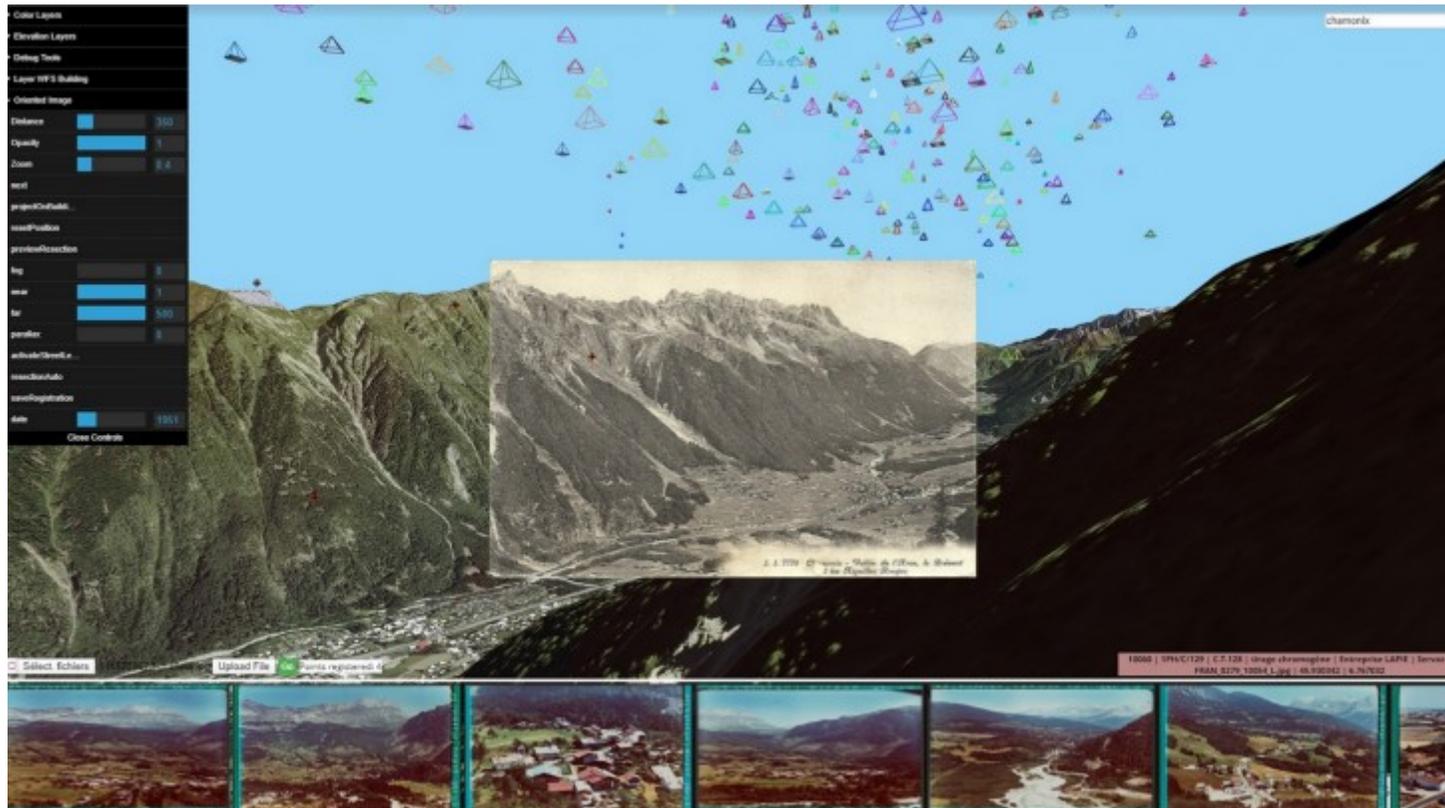
Passage à l'échelle du nombre de textures

Passage à l'échelle du rendu basé image et gestion de la diachronie et de l'hétérogénéité des données [Castro 2021]



Passage à l'échelle du nombre de textures

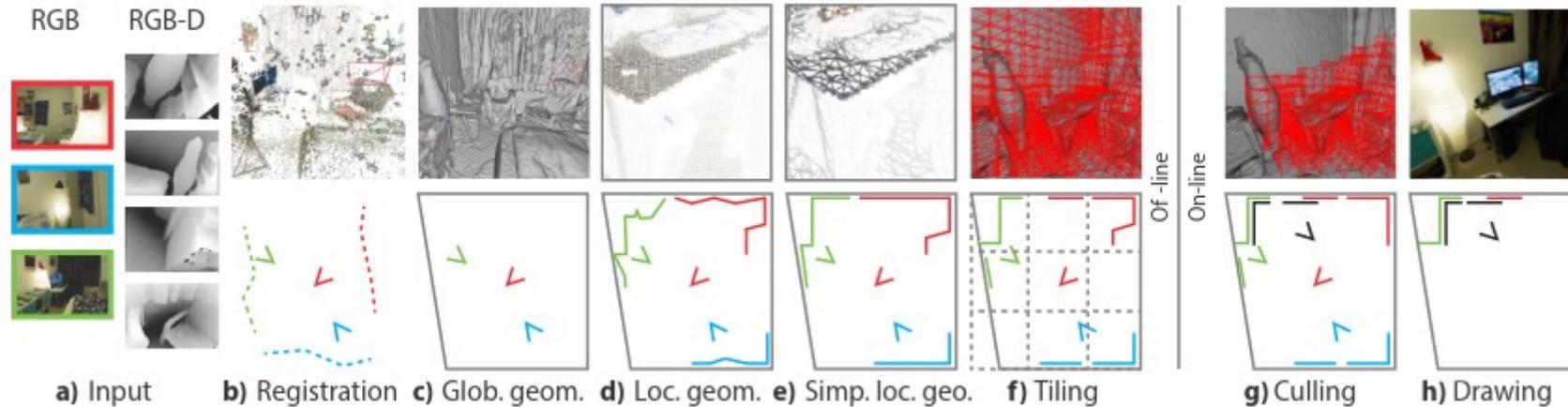
→ Peut on projeter 10, 100, 1000... textures simultanément **sans prétraitement** et **sans duplication** des géométries?



J. J. 7770 Chamonix - Vallée de l'Arve, le Brévent et les Aiguilles Rouges Jullien frères, Phot. Editeurs, Genève

Passage à l'échelle du nombre de textures

Scalable Inside-Out Image-Based Rendering, [Hedman 2016]



Calcul et rendu d'une scène 3D pour chaque image, tuilage

→ Comment repousser le nombre maximal de textures reprojétées simultanément **sans prétraitement** et **sans duplication** des géométries?

Limitations du nombre de textures

T = nombre de textures utilisées

Emplacements des textures utilisées	2T texture units		MAX_TEXTURE_IMAGE_UNITS (16)
Mises à jour des cartes de profondeur	T RenderToTexture		Realtime fps (16ms)
Texturation projective des Nuages de points [Devaux 2016]	9T varyings		MAX_VARYING_COMPONENTS (125)

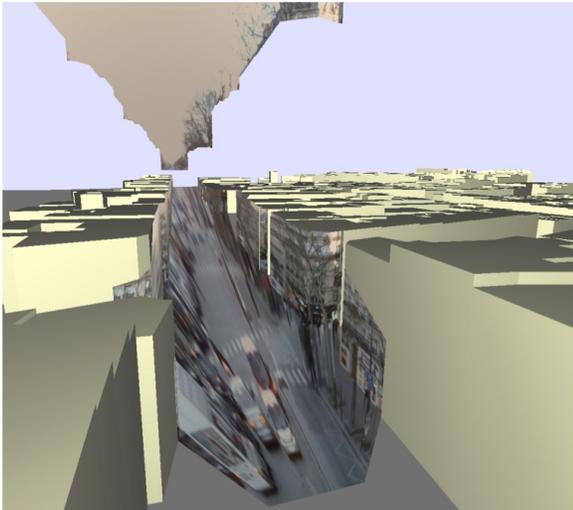
→ Limite à T=8 textures par rendering pass

Comment repousser ces limites ?

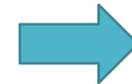
[Castro 2021]

Texturation projective avec shadowmap

- **Texturation projective** : coordonnées de texture recalculées à la volée pour chaque texture, et mélange (blending) des texels visibles
 - → indépendance de la géométrie et de la texturation (pas de précalcul)
- **Shadowmap** : calcul d'une depthmap de chaque texture à chaque modification de la scène (chargement/déchargement/changement de LOD...), pour ne texturer que la première surface visible



Sans shadow mapping



Avec shadow mapping

Texturation Multiple



- **Approche basique** : 2 samplers par texture projetée :
 - T textures RGB (photos)
 - T textures de profondeur (shadowmaps)
- **Approche proposée** : 2 texture2DArrays (RGB et shadowmap)
 - Indexation et mise à jour plus flexible (une texture est juste une couche du texture2DArray)
 - Légère perte de performance possible (à évaluer)

Emplacements des textures utilisées	2T	2	MAX_TEXTURE_IMAGE_UNITS (16)
--	----	---	---------------------------------

Mise à jour progressive des shadowmaps

- **Scène avec des données chargées dynamiquement (LOD, culling...)**
 - → mises à jour nécessaires des cartes de profondeur impactées (frustum culling)
 - Jusqu'à N renders-to-texture (RTT) par frame
- **Proposition: mise à jour progressive des shadowmaps**
 - Limiter à T_{max} le nombre de RTT par frame
 - priorisés par l'importance relative de la texture à recalculer sur le point de vue courant
 - (RTT dans une couche du texture2Darray de profondeur)



Mises à jour des cartes
de profondeur

T RTT

T_{max} RTT

Realtime fps (16ms)

Texturation projective des nuages de points

- Permet de bénéficier de la précision du nuage de point
- Par estimation d'une homographie de reprojection par point [Devaux 2016]

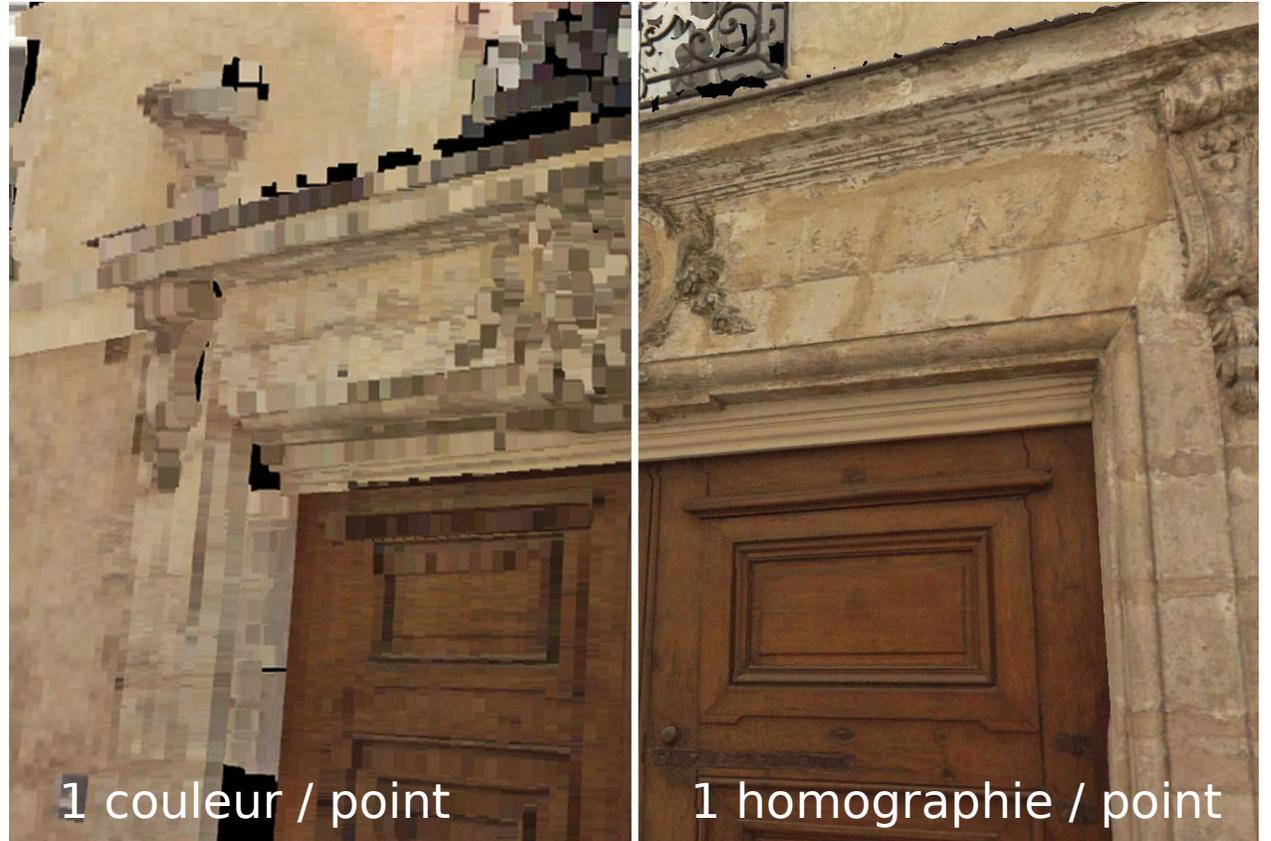
VS

$$H_i = M'_i \left(I + \frac{(C - C'_i)N^T}{N^T(P - C)} \right) M^{-1}$$

FS

$$p' = H_i p$$

→ 9T varyings H_i



Realtime projective multi-texturing of pointclouds and meshes for a realistic street-view web navigation, [Devaux 2016]

Texturation projective des nuages de points

- Réduction du nombre de varyings : linéaire (9T) → constant (3)

VS

$$H_i = M'_i \left(I + \frac{(\mathbf{C} - \mathbf{C}'_i) \mathbf{N}^T}{\mathbf{N}^T (\mathbf{P} - \mathbf{C})} \right) M^{-1}$$

FS

$$\mathbf{p}' = H_i \mathbf{p}$$

.js $\mathbf{E}_i = M'_i (\mathbf{C} - \mathbf{C}'_i)$

$$H'_i = M'_i M^{-1}$$

VS

$$\mathbf{n} = \frac{(M^{-1})^T \mathbf{N}}{\mathbf{N}^T (\mathbf{P} - \mathbf{C})}$$

FS

$$\mathbf{p}' = H'_i \mathbf{p} + (\mathbf{n}^T \cdot \mathbf{p}) \mathbf{E}_i$$

**Texturation projective
des Nuages de points
[Devaux 2016]**

9T varyings H_i

3 varyings \mathbf{n}

MAX_VARYING_COMPONENTS
(125)

Limitations du nombre de textures

T = nombre de textures utilisées

Emplacements des textures utilisées	2T texture units	2 texture units	MAX_TEXTURE_IMAGE_UNITS (16)
Mises à jour des cartes de profondeur	T RTT	T_max RTT	Realtime fps (16ms)
Texturation projective des Nuages de points [Devaux 2016]	9T varyings	3 varyings	MAX_VARYING_COMPONENTS (125)

Limitations restantes (= dépendances linéaires en T)

- Nombre d'uniforms (paramètres des caméras)
- Performance : nombre de Texture lookups (-> in-shader culling of textures ?)

En pratique : **T = 8 -> 40 textures simultanées**

[Castro 2021]

Comment traiter la diachronie et l'hétérogénéité des données ?

- Données lidar très précises, **actuelles**, objets fixes et **mobiles**
→ donc incohérentes avec les photographies historiques

Les points des arbres cachent les façades et le ciel (shadowmap)



Les points des arbres sont texturés avec des pixels de façade.

Comment traiter la diachronie et l'hétérogénéité des données ?

- Données lidar très précises, actuelles, objets fixes et mobiles
 - Modèle 3D de ville : moins précis, mais objets fixes uniquement
- **Classification à la volée** des points sur la distance au modèle 3D :
- Proche = détail d'un objet fixe, à visualiser (texturation projective)
 - Éloigné = objet mobile, ne pas visualiser



Comment traiter la diachronie et l'hétérogénéité des données ?

- **Hypothèse** : la géométrie des bâtiments et du sol n'a pas évolué et la géométrie actuelle est donc valide au temps de la photographie historique
- Classification à la volée : test de shadowmapping sur le sol+bâti, sans les points



$$\begin{array}{ccc} dist \leq minDist + \epsilon & \longrightarrow & minDist - \epsilon \leq dist \leq minDist + \epsilon \\ \text{Soft shadowmap} & & \text{Interval shadowmap} \end{array}$$

Comment traiter la diachronie et l'hétérogénéité des données ?

- **Limitations**

- Objets mobiles sur photos historiques
→ projetés sur le sol+bâti
- Objets fixes actuels non présents sur les photos historiques
→ occlusions, mauvaises texturations.

- **Pistes d'amélioration**

- Segmentation sémantique du lidar et des images ?
En prétraitement ?
- Utilisation d'un modèle 4D de ville ?



Démo Vidéo

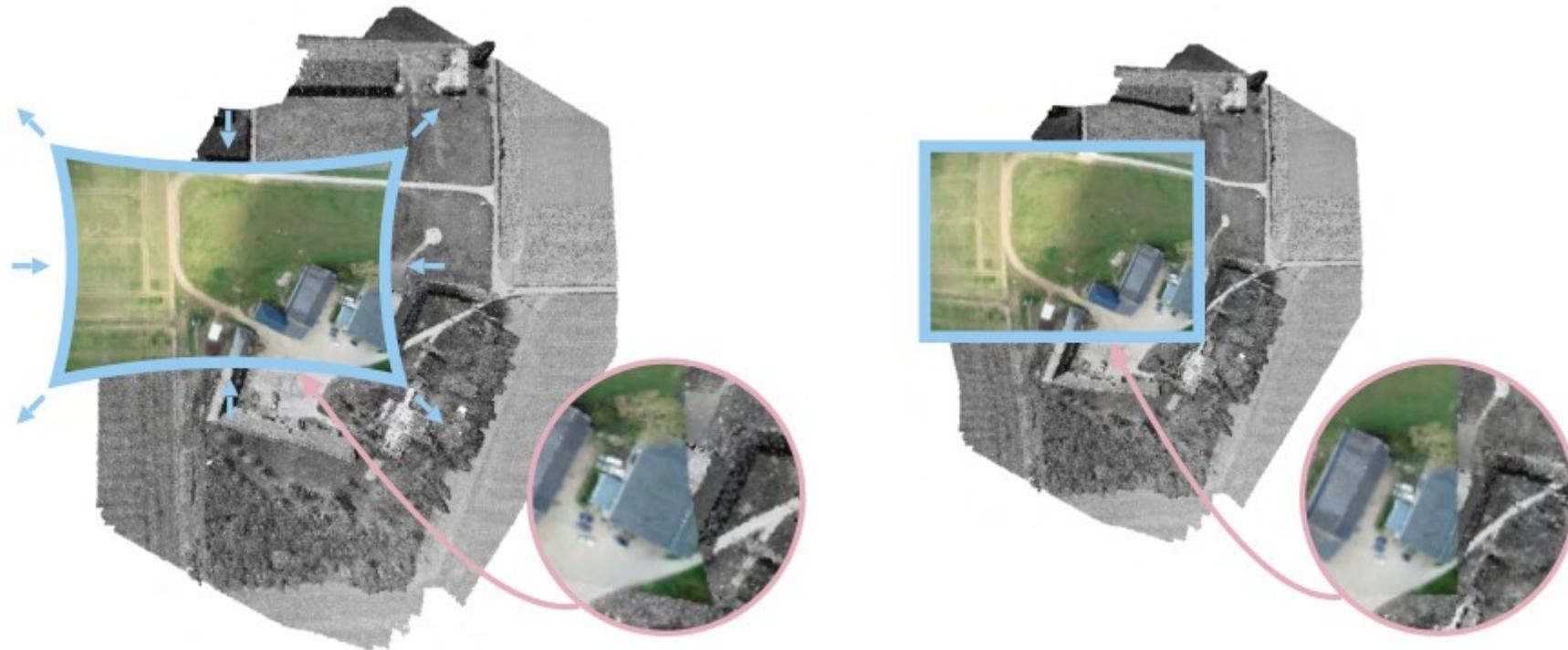
Résultats sur Chalon

Fonds Charles Gros (Musée Nicéphore Niépce)



Prise en compte de la distorsion géométrique des images

Geometric Distortion of Historical Images for 3D Visualization [Paiz-Reyes 2020]



**Correction de la distorsion
à la volée**

**+ Visualisation à travers la caméra
dézoomée avec distorsion**

Prise en compte de la distorsion géométrique des images



HistoVIS

<https://epaizreyes.github.io/histovis/examples/multiple.html>

Application interactive de géovisualisation permettant de naviguer et d'explorer ces données

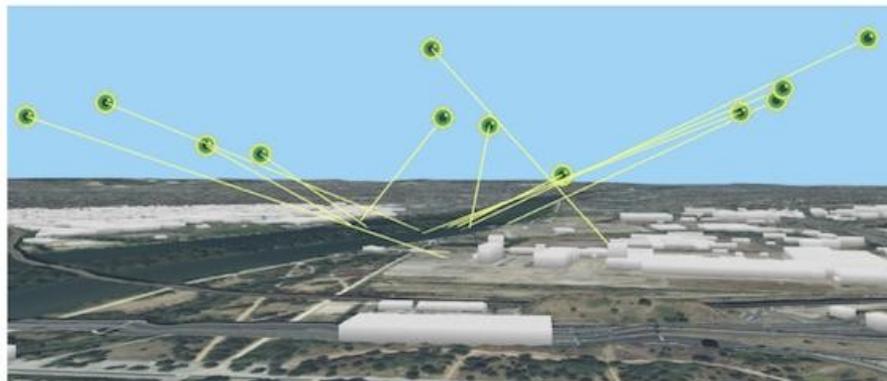
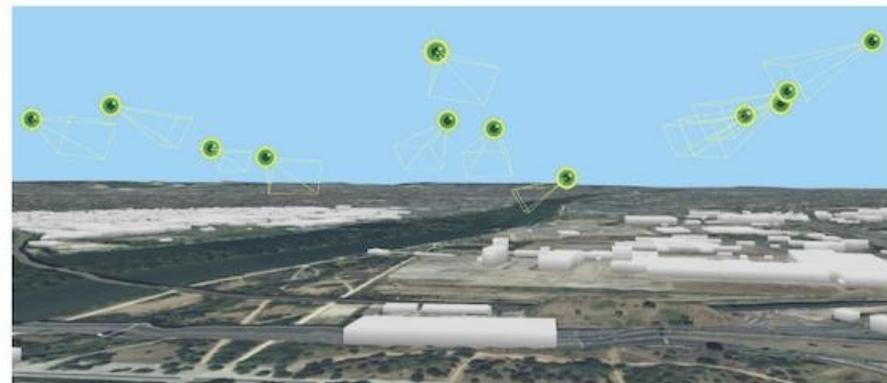
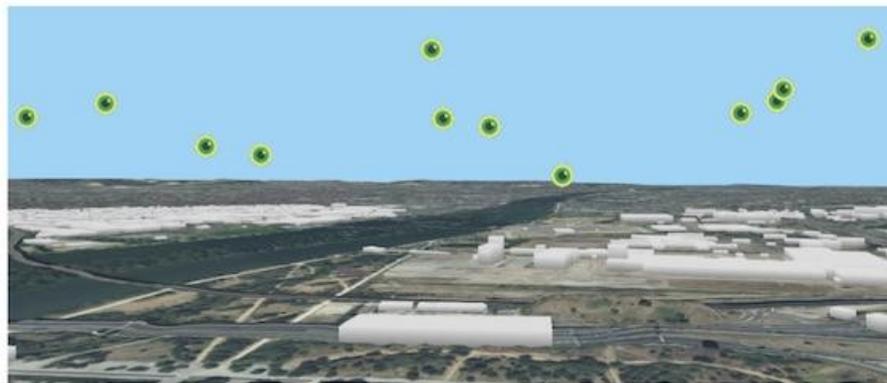
Cluttering Reduction for Interactive Navigation and Visualization of Historical Images [Paiz-Reyes ICC 2021]



HistoVIS

**Représentation
individuelle
des prises de vue :**

- un icône et/ou
- une géométrie

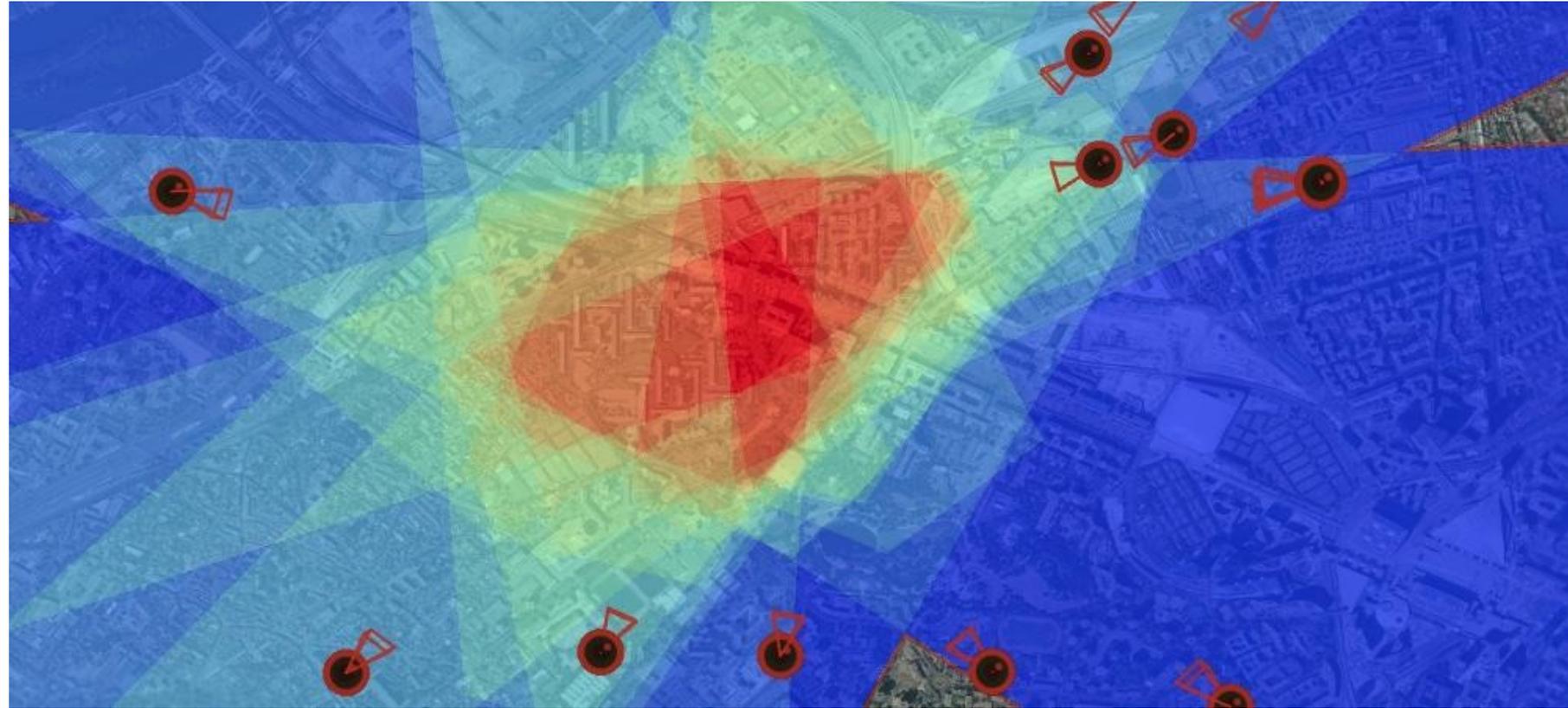


HistoVIS

**Représentation
agrégée
des prises de vue :**
- **carte de chaleur**

→ **Densité de la
couverture
cartographique**

→ **Qu'est ce qui est
vu ?**



HistoVIS

**Aperçus placés
autour de la vue 3D
des images visibles**

**Regroupés par
clusters dynamiques
de similarité spatiale**

**Position du cluster à
au plus près de la
projection de sa
position dans la vue
3D**

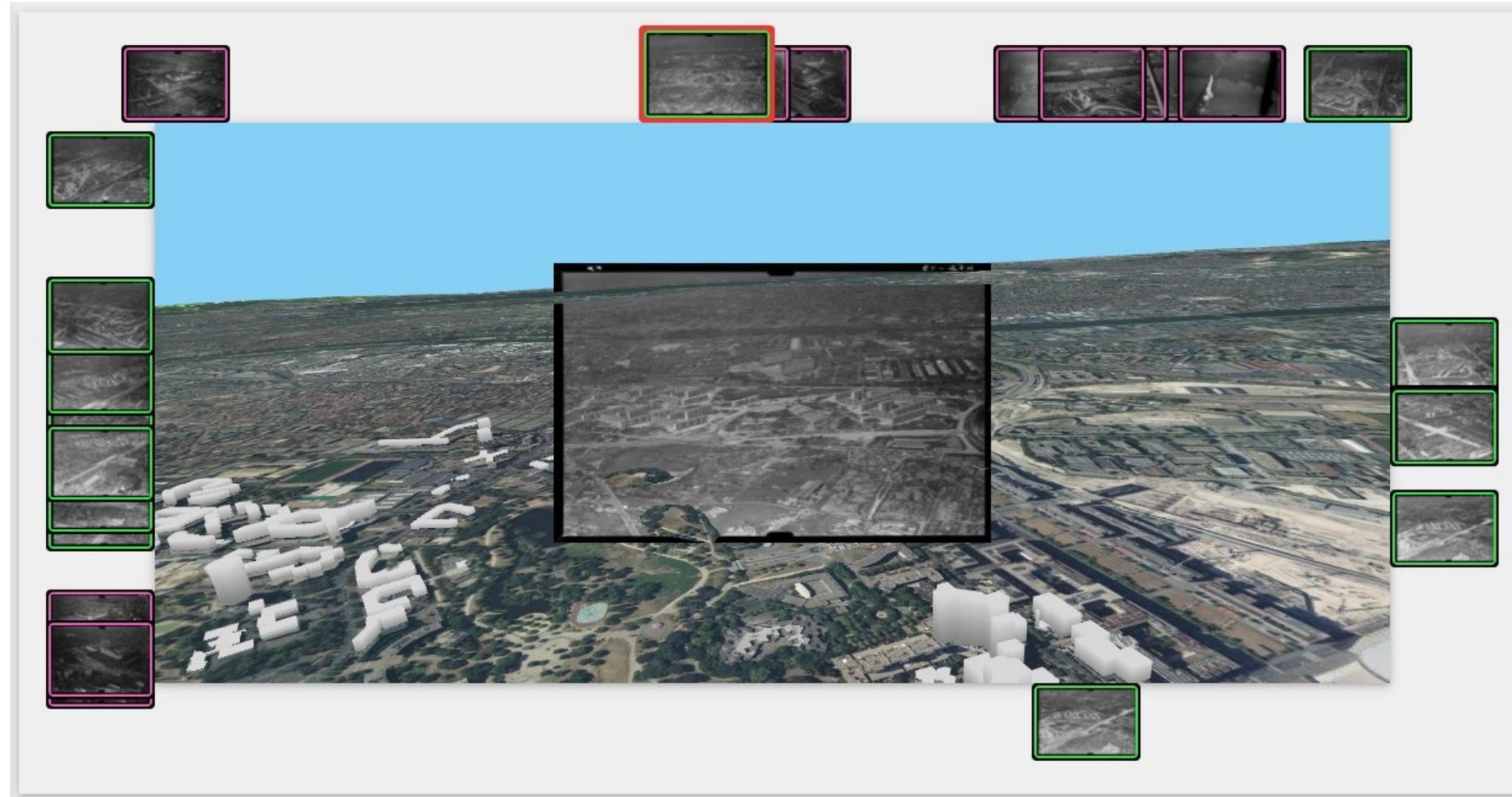


HistoVIS

**Aperçus placés
autour de la vue 3D
des images visibles**

**Regroupés par
clusters dynamiques
de similarité spatiale**

**Position du cluster à
au plus près de la
projection de sa
position dans la vue
3D**



HistoVIS

Bordure des images projetées :

- Fondu le plus continu possible entre les images projetées

ou

- Accentuation des bords de chaque image

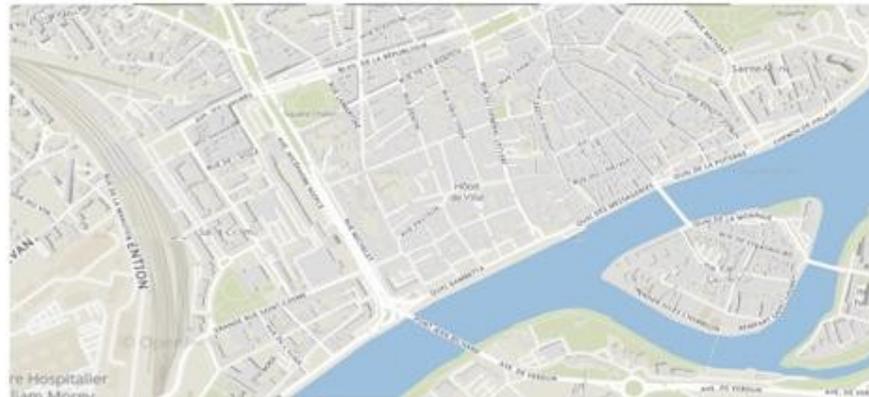


HistoVIS

En l'absence de pixels d'images historiques :

→ Utilisation de la couleur d'une couche actuelle

→ Stylisation
- abstraite/réaliste
- homogénéisation possible du style avec les images historiques (sépia, noir et blanc...)



HistoVIS : navigation interactive

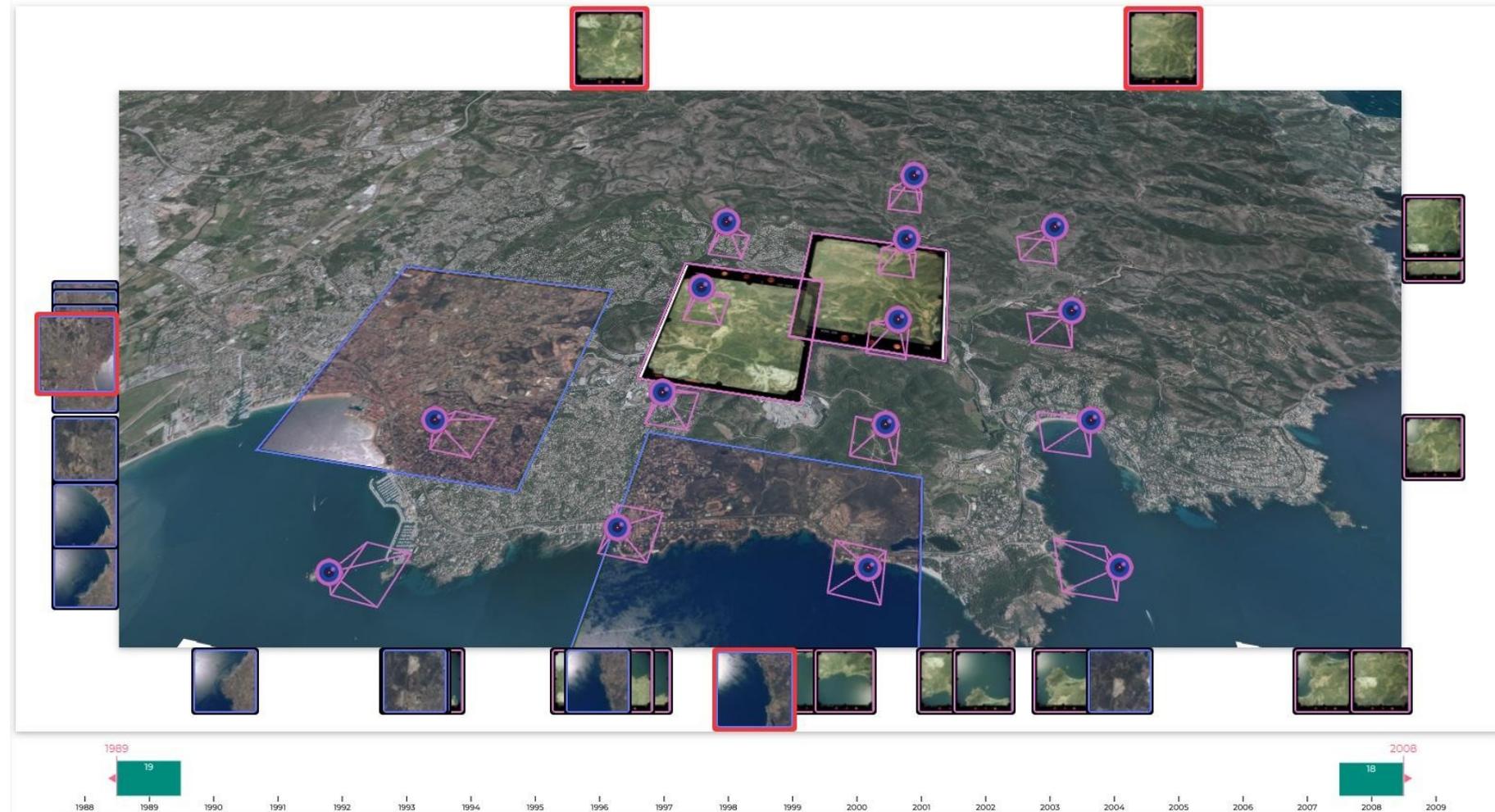
Interactions :

- marqueurs
- aperçus des clusters
- la frise temporelle

→ pour projeter ou non les images dans la scène

→ pour déplacer le point de vue

→ filtrage temporel des images visibles



HistoVIS

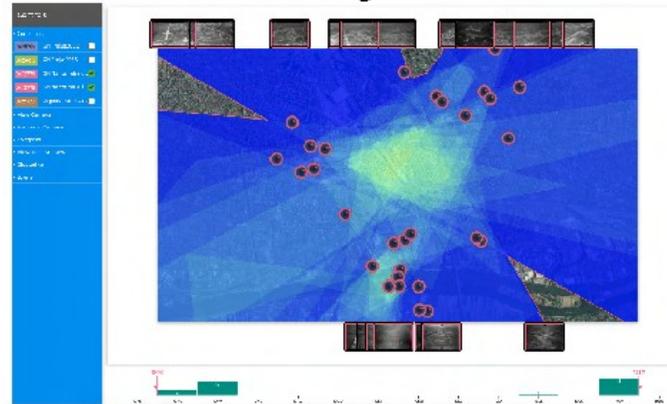
Exemple d'usage pour l'exploration de données

Photos non spatialisées ?

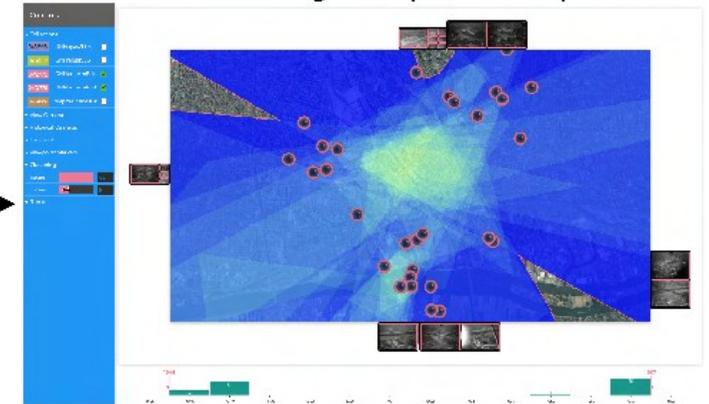
→ découvertes par similarité de contenu image (ou de métadonnées)

[Gominski 21]

Task 1: loading the dataset



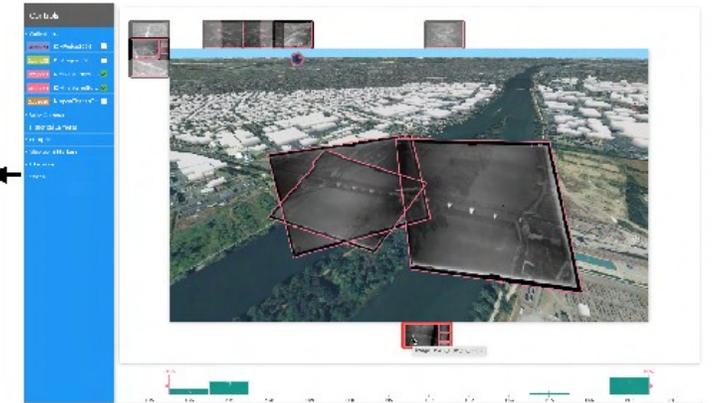
Task 2: clustering the spatialized photos



Task 4: homogenizing the context



Task 3: visualizing a set of photographs



Task 5: browsing non-spatialized photos



Task 6: visualizing another area



Conclusion

- Développement d'un démonstrateur, HISTOVIS permettant une navigation immersive et une exploration interactive de fonds photographiques
<https://epaizreyes.github.io/histovis/examples/multiple.html>
- Premiers résultats sur la gestion de la diachronie et de l'incertitude dans le rendu de données hétérogènes massives



Perspectives

- **Diachronie**

- Modèle 3D+T de la ville, datation (au pixel?) des photographies

- **Modélisation et rendu**

- Incertitudes du modèle 3D (imprécision, incomplétude...). Prise en compte du LOD
- Incertitudes du géoréférencement des photographies
- Contrôle du style des images reprojctées et mélangées : homogénéisation, différenciation ?

- **Passage à l'échelle :**

- Combiner notre approche et Inside-Out IBR [Hedman 16]
- Virtualisation des Textures (Mégatexture) et culling des textures sur le GPU

- **Réalité Augmentée, nouveaux usages**

Merci !



<http://www.itowns-project.org/>



<http://alegoria.ign.fr/>