

# Video Frame Interpolation

Étude de cas

John Robinson

# Contenu de cette presentation

- Comprendre le problème
- Revue d'articles
  - Deep bayesian video frame interpolation
  - Exploring Motion Ambiguity and Alignment for High-Quality Video Frame Interpolation
  - IFRNet: Intermediate Feature Refine Network for Efficient Frame Interpolation
  - Uncertainty-Guided Spatial Pruning Architecture for Efficient Frame Interpolation
  - Clearer Frames, Anytime: Resolving Velocity Ambiguity in Video Frame Interpolation
- Que retenir de ces recherches ?
- Conclusion
- Prochaine étape

# Interpolation d'images

# Le problème

En se basant sur une série d'images

$$\mathcal{I} = \{I_{-k}, \dots, I_0, I_1, \dots, I_k\}$$

Construire un modèle  $\mathcal{F}$  capable de générer une image intermédiaire.

$$I_t = \mathcal{F}(\mathcal{I}, t), \quad 0 < t < 1$$

$k$  paramétrise le modèle et le training set,

$$\mathcal{D} = \left\{ \bigcup_{l=1}^k I_{i \pm l}, I_i \right\}_{i=k}^{N-k}$$

- $k = 1$ , triplets
- $k = 2$ , quintuplets
- $k = 3$ , septuplets

Le **deep learning** nous permet d'approcher ce problème de **regression**.

## Regression "pure"

Le modele  $\mathcal{F}$  tente de capturer la relation directe entre l'output  $I_t$  et les images adjacentes dans le dataset

Cette formulation offre peu de flexibilité,  $t = 0.5$

# Optic Flow

On considère ici une étape intermédiaire, celle de l'optic flow, qui caractérise le mouvement apparent de la scène.

La première étape consiste en l'estimation d'un certain nombre de d'optic flows, souvent 2.

$$\phi = \{F_{i \rightarrow t}\}_i^K$$

où

$$F_{i \rightarrow t} \approx g(\mathcal{I}, t)$$

Le modele  $\mathcal{F}$  interpole donc en fonction des images et des flows.

$$I_t = \mathcal{F}(\mathcal{I}, \phi)$$

Cette approche permet une plus grande flexibilité quant a  $t$ . Cependant, approximer l'optic flow en ne se basant que sur les images reste imprécis (problème d'occlusion, etc...).

## Solution Actuelle, Modèle

Le modele actuel execute une régression "pure" et se base sur CAIN

## Solution Actuelle, Challenges

Difficultés avec les **paternes répétitifs** et les situations **d'ambiguïté**.

Cette solution est donc clairement perfectible.



# Métriques et Evaluation

De nombreuses métriques telles que le PSNR et la SSIM

- Peak Signal-to-Noise Ratio (dB)

$$PSNR(I_1, I_2) = 10 \log_{10} \left( \frac{MAX^2(I_1)}{MSE(I_1, I_2)} \right)$$

Compare la qualité de  $I_2$  par rapport à  $I_1$

- Structural Similarity (entre 0 et 1)

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2\mu_y^2 + c_1)(\sigma_x^2\sigma_y^2 + c_1)}$$

Compare la structure de l'image  $x$  à celle de l'image  $y$ .

Ces métriques n'expliquent pas tous les aspects de la qualité d'une image, l'analyse qualitative reste donc de vigueur.

## **This is a sub-title 1-2-3**

dazeiojfpoajfpozajfopazjfzaojfapzofjzpj

This slide is centered vertically with `class: middle`.

# Some maths

$$y = mx + b$$

# Some code

```
def fibonacci(n):  
    if n <= 1:  
        return n  
    else:  
        return fibonacci(n-1) + fibonacci(n-2)
```

# References

- Liste des articles
  - Yu, Zhiyang, & al. "Deep bayesian video frame interpolation." Oct 2022.
  - Choi, Kim, & al. "Channel Attention Is All You Need for Video Frame Interpolation" 2020.
  - Zhou, Li, & al. "Exploring Motion Ambiguity and Alignment for High-Quality Video Frame Interpolation" Mar 2022
  - Kong, Jiang, & al. "IFRNet: Intermediate Feature Refine Network for Efficient Frame Interpolation" May 2022
  - Cheng, Jiang, & al. "Uncertainty-Guided Spatial Pruning Architecture for Efficient Frame Interpolation" Oct 2023
  - Zhong, Krishnan, & al. "Clearer Frames, Anytime: Resolving Velocity Ambiguity in Video Frame Interpolation" Nov 2023

- Autres références

