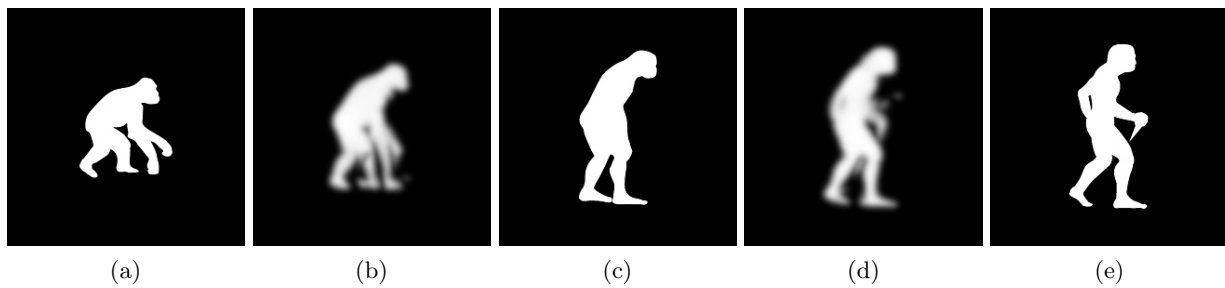


Rapport du TP2 du module de Recherche

Transport optimal

Guillaume DURET (p2021346)

Figure 1: Affichage de l'évolution de l'homme avec une interpolation entre les images d'origine



10 Décembre 2020

1 Introduction

Ce TP a pour but d'utiliser la théorie du transport optimal pour interpoler des images et des formes. Pour cela il est implémenté l'algorithme de Sinkhorn qui permet d'interpoler à l'aide des barycentres de Wasserstein

2 Théorie

Le problème posé de Kantorovich pour le transport optimal est le problème défini dans l'équation 1 avec C une matrice de Cout (ce sera une métrique L2 pour le TP) et M le plan de transformation que l'on cherche à estimer sachant que M appartient à $U(r, c)$ (définition 2) qui permet de vérifier la conservation de la masse.

$$\min_{M \in U(r, c)} \langle C, M \rangle \quad (1)$$

$$U(r, c) = \{X \in M_+^{m \times n}; X \mathbf{1} = r \text{ et } X^T \mathbf{1} = c\} \quad (2)$$

De plus l'idée de Cuturi a été d'ajouter un terme de régularisation entropique (équation 4) pour que le problème devienne la formule 3

$$W = \min_{M \in U(r, c)} \langle C, M \rangle - \epsilon E(M) \quad (3)$$

$$E(M) = - \sum M_{ij} (\log(M_{ij}) - 1) \quad (4)$$

Cette formule est appelée distance régularisée de Wasserstein, dans notre cas le but est de pouvoir interpoler des images avec le barycentre régularisé de Wasserstein qui est défini comme n'importe quel barycentre mais en utilisant la distance définie plus tôt. Ainsi pour s histogrammes il est défini le barycentre avec l'équation 5

$$P(\lambda) = \arg \min_{p \in \sum_N} \sum_s \lambda_s W(p, p_s) \quad (5)$$

Ce barycentre permettra ainsi de pouvoir interpoler les images p_s qui sont vu comme des histogrammes. Le paramètre λ permet de pouvoir contrôler l'interpolation et associer un poids à chaque image, la somme des λ_s sur les s doit être égal à 1 tous comme la somme sur les s des p_s

De plus le problème décrit dans la formule 3 peut être reformuler de la forme de la formule 6 avec ξ et $KL()$ qui sont défini dans les définitions 7 et 8

$$\min_{M \in U(r, c)} KL(C, \xi) \quad (6)$$

$$\xi = \exp\left(-\frac{C}{\epsilon}\right) \quad (7)$$

$$KL(M, \xi) = \sum M_{ij} (\log\left(\frac{M_{ij}}{\xi_{ij}}\right) - 1) \quad (8)$$

Il se trouve que ce problème peut être résolu efficacement par l'algorithme de Sinkhorn, et cet algorithme peut être utilisé pour calculer le barycentre décrit dans l'équation 5. Le paramètre K désigne une convolution par une gaussienne.

$$P^{(l)}(\lambda) = \prod_s (K^T a_s^{(l)})^{(\lambda_s)} \quad (9)$$

$$b_s^{(l+1)} = \frac{P^{(l)}(\lambda)}{K a_s^{(l)}} \quad (10)$$

$$a_s^{(l+1)} = \frac{p_s}{Kb_s^{(l+1)}} \quad (11)$$

Cet algorithme sera donc utilisé pour pouvoir interpoler des images entre elles

3 Influence des paramètres

Il sera utilisé 3 images noirs et blancs (figure 2 qui seront interpolé suivant la méthode décrite dans la section 1.

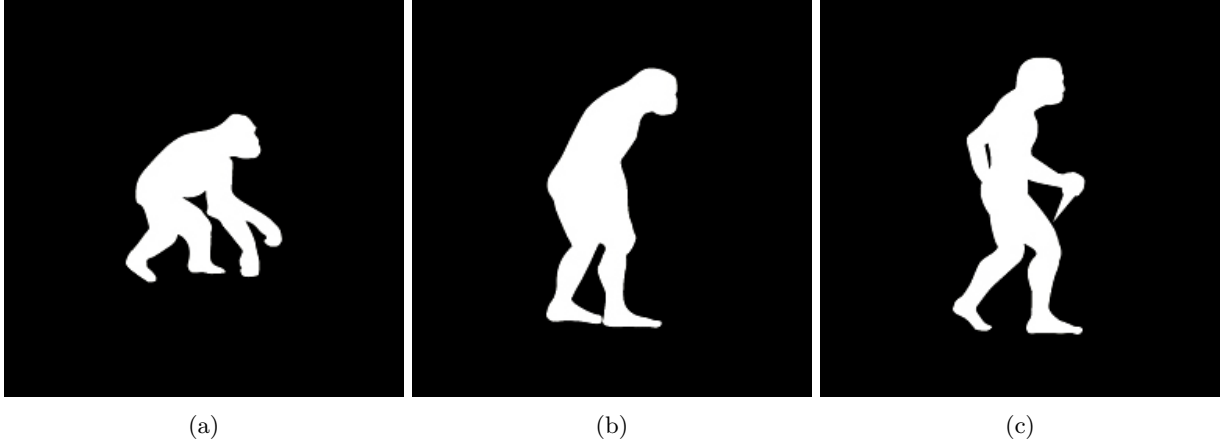
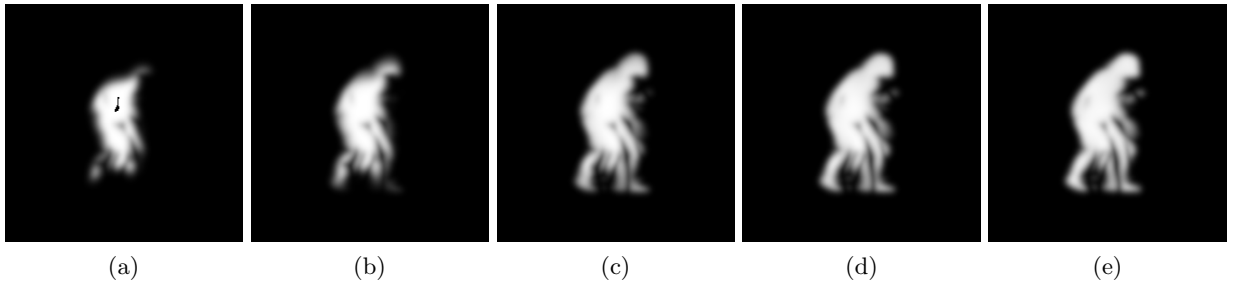


Figure 2: Affichage des 3 images d'origine

Le premier paramètre à observer est le paramètre l qui détermine le nombre d'itération de l'algorithme, en effet l'algorithme affirme que plus le nombre d'itérations est élevé plus le barycentre évalué tend vers le barycentre régularisé de Wasserstein qui est le résultat voulu. Il est possible d'observer le résultat de l'interpolation sur la figure 3

Figure 3: Affichage de l'évolution du barycentre calculé selon le nombre d'itération l avec $\lambda = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. Le paramètre l pour les images est de la gauche vers la droite 1,2,5,10 et 20

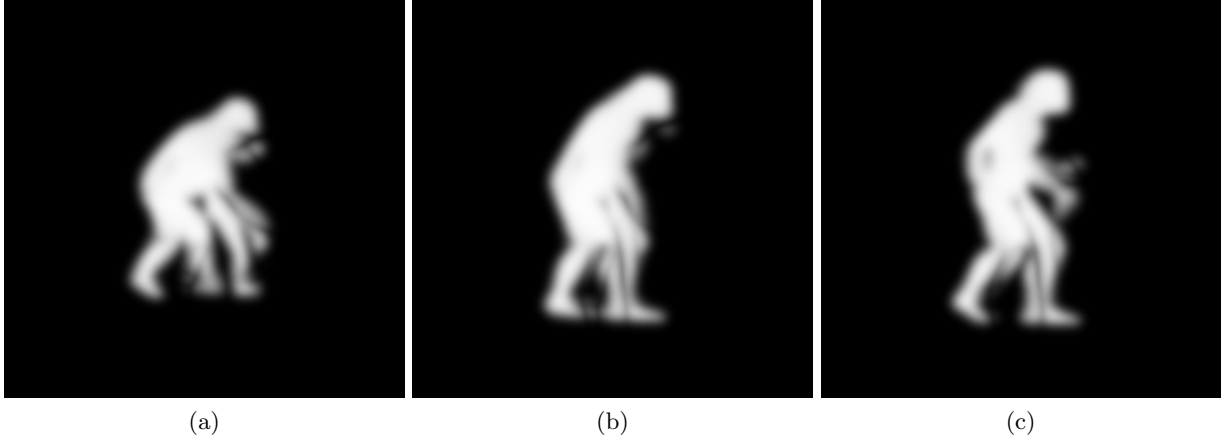


Cette figure illustre bien que l'algorithme semble converger vers une solution stable qui se trouve être le barycentre de Wasserstein. De plus visuellement on voit bien que le résultat a combiné les formes des 3 images de base. Le résultat se retrouve naturellement flou à cause des convolutions gaussiennes qui lissent l'image. On peut noter que ce lissage n'est pas toujours un défaut car on peut imaginer que dans une interpolation 3D en des objets le lissage ne réaliserai que de meilleurs résultats (sauf si les objets ne sont pas C^1 ou C^2 ...)

Pour la suite on supposera que tous les résultats sont obtenus avec un paramètre l assez élevé pour considérer que l'algorithme a convergé.

Le second paramètre à modifier est le paramètre λ qui permet de pondérer les images de base pour obtenir des résultats intermédiaires. En effet au lieu de donner des poids égaux à $\frac{1}{3}$ il est possible par exemple de pondérer une image à $\frac{2}{3}$ et les deux autres à $\frac{0.5}{3}$ ce qui permet d'obtenir la figure 4

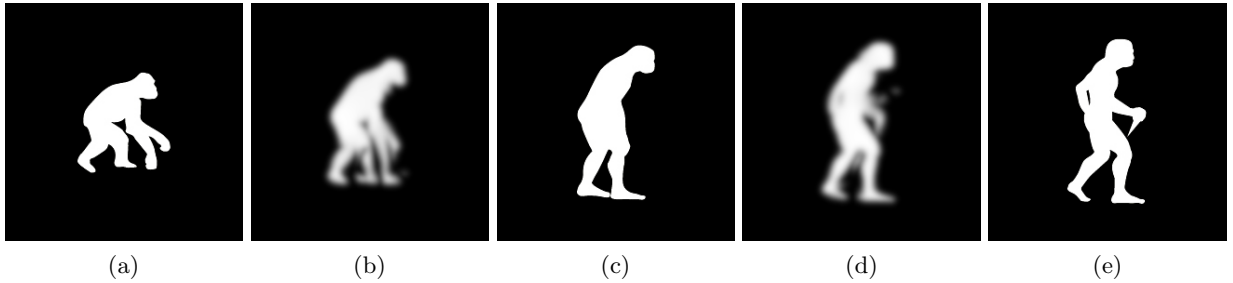
Figure 4: Affichage de l'évolution du barycentre calculé selon le paramètre $\lambda = (\frac{2}{3}, \frac{0.5}{3}, \frac{0.5}{3})$ avec le coefficient à $\frac{2}{3}$ appliqué de la gauche vers la droite.



Il est donc observable que par rapport à la figure 3 que 4 possède des résultats différents et que la prépondérance d'une des trois images impliquent que le résultat est plus proche de celle-ci.

De plus il est même possible de mettre un coefficient de λ à zéros pour qu'une des images soit totalement ignorée et qu'il y ait une interpolation entre 2 images seulement. Dans notre cas cela semble très cohérent avec nos images qui présentent une évolution de l'homme. La figure 5 présente l'utilisation de la méthode pour interpoler des formes intermédiaires entre les images d'origines de façon cohérente.

Figure 5: Affichage de l'évolution de l'homme avec une interpolation entre les images d'origine



De plus cette méthode qui par l'utilisation de transport optimal, possède la particularité de réaliser une interpolation en prenant en compte les formes des objets. En effet il est possible d'illustrer ce propos à l'aide de l'interpolation entre deux nouvelles images. Les figures 6 et 7 montrent l'interpolation de ces deux images en 5 étapes de pondérations qui sont: $\frac{1}{6}, \frac{2}{6}, \frac{3}{6}, \frac{4}{6}$ et $\frac{5}{6}$.

Figure 6: Affichage d'une des image d'origine à gauche et des 3 premières interpolations (3 figures à droite).

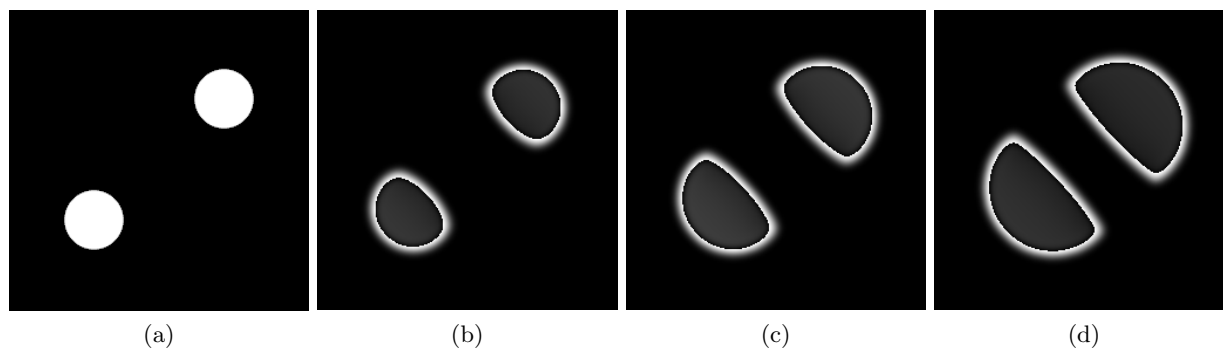
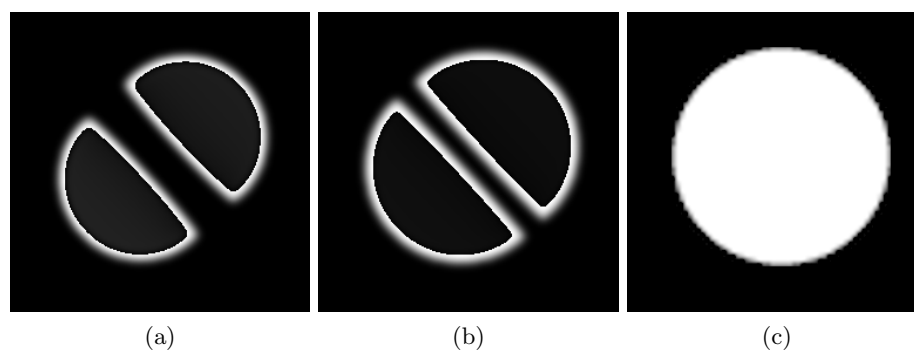


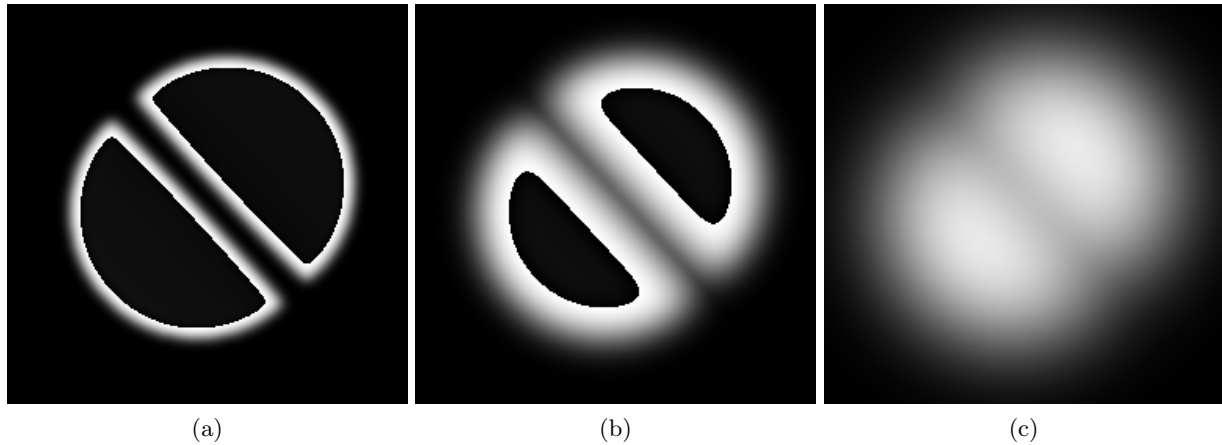
Figure 7: Affichage de la seconde image d'origine à droite et des 2 dernières interpolations (2 figure à gauche).



Cette interpolation illustre parfaitement de changement de forme et le transport optimal de matière entre les deux images.

Le dernier paramètre sur lequel on peut jouer est la largeur du filtre gaussien. Ce paramètre a donc pour conséquence d'augmenter le lissage et le flou du résultat mais celui-ci a un impact sur la convergence de l'algorithme. S'il est trop faible l'algorithme ne sembla pas donner de résultat, s'il est trop grand l'algorithme converge rapidement mais le résultat s'en retrouve lissé et moins précis sur la forme.

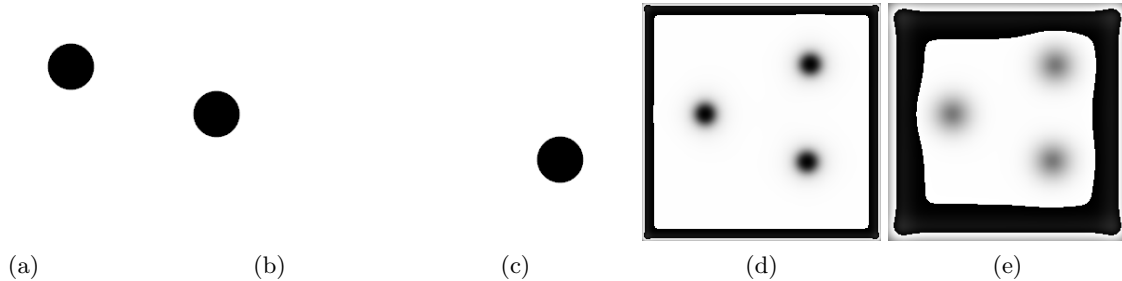
Figure 8: Comparaison du résultat obtenu avec 3 longueurs de filtre différentes qui sont de gauche à droite 25, 250 et 1600.



Une hypothèse sur ce phénomène serait que le filtre n'arrive pas à étaler la matière avec un filtre gaussien faible et donc n'arrive pas à converger.

De plus ce phénomène peut être utilisé pour boucher des trous dans une image. La figure 9 illustre l'influence du filtre gaussien sur sa capacité à remplir des trous.

Figure 9: Tentative de la méthode pour supprimer des trous dans les images. Les trois premières figures à gauche sont les images d'origine, la quatrième est l'interpolation avec le paramètre de la gaussienne à 25 et la dernière figure a le paramètre de la gaussienne à 225



On observe donc que la méthode arrive mieux à supprimer les trous avec un paramètre gaussien à 225. Il est cependant remarquable qu'il semble y avoir des défauts sur les bords des images avec un filtre gaussien trop grand ; cela est en fait dû au fait que le filtre ne gère pas complètement les bords de l'image et que mettre des pixels blancs sur ceux-ci crée ce problème.

4 Conclusion

L'utilisation du transport optimal amène une nouvelle façon de résoudre des problèmes. Durant ce TP il a pu être montré que le transport optimal est particulièrement adapté à l'interpolation des formes ce qui n'est pas possible avec des méthodes classiques qui ne font que d'interpoler l'amplitude des pixels alors qu'ici les pixels sont déplacés.

De plus l'utilisation du transport optimal peut être utilisée avec des formes 3D et possède des applications médicales ou encore dans le machine Learning.