Graph-based compression of omnidirectional images

Thomas Maugey
thomas.maugey@inria.fr

Workshop EPFL-Inria
January 2020

Context

Associate team

- Title: Graph-based Omnidirectional image Processing (GOP)
- Date: Jan 2017- Dec 2019
- Partners: LTS4, EPFL (Pascal Frossard) and SIROCCO, Inria (Thomas Maugey)

Labex Cominlabs project

- Title: Interactive communication (Intercom)
- Date: Jan 2017- Dec 2020
- Partners: IMT Atlantique, L2S, Inria

Rennes Metropole and InriaHub

- Title: Multi-omnidirectional capture
- Date: Jan 2015- Dec 2018

Outline

Graph-based compression of omnidirectional images

Graph-based compression of omnidirectional images

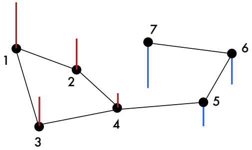
Omnidirectional image definition

Outline

Graph-based compression of omnidirectional images

Omnidirectional image definition

Graph signal processing interest

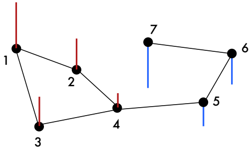


Outline

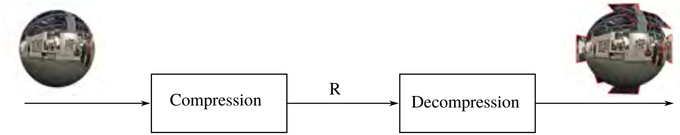
Graph-based compression of omnidirectional images

Omnidirectional image definition

Graph signal processing interest



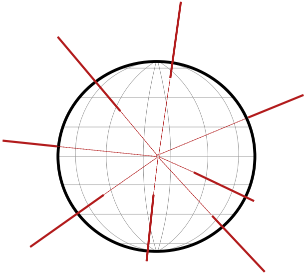
Compression schemes design



Omnidirectional image definition

Definition

An image that represents the light activity arriving to a point (the image center) from every direction (360^o field of view).



Applications

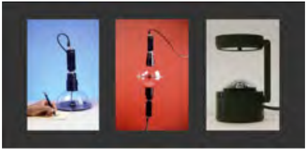
- Virtual reality,
  e.g., Head-Mounted Display (HDM)

- Free viewpoint Navigation
- Robotics

Omnidirectional capture

- **Multi-camera** capture (180^o or 360^o field of view)

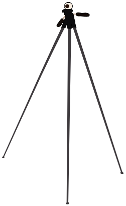
- **Catadioptric** system (180^o field of view)



- **Fish-eye** lenses (180^o field of view)

Multi omnidirectional capture for Free Viewpoint

A omnidirectional camera captures the light ray coming from all directions.



T. Maugey , L. Guillo, C. Le Cam, FTV360: a Multiview 360-degree Video Dataset with Calibration Parameters, ACM Multimedia Systems Conference, Amherst, MA, US, June 2019.
Publicly available at https://project.inria.fr/ftv360

Multi omnidirectional capture for Free Viewpoint

A omnidirectional camera captures the light ray coming from all directions.

At a given position δ, it enables any

$$\mathbf{r}(\delta) \in [-\pi/2, \pi/2] \times [-\pi, \pi] \times [-\pi, \pi]$$

We position many cameras at position δ_i, it enables any

$$\mathbf{t} = \delta_i \text{ and } \mathbf{r}(\delta_i) \in [-\pi/2, \pi/2] \times [-\pi, \pi] \times [-\pi, \pi]$$

T. Maugey , L. Guillo, C. Le Cam, FTV360: a Multiview 360-degree Video Dataset with Calibration Parameters, ACM Multimedia Systems Conference, Amherst, MA, US, June 2019. Publicly available at https://project.inria.fr/ftv360

Multi omnidirectional capture for Free Viewpoint

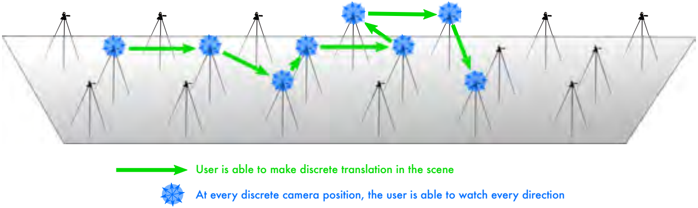
A omnidirectional camera captures the light ray coming from all directions.

At a given position δ, it enables any

$$\mathbf{r}(\delta) \in [-\pi/2, \pi/2] \times [-\pi, \pi] \times [-\pi, \pi]$$

We position many cameras at position δ_i, it enables any

$$\mathbf{t} = \delta_i \text{ and } \mathbf{r}(\delta_i) \in [-\pi/2, \pi/2] \times [-\pi, \pi] \times [-\pi, \pi]$$



T. Maugey , L. Guillo, C. Le Cam, FTV360: a Multiview 360-degree Video Dataset with Calibration Parameters, ACM Multimedia Systems Conference, Amherst, MA, US, June 2019. Publicly available at https://project.inria.fr/ftv360

Omnidirectional image representation

The captured image is not directly exploitable

Fish-eye capture

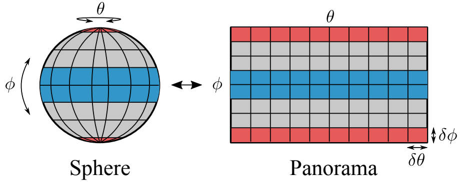
Catadioptric capture

Some post-processing are needed (*e.g.*, stitching) $\rightarrow$ spherical image

The spherical image is then generally mapped into a 2D planar image

Equirectangular representation

Equirectangular or Panorama description



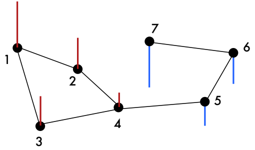
- Most popular
- Suitable for image processing applications

But

- Radial distortions

Basics on Graph Signal Processing

A signal $\mathbf{x}$ defined on a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$



$v_i \in \mathcal{V}$ = pixel position
$e_{ij} \in \mathcal{E}$ and $w_{ij} \in \mathcal{W}$ = neighborhood information
x_i = the pixel color

- Adjacency matrix $\mathbf{A}$

$$a_{ij} = \begin{cases} w_{ij} \text{ if } e_{ij} \in \mathcal{E} \\ 0 \text{ otherwise} \end{cases}$$

- Degree matrix $\mathbf{D}$

$$d_{ij} = \begin{cases} \mathrm{degree}(v_i) \text{ if } i = j \\ 0 \text{ otherwise} \end{cases}$$

- Laplacian matrix $\mathbf{L}$

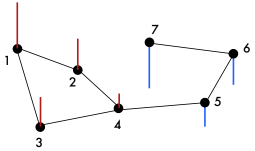
$$\mathbf{L} = \mathbf{D} - \mathbf{A}$$

Shuman, D. I., Narang, S. K., Frossard, P., Ortega, A., and Vandergheynst, P. (2013). The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. IEEE signal processing magazine, 30(3), 83-98.

Basics on Graph Signal Processing

Laplacian measures locally "how much" a pixel is different from the weighted average of the neighbors:

$$\mathbf{L}\mathbf{x} = \left(d_i x_i - \sum_{j \in \text{Neighborhood}} w_{ij} x_j \right)_i$$



It measures the variation of a signal $\mathbf{x}$ on a graph $\mathcal{G}$.

- Laplacian eigenvalues $\lambda_i =$ “frequency" or total variation
- Laplacian eigenvectors $\mathbf{u}_i =$ transform basis

Some interpretations

- two pixels are connected if they are able to help to predict each other
- the corresponding weight captures "how much" they contribute to the prediction

Shuman, D. I., Narang, S. K., Frossard, P., Ortega, A., and Vandergheynst, P. (2013). The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. IEEE signal processing magazine, 30(3), 83-98.

Graph spectral compression

At the encoder:

1) Graph transform

$$\mathbf{c} = \mathbf{U}^\top \mathbf{x}$$

2) Quantization

$$\hat{\mathbf{c}} = q(\mathbf{c})$$

3) Entropy coding

$$i = f(\hat{\mathbf{c}}) \quad \text{with} \quad \mathcal{R} = |i|$$

R is closely related to $||\hat{\mathbf{c}}||_0$

At the decoder:

4) Entropy decoding

$$\hat{\mathbf{c}} = f^{-1}(i)$$

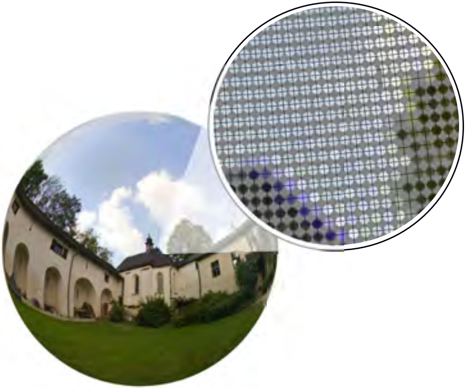
5) Inverse transform

$$\hat{\mathbf{x}} = \mathbf{U}\hat{\mathbf{c}} \quad \text{with} \quad \mathcal{D} = ||\mathbf{x} - \hat{\mathbf{x}}||_2^2$$

Compression tradeoff $\min \ \mathcal{D} + \gamma\mathcal{R}$

GSP interest #1

GSP enables to embed the physical distance into the graph weights



Hypothesis: pixels that are close on the sphere are more correlated

$$w_{ij} = \exp\left(-\frac{d_{ij}^2}{2\sigma}\right)$$

Coder 1: Partition on graph

New Hypothesis: pixels that are close on the sphere are more correlated ***but sometimes not***

$$w_{ij} = \exp\left(-\frac{d_{ij}^2}{2\sigma}\right) \quad \text{or} \quad w_{ij} = 0$$

We cut the graph such that

$$\begin{aligned} \min_{\tilde{\mathcal{G}}=\{\mathcal{G}_i\}} \quad & \mathcal{D}(\tilde{\mathcal{G}}) + \gamma(\mathcal{R}_C(\tilde{\mathcal{G}}) + \mathcal{R}_B(\tilde{\mathcal{G}})) \\ \text{subject to} \quad & N_i < N_{max}, \; \forall i \end{aligned} \tag{1}$$

M. Rizkallah, F. De Simone, T. Maugey, C. Guillemot, P. Frossard, Rate Distortion Optimized Graph Partitioning for Omnidirectional Image Coding EUSIPCO, Athens, Greece, Sept. 2018. Best student paper

Coder 1: Partition on graph

New Hypothesis: pixels that are close on the sphere are more correlated ***but sometimes not***

$$w_{ij} = \exp\left(-\frac{d_{ij}^2}{2\sigma}\right) \quad \text{or} \quad w_{ij} = 0$$

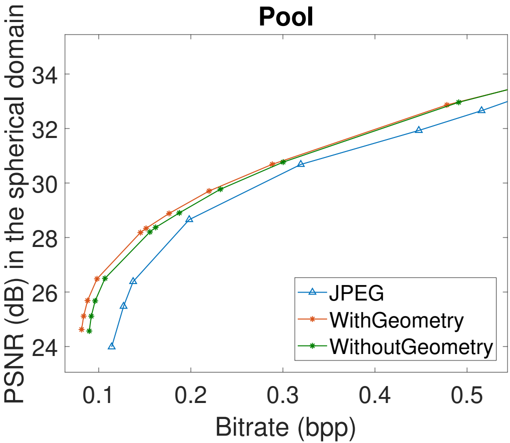
We cut the graph such that

$$\begin{aligned} \min_{\tilde{\mathcal{G}}=\{\mathcal{G}_i\}} \quad & \mathcal{D}(\tilde{\mathcal{G}}) + \gamma(\mathcal{R}_C(\tilde{\mathcal{G}}) + \mathcal{R}_B(\tilde{\mathcal{G}})) \\ \text{subject to} \quad & N_i < N_{max}, \; \forall i \end{aligned} \qquad (1)$$

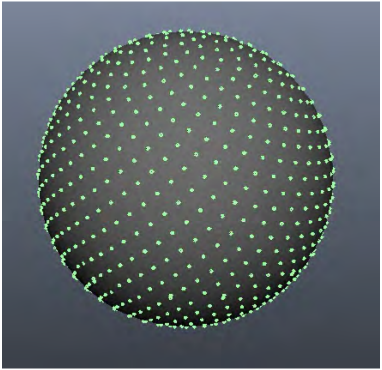
$\longrightarrow$

M. Rizkallah, F. De Simone, T. Maugey, C. Guillemot, P. Frossard, Rate Distortion Optimized Graph Partitioning for Omnidirectional Image Coding EUSIPCO, Athens, Greece, Sept. 2018. Best student paper

Coder 1: results



Uniform sampling



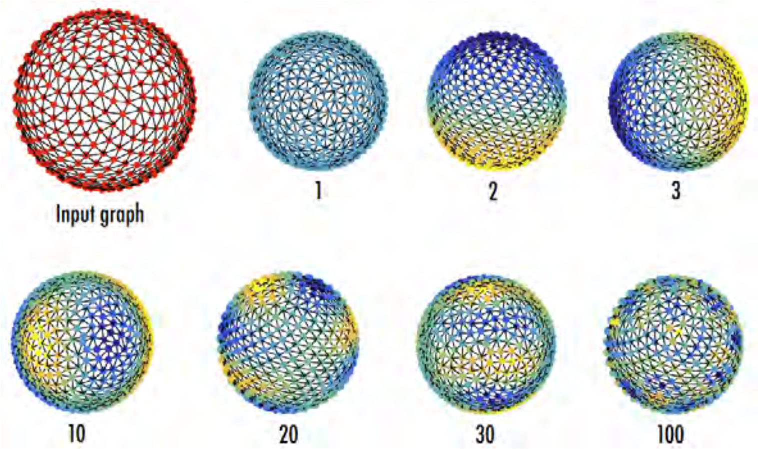
- Equidistant point
- Connectivity preserved

But

- Not a 2D image anymore

GSP interest #2

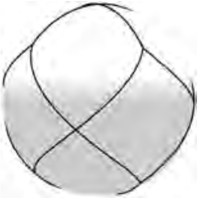
GSP enables to adapt to new topologies

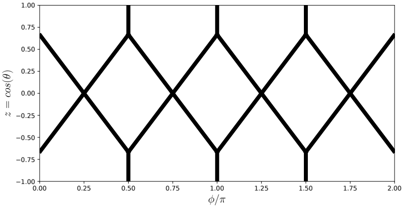


Some graph transform vectors $\mathbf{u}_i$

Coder 2: healpix-based

Healpix sampling



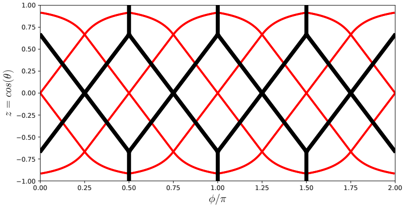


cylindrical projection

N. Mahmoudian Bidgoli, T. Maugey , A. Roumy, Intra-coding of 360-degree images on the sphere Picture Coding Symposium (PCS), Ningbo, China, Nov. 2019 And extension in collaboration with R. Azevedo

Coder 2: healpix-based

Healpix sampling

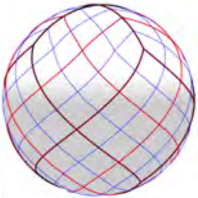


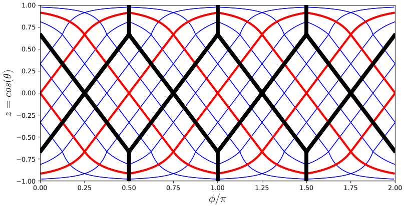
cylindrical projection

N. Mahmoudian Bidgoli, T. Maugey , A. Roumy, Intra-coding of 360-degree images on the sphere Picture Coding Symposium (PCS), Ningbo, China, Nov. 2019 And extension in collaboration with R. Azevedo

Coder 2: healpix-based

Healpix sampling



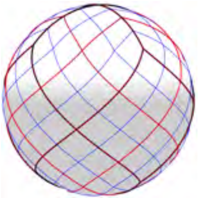


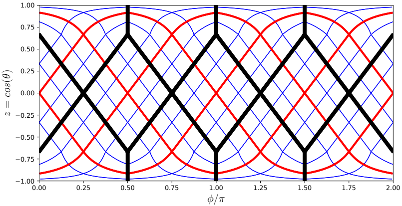
cylindrical projection

N. Mahmoudian Bidgoli, T. Maugey , A. Roumy, Intra-coding of 360-degree images on the sphere Picture Coding Symposium (PCS), Ningbo, China, Nov. 2019 And extension in collaboration with R. Azevedo

Coder 2: healpix-based

Healpix sampling





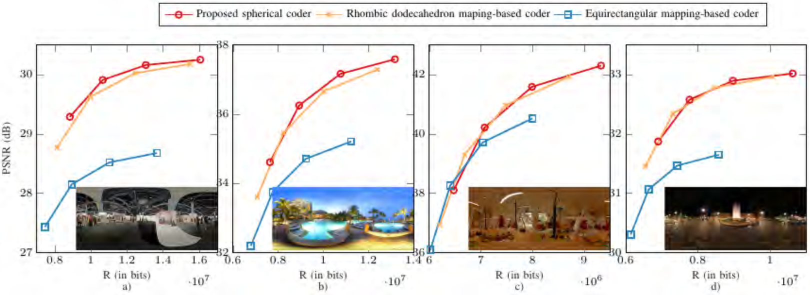
cylindrical projection

Spherical blocks

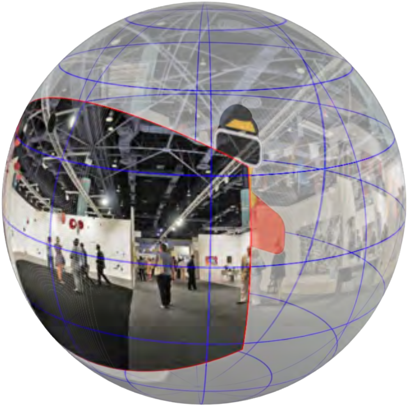
Enable to define Spherical blocks, and thus extend the classical coding tools such as **prediction**

N. Mahmoudian Bidgoli, T. Maugey , A. Roumy, Intra-coding of 360-degree images on the sphere Picture Coding Symposium (PCS), Ningbo, China, Nov. 2019 And extension in collaboration with R. Azevedo

Coder 2: results

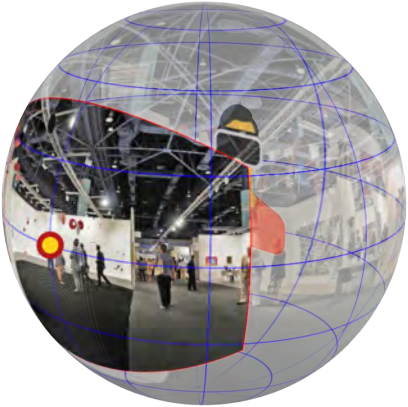


Coder 3: Random access



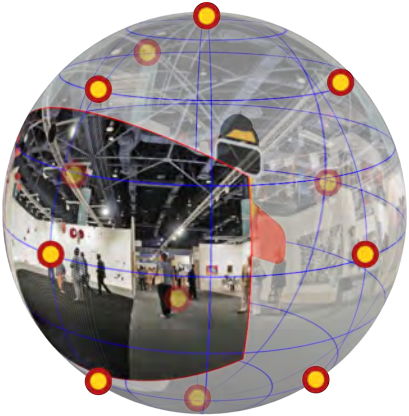
Only a small image portion is visible at a given time

Coder 3: Random access



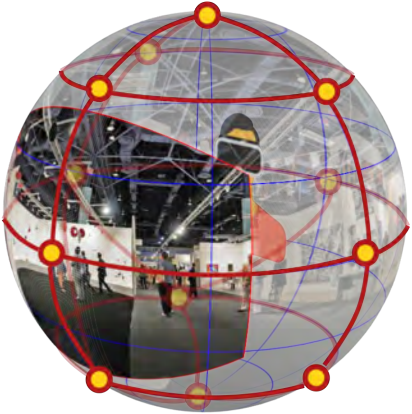
a vertex = a direction

Coder 3: Random access



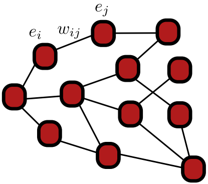
the vertices = the achievable directions

Coder 3: Random access



the edges = the possible transitions

Coder 3: Random access



The **navigation graph**, where the weights correspond to the coding rate needed to make the view transition $w_{ij} = \mathcal{R}_{ij}$.

Classical compression: for each position i,

$$\text{Storage: } \mathcal{S} = \sum_j \mathcal{R}_{ij} \quad \text{Transmission rate: } \mathcal{R} = \mathcal{R}_{ij}$$

$$\text{or} \quad \text{Storage: } \mathcal{S} << \sum_j \mathcal{R}_{ij} \quad \text{Transmission rate: } \mathcal{R} >> \mathcal{R}_{ij}$$

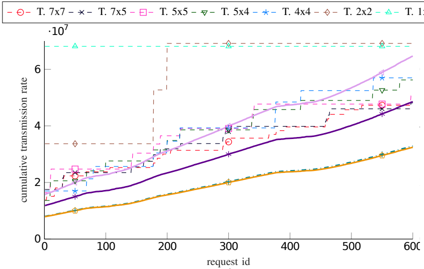
Our incremental coding solution: for each position i,

$$\text{Storage: } \mathcal{S} = \max_j \mathcal{R}_{ij} \quad \text{Transmission rate: } \mathcal{R} = \mathcal{R}_{ij}$$

E. Dupraz, T. Maugey, A. Roumy, M. Kieffer, Rate-Storage Regions for Extractable Source Coding with Side Information, in Physical Communication, Elsevier, Vol. 37 2019.

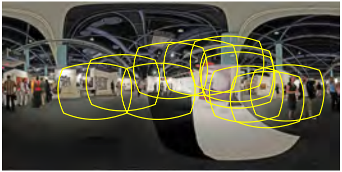
Storage cost Much lower than exhaustive storage
Reasonable overhead with respect to the “no random access" case

Transmission cost over time



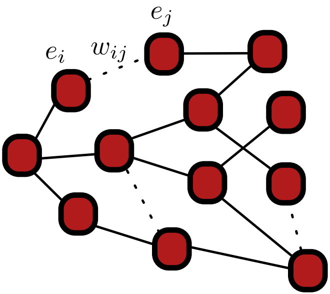
a)

One example of client's request over time



N. Mahmoudian-Bidgoli, T. Maugey, A. Roumy, Interactive compression of 360-degree images, submitted to IEEE TMM.

Ongoing work



Graph optimization:

$$\min_{\mathcal{G}} \quad \mathcal{S} + \gamma \mathcal{R}$$

M. Pham, A. Roumy, T. Maugey, E. Dupraz, P. Frossard, Graph optimization for multi-source compression with random access, in preparation.

Conclusion

Omnidirectional images greatly benefit from Graph signal processing theory, especially for compression tasks.

Future directions:

- Extension to temporal aspects
- Include learning tools, such as DeepSphere for compression

Thanks, questions?