Co-segmentation non-supervisée d'images utilisant les distances de Sinkhorn

Julien RABIN[1], Nicolas PAPADAKIS[2]

[1]Normandie Université, Université de Caen, GREYC, UMR 6072
ENSICAEN, 6, Bd du Maréchal Juin, 14050 CAEN

[2]CNRS, IMB, UMR 5251, Université de Bordeaux
351, cours de la Libération F-33405, Talence

julien.rabin@unicaen.fr, nicolas.papadakis@math.u-bordeaux.fr

Résumé – Nous proposons une formulation convexe et robuste du problème de co-segmentation de paire d'images non supervisé. Ce modèle définit l'adéquation statistique des régions segmentées dans le cadre du transport optimal, en mesurant le coût de transport entre les histogrammes de descripteurs (ici la couleur). Afin de réduire la complexité de mise en œuvre de ce modèle, les coûts de transport optimaux sont approchés par les distances de Sinkhorn, qui sont formulées comme la régularisation entropique du transport optimal. Un algorithme itératif exploitant la formulation primale-duale du problème est utilisé pour résoudre le problème de manière efficace et exacte.

Abstract – In this work, a convex and robust formulation of the unsupervised co-segmentation problem is introduced for pair of images. The proposed model relies on the optimal transport theory to asset the statistical similarity of the segmented regions' features (color histograms in this work). The optimal transport cost is approximated by Sinkhorn distance to reduce the optimization complexity. A primal-dual algorithm is used to solve the problem efficiently, without making use of sub-iterative routines.

1 Présentation du problème

Cet article a pour objet la segmentation d'image non supervisée, qui est sans doute l'un des problèmes les plus étudiés du domaine de la reconnaissance de formes. Plus particulièrement, nous traitons du problème de la *co-segmentation* (voir [8] pour une synthèse), où l'on cherche à segmenter simultanément plusieurs images qui contiennent le même objet d'intérêt.

Motivation Nous utilisons dans ce travail le modèle variationnel proposé par [7] qui repose sur la distance de Wasserstein (voir e.g. [2]) pour comparer les histogrammes couleurs des deux régions segmentées. Alors que les métriques usuelles (normes ℓ^1 ou ℓ^2) sont très sensibles au changement de quantification [8], celles reposant sur le transport optimal y sont naturellement très robustes. La formulation convexe de [7] autorise l'utilisation des outils de l'optimisation convexe (Generalized Forward Backward Splitting [6]) mais requiert le calcul extrêmement coûteux et itéré de l'opérateur proximal de la distance de Wasserstein, exprimé comme un programme quadratique.

Notre objectif est de généraliser ce modèle –dont les qualités ont été démontrés– à d'autres coûts de transport, tout en accélérant l'optimisation de ce problème. Pour cela, nous exploitons les travaux sur la régularisation entropique du transport optimal (appelée « distances de Sinkhorn ») [2], et en particulier sur les récents résultats de [3] sur la forme duale, en étendant le modèle de segmentation proposé dans [5].

Contributions Le modèle généralisé proposé permet : (i) de considérer d'autres mesures de similarité exploitant le transport optimal que la seule distance quadratique de Wasserstein ; (ii) de définir un modèle plus robuste aux changements de caractéristiques de l'objet entre les vues (illumination, balance des blancs, *etc*), grâce à la régularisation entropique ; (iii) d'être résolu par un algorithme itératif accéléré de type primal-dual, sans avoir recours à des sous-itérations.

2 Modèle étudié

Pour des raisons de simplicité, nous nous limiterons ici à la co-segmentation de deux images, notées I^1 et I^2, en utilisant des attributs de couleurs exclusivement. Une image $I^k : \Omega_k \mapsto \mathbb{R}^d$ est définie sur une grille discrète $\Omega_k \subset \mathbb{Z}^2$ de $N_k = |\Omega_k|$ pixels en couleurs ($d = 3$). On note $N = N_1 + N_2$ la taille du problème, et $\underline{N} = \min\{N_1, N_2\}$ la plus grande taille de région commune aux deux images.

Notations On considère indifféremment des espaces vectoriels de vecteurs ou de matrices où $\langle ., . \rangle$ indique le produit scalaire canonique. La transposition d'une matrice A est notée $A^\top$. On note : $\mathbf{0}_n, \mathbf{1}_n \in \mathbb{R}^n$ les vecteurs/matrices remplis de 0 et de 1 respectivement, Id l'opérateur identité, $\iota_{\mathcal{X}}(.)$ la fonction indicatrice de l'ensemble $\mathcal{X}$, $\delta_x(.) = \delta(. - x)$ la distribution de Dirac localisée en x, et pour finir ∇ l'opérateur de gradient.

Modèle non convexe La segmentation de I^k est définie de manière explicite en chaque pixel $x \in \Omega^k$ par une variable binaire $u^k(x) \in \{0,1\}$, indiquant que le pixel x dans l'image I^k appartient (si égal à 1) ou pas (si 0) à la région d'intérêt. On souhaite segmenter les deux images de manière conjointe, de telle sorte que les statistiques des deux régions soient proches, en considérant le modèle variationnel suivant [7] :

$$J(u) := S(\mu^1, \mu^2) + \sum_{k=1}^{2} \rho\, TV(u^k) - \delta \|u^k\|_0 \qquad (1)$$

où
- $\rho, \delta \geq 0$ sont des paramètres du modèle ;
- $\mu^k = \mu(u^k)$ est la distribution de probabilité discrète (*ddp*) de la région décrite par u^k dans l'image I^k :

$$\mu^k : y \in \mathbb{R}^d \mapsto \frac{\sum_{x\in\Omega_k} u^k(x)\delta_{I^k(x)}(y)}{\sum_{x\in\Omega_k} u^k(x)} \in [0,1]; \quad (2)$$

- S est une mesure de similarité entre *ddp* (divergence de Kullback, distance de Wasserstein, distance ℓ^1, *etc*) ;
- TV est la formulation discrète de la *variation totale*, qui pénalise le périmètre des lignes de niveaux et ainsi promeut les régions avec des contours réguliers ;
- $\|u^k\|_0 = \langle u^k,\, \mathbf{1}_{N_k}\rangle$ lorsque $u^k \in \{0,1\}^{N_k}$ est la taille de la région segmentée dans I^k ; c'est un terme de « gonflement » qui encourage la segmentation de larges régions.

Relaxation convexe La résolution de (1) est NP-difficile en raison de l'ensemble de contraintes non-convexes $\{0,1\}^{N_k}$ sur les variables de segmentation u^k qu'il est usuel d'approcher par son enveloppe convexe [5]. Les variables u^k sont alors « relaxées » dans l'intervalle $[0,1]^{N_k}$: ce sont des fonctions de pondération définissant de manière implicite la segmentation (obtenue par la suite à l'aide d'un seuillage fixé à $\frac{1}{2}$).

D'autre part, et de manière à obtenir une formulation convexe du problème considéré, il est nécessaire de s'affranchir de la normalisation des *ddp* μ^k (2), qui est non linéaire. Pour cela, on impose que les deux variables de segmentation u^k aient la même somme (domaine $\mathcal{S}$) et soient normalisées (domaine Δ) :

$$(u^1, u^2) \in \mathcal{S} \cap \Delta, \quad \text{avec} \qquad (3)$$
$$\begin{cases} \mathcal{S} := \left\{(x,y) \in \mathbb{R}_+^{N_1} \times \mathbb{R}_+^{N_2}, \langle x,\, \mathbf{1}_{N_1}\rangle = \langle y,\, \mathbf{1}_{N_2}\rangle\right\} \\ \Delta := \left\{(x,y) \in [0,1]^{N_1} \times [0,1]^{N_2}\right\} \end{cases}.$$

Histogrammes Afin de diminuer la complexité de la comparaison des distributions empiriques μ^k, il est également usuel de réduire leur dimension en utilisant des histogrammes h^k, généralement obtenus en combinant un algorithme de classification (K-moyennes, super pixels, *etc*) sur M_k *classes* et une quantification par affectation « au plus-proche voisin ». L'histogramme h^k de la région segmentée dans I^k dépend alors linéairement de u^k en fonction de la matrice d'affectation H_k

$$h^k := H_k u^k \in \mathbb{R}_+^{M_k} \quad \text{avec } H_k \in \mathbb{R}^{M_k \times N_k} \qquad (4)$$

où $H_k(i,x) = 1$ si le pixel $I^k(x)$ appartient à la classe d'indice $1 \leq i \leq M_k$, et 0 sinon. On note par la suite H la matrice $H := \begin{bmatrix} H_1 & \mathbf{0}_{M_1\times N_2} \\ \mathbf{0}_{M_2\times N_1} & H_2 \end{bmatrix} \in \mathbb{R}^{M\times N}$.

Variation totale Nous utilisons la définition isotrope usuelle de la variation totale discrète que l'on notera $TV(u) = \|Du\|_{1,2}$. L'opérateur de gradient discret $D_k u^k(x) = v^k = (v_1^k(x), v_2^k(x))$ est défini par des différences finies :

$$v_1(i,j) = u(i,j) - u(i-1,j), \quad v_2(i,j) = u(i,j) - u(i,j-1)$$

et la norme $\ell_{1,2}$ correspond à la somme des normes ℓ^2 sur Ω_k

$$\|v^k\|_{1,2} = \sum_{x\in\Omega_k} \sqrt{v_1^k(x)^2 + v_2^k(x)^2}.$$

On note D la matrice $D = \begin{bmatrix} D_1 & \mathbf{0}_{2N_1\times N_2} \\ \mathbf{0}_{2N_2\times N_1} & D_2 \end{bmatrix} \in \mathbb{R}^{2N\times N}$.

Modèle convexe étudié En supposant que la mesure de similarité S est convexe, le problème de cosegmentation peut se formuler par le problème d'optimisation convexe suivant

$$u^\star \in \underset{u\in\mathcal{S}\cap\Delta}{\operatorname{argmin}}\ S(Hu) + \rho\|Du\|_{1,2} - \delta\langle u,\, \mathbf{1}_N\rangle. \qquad (5)$$

Le choix de S est crucial, tant du point de vue des performances du modèle que de celui algorithmique. Nous étudions dans la section suivante comment définir une mesure de similarité robuste dans le cadre du transport optimal.

3 Distances de Sinkhorn

Nous considérons ici le problème discret du transport optimal de masse de Monge-Kantorovitch (voir *e.g.* [2]) entre deux distributions de masse h^1 et h^2. Il n'est pas nécessaire que les deux histogrammes aient la même dimension, ou qu'ils correspondent au même partitionnement de l'espace des descripteurs, mais il est néanmoins indispensable qu'ils aient la même masse totale. On note ainsi $\mathcal{H}$ l'ensemble suivant

$$\mathcal{H} := \left\{(h^1, h^2) \in \mathbb{R}_+^{M_1} \times \mathbb{R}_+^{M_2}, \langle h^1,\, \mathbf{1}_{M_1}\rangle = \langle h^2,\, \mathbf{1}_{M_2}\rangle\right\},$$

et le sous-ensemble $\mathcal{H}_m$ où la masse totale est bornée par m

$$\mathcal{H}_m := \left\{(h^1, h^2) \in \mathcal{H}, \langle h^1,\, \mathbf{1}_{M_1}\rangle = \langle h^2,\, \mathbf{1}_{M_2}\rangle \leq m\right\}.$$

Dans le contexte du problème étudié, on considérera pour m la taille en pixel de la plus petite des deux images, soit $m = \underline{N}$.

Transport optimal On considère une table de coût $C \in \mathbb{R}^{M_\times}$, où $C_{i,j}$ indique le coût de transport d'une unité de masse depuis la classe d'indice i de h^1 vers la classe d'indice j de h^2. On a donc $M_\times := M_1 \times M_2$ transferts possibles. Le problème du transport optimal consiste à définir un plan de transport $P \in \mathbb{R}_+^{M_\times}$ transférant la totalité de la masse de h^1 vers h^2 avec un coût minimal, où chaque entrée $P_{i,j}$ correspond à la quantité de masse puisée parmi $h^1(i)$ et envoyée vers $h^2(j)$. Le calcul du coût de transport optimal ainsi défini correspond au problème d'optimisation linéaire de Monge-Kantorovich

$$\forall\,(h^1, h^2) \in \mathcal{H}, \quad \mathbf{MK}(h^1, h^2) := \min_{P\in\mathcal{P}(h^1,h^2)} \langle P,\, C\rangle \qquad (6)$$

où $\mathcal{P}(.,.)$ est l'ensemble des plans de transport admissibles

$$\mathcal{P}(x,y) := \left\{P \in \mathbb{R}_+^{M_\times}, P\mathbf{1}_{M_2} = x \text{ et } P^\top \mathbf{1}_{M_1} = y\right\}. \qquad (7)$$

Lorsque $C_{i,j} = \|x_i - y_j\|^2$ (distance entre les classes), le coût de transport est appelé distance de Wasserstein quadratique.

Il est possible d'écrire (6) comme un programme linéaire sous forme primale / duale, en y intégrant la contrainte sur $\mathcal{H}$

$$\mathbf{MK}(h) = \min_{\substack{p\in\mathbb{R}_+^{M_\times} \\ L^\top p = h}} \langle p,\, c\rangle + \iota_{\mathcal{H}}(h) \;=\; \max_{\substack{\beta\in\mathbb{R}^M \\ L\beta\leq c}} \langle h,\, \beta\rangle, \quad (8)$$

où $h = \begin{bmatrix} h^1 \\ h^2 \end{bmatrix}$ est la concaténation des histogrammes, et les vecteurs p et c correspondent aux matrices bi-stochastique P et de coût C (obtenus par concaténation de leurs vecteurs colonnes). Les M contraintes linéaires sur p de la relation $L^\top p = h$ sont décrites à l'aide de la matrice $L^\top \in \mathbb{R}^{M\times M_\times}$ qui est définie de telle sorte que $(L\beta)_{i,j} = \beta^1_i + \beta^2_j$. On remarquera que dans la formulation duale (8), la contrainte $\iota_{\mathcal{H}}(h)$ est absorbée par la variable duale β qui est définie à une constante près.

Distance de Sinkhorn Nous intéressons à la régularisation entropique du transport optimal proposée par Cuturi dans [2]

$$\mathbf{MK}_\lambda(h) := \min_{p\in\mathbb{R}_+^{M_\times},\, L^\top p = h} \langle p,\, c + \tfrac{1}{\lambda}\log p\rangle + \iota_{\mathcal{H}}(h). \quad (9)$$

Cuturi montre que le coût de transport obtenu, appelé dorénavant « distance de Sinkhorn » en raison de l'algorithme utilisé, est plus rapide à optimiser et même plus robuste que la distance de Wasserstein en présence de données aberrantes. Dans [3], il est également montré que cette régularisation entropique est même nécessaire en pratique pour obtenir un plan de transport de transport stable vis-à-vis de la quantification.

Lorsque $\lambda \to \infty$, on retrouve la métrique **MK** (Eq. (8)), tandis que si $\lambda \to 0$ le plan de transport solution de (9) est séparable $P = \frac{1}{m} h^1 {h^2}^\top$, avec $m = \langle h^1,\, \mathbf{1}_{M_1}\rangle = \langle h^2,\, \mathbf{1}_{M_2}\rangle$. Pour notre problématique, il est intéressant de travailler sur l'ensemble restreint $\mathcal{H}_m$, ce qui motive la définition suivante

$$\underline{\mathbf{MK}}_{\lambda,m}(h) := \min_{p\in\mathbb{R}_+^{M_\times},\, L^\top p = h} \langle p,\, c + \tfrac{1}{\lambda}\log \tfrac{p}{m}\rangle + \iota_{\mathcal{H}_m}(h). \quad (10)$$

où la matrice p est normalisée par m, la masse totale maximale.

Formulation duale Rappelons que la conjuguée de Legendre-Fenchel $F^* : Y \mapsto \mathbb{R}$ d'une fonction convexe $F : X \mapsto \mathbb{R}$ s'écrit $F^*(y) = \sup_{x\in X}\langle x,\, y\rangle - F(x)$, et $F^{**} = F$. On a par exemple, d'après la formulation duale de (8)

$$\mathbf{MK}^*(\beta) = \iota_{L\beta\leq c}(\beta). \quad (11)$$

Dans [3], il est démontré que la transformée de Legendre de $\mathbf{MK}_\lambda$ s'écrit (en notant pour simplifier $\mathbf{1}$ au lieu de $\mathbf{1}_{M_\times}$)

$$\mathbf{MK}^*_\lambda(\beta) = \tfrac{1}{\lambda}\langle Q_\lambda(\beta), \mathbf{1}\rangle \quad \text{où } Q_\lambda(\beta) := e^{\lambda(L\beta - c) - \mathbf{1}} \quad (12)$$

et où la fonction exponentielle est appliquée à chaque élément des vecteurs/matrices. On peut également montrer que, dans le cas normalisé, on retrouve un résultat similaire [5] :

$$\underline{\mathbf{MK}}^*_{\lambda,m}(\beta) = \frac{m}{\lambda}\begin{cases} \langle Q_\lambda(\beta), \mathbf{1}\rangle & \text{si } \langle Q_\lambda(\beta), \mathbf{1}\rangle \leqslant 1 \\ \log\langle Q_\lambda(\beta), \mathbf{1}\rangle + 1 & \text{si } \langle Q_\lambda(\beta), \mathbf{1}\rangle \geqslant 1 \end{cases} \quad (13)$$

Ce résultat remarquable permet d'utiliser le transport optimal comme terme de fidélité dans des problèmes variationnels sans avoir à calculer explicitement le plan de transport.

Gradient Le gradient de la conjuguée de la distance de Sinkhorn normalisée s'écrit (en notant simplement Q pour $Q_\lambda(\beta)$)

$$\nabla\underline{\mathbf{MK}}^*_{\lambda,m}(\beta) = \begin{cases} m\ \left(Q\mathbf{1}_{M_2}, Q^\top\mathbf{1}_{M_1}\right) & \text{si } \langle Q, \mathbf{1}\rangle \leqslant 1 \\ \frac{m}{\langle Q, \mathbf{1}\rangle}\left(Q\mathbf{1}_{M_2}, Q^\top\mathbf{1}_{M_1}\right) & \text{si } \langle Q, \mathbf{1}\rangle \geqslant 1 \end{cases} \quad (14)$$

C'est une fonction Lipschitzienne de constante bornée supérieurement par $2\lambda m$. Nous allons exploiter ce résultat pour résoudre notre problème de co-segmentation (5).

4 Algorithme

Formulation Primale Dans [7], le problème (5) est résolu par un algorithme explicite-implicite (GFB [6]), qui requiert le calcul des opérateurs proximaux des différents termes (notés *Prox* dans la suite). Dans le cas où l'on sait calculer de manière explicite ces opérateurs, on dit alors que les fonctions associées sont *simples*. Toutefois, dans le cas général, comme ici pour la distance de Wasserstein, cela implique la résolution à chaque itération d'un problème d'optimisation convexe (programme quadratique dans [7]).

Formulation Primale-Duale Nous proposons ici de reformuler (5) comme un problème de point-selle qui peut être efficacement résolu par un algorithme implicite-explicite de type primal-dual [1]. Pour cela, nous exprimons certains des termes convexes en fonction de leur conjuguée de Legendre-Fenchel. Il est en outre nécessaire de passer par une séparation de variables pour les termes faisant intervenir un opérateur linéaire non inversible, comme par exemple la variation totale : $TV(u) = \sup_v\langle v,\, Du\rangle - \iota_{\|\cdot\|_{\infty,2}\leq 1}(v)$, où ℓ^∞ est la norme duale de ℓ^1.

Proposition 1. *En choisissant $S = \mathbf{MK}_\lambda$, le problème (5) est équivalent au suivant*

$$\min_u \max_v \langle Ku, v\rangle + \iota_\Delta(u) - \delta\langle u,\, \mathbf{1}_N\rangle - F^*(v) - G^*(v) \quad (15)$$

où
- $u = (u^1, u^2) \in \mathbb{R}^N$ est la variable primale,
- $v = (v^1, v^2, h^1, h^2) \in \mathbb{R}^{2N+M}$ est la variable duale ;
- $K^\top = [D^\top, H^\top] \in \mathbb{R}^{N\times(2N+M)}$ est une matrice ;
- ι_Δ est l'indicatrice de Δ, non-lisse mais *simple* ;
- $\langle u,\, -\delta\mathbf{1}_N\rangle$ est linéaire, de gradient constant $-\delta\mathbf{1}_N$;
- $G^*(v) = \sum_{k=1}^2 \iota_{\|\cdot\|_{\infty,2}\leq\rho}(v^k)$ est non-lisse mais *simple* ;
- $F^*(v) = \underline{\mathbf{MK}}^*_{\lambda,N}(h^1, h^2)$ (Eq. (13)) est la conjuguée de Legendre-Fenchel de $\underline{\mathbf{MK}}_{\lambda,N}$ (Eq. (10)), à gradient Lipschitz de constante $L_{F^*} \leq 2\lambda N$ (Eq. (14)).

Démonstration. Commençons par observer que la contrainte $u \in \Delta\cap\mathcal{S}$ peut être exprimée par la somme de deux indicatrices $\iota_{\Delta\cap\mathcal{S}} = \iota_\Delta + \iota_{\mathcal{S}}$, où ι_Δ apparaît au problème primal tandis que $\iota_{\mathcal{S}}$ est incorporé par la formulation duale. Ensuite, pour toute matrice $H \in \mathbb{R}_+^{M\times N}$ dont les colonnes somment à 1 (ce que vérifie la matrice d'affectation H définie en Eq. (4)), on a

$$\forall u \in \mathbb{R}^N,\ \iota_{\mathcal{S}\cap\Delta}(u) = \iota_\Delta(u) + \iota_{\mathcal{S}}(u) = \iota_\Delta(u) + \iota_{\mathcal{H}_N}(Hu)$$

car $\langle H_k u^k,\, \mathbf{1}_{M_k}\rangle = \langle u^k,\, \mathbf{1}_{N_k}\rangle \leq N_k\ \forall u \in \Delta$. On peut alors écrire $S(Hu)+\iota_{\Delta\cap S}(u) = \mathbf{MK}_\lambda(Hu)+\iota_\Delta(u)+\iota_{\mathcal{H}_N}(Hu) = \iota_\Delta(u) + \max_h \langle Hu,\, h\rangle - \mathbf{MK}^*_{\lambda,N}(h)$. En combinant ce résultat avec la séparation de variables pour les termes de variation totale, on obtient le résultat souhaité. □

Algorithme Primal-Dual Pour résoudre le problème (15), on utilise l'algorithme primal-dual accéléré proposé dans [4] (par simplicité, présenté ici avec un seul paramètre $r > 0$)

$$\begin{cases} u^{(t+1)} &= Proj_\Delta\left(u^{(t)} - T(K^\top v^{(t)} - \delta \mathbf{1}_N)\right) \\ \tilde{u}^{(t+1)} &= 2u^{(t+1)} - u^{(t)} \\ v^{(t+1)} &= Prox_{G^*}\left(v^{(t)} + \Sigma(K\tilde{u}^{(t)} - \nabla F^*(v^{(t)}))\right) \end{cases} \tag{16}$$

avec $T = \tau\,\mathrm{Id}_N$ et $\Sigma = \mathrm{diag}\,(\sigma_v \mathbf{1}_{2N}, \boldsymbol{\sigma}_h)$ deux matrices de préconditionnement diagonales (assurant la convergence)

$$\frac{1}{\tau} = 3r, \quad \frac{1}{\sigma_v} = \frac{2}{r}, \quad \text{et} \quad \frac{1}{\boldsymbol{\sigma}_h} = L_{F^*}\mathbf{1}_M + \frac{1}{r}H\mathbf{1}_N$$

où l'inversion $\frac{1}{\boldsymbol{\sigma}} = \left(\frac{1}{\sigma(\ell)}\right)_{1\leq \ell\leq M}$ est effectuée par élément.

Opérateurs de projection et proximal Le projecteur orthogonal sur l'hypercube Δ se calcule aisément

$$Proj_\Delta(u) = \min(\max(u, \mathbf{0}_N), \mathbf{1}_N).$$

Rappelons que l'opérateur proximal de Moreau pour une fonction F se définit ainsi : $Prox_F(x) = \mathrm{argmin}_y \frac{1}{2}\|x-y\|^2 + F(y)$. L'opérateur proximal de G^* s'exprime lui aussi aisément car G^* est une somme de fonctions séparables ; de plus, ces fonctions sont les indicatrices d'ensembles convexes dont les opérateurs proximaux sont les projecteurs. On a donc

$$Prox_{G^*}(v^1, v^2) = \left(Proj_{\|\cdot\|_{\infty,2}\leq\rho}(v^1), Proj_{\|\cdot\|_{\infty,2}\leq\rho}(v^2)\right)$$
$$\text{avec } Proj_{\|\cdot\|_{\infty,2}\leq\rho}(v^k) = \left(\frac{v^k(x)}{\max\{\|v^k(x)\|/\rho, 1\}}\right)_{x\in\Omega_k}.$$

5 Expériences et discussion

Nous reprenons le protocole expérimental utilisé dans [8, 7]. Il consiste à segmenter deux photographies d'un même objet placé sur deux fonds différents (figure 1). Les histogrammes couleurs sont obtenus après quantification linéaire sur $M_k = 16$ valeurs de chacun des canaux couleurs RGB. La matrice de coût C est la matrice de distance quadratique entre les couleurs. Les paramètres de notre modèle sont $\rho = 0.3$, et $\delta = 10^{-2}$, avec la distance de Sinkhorn définie par $\lambda = 10^3$. On utilise l'algorithme (16) avec $r = 1$ et 10^5 itérations. La figure 1 montre que l'on arrive à obtenir des résultats très similaires à [7], qui utilise une pré-segmentation en super pixels pour diminuer la dimension du problème.

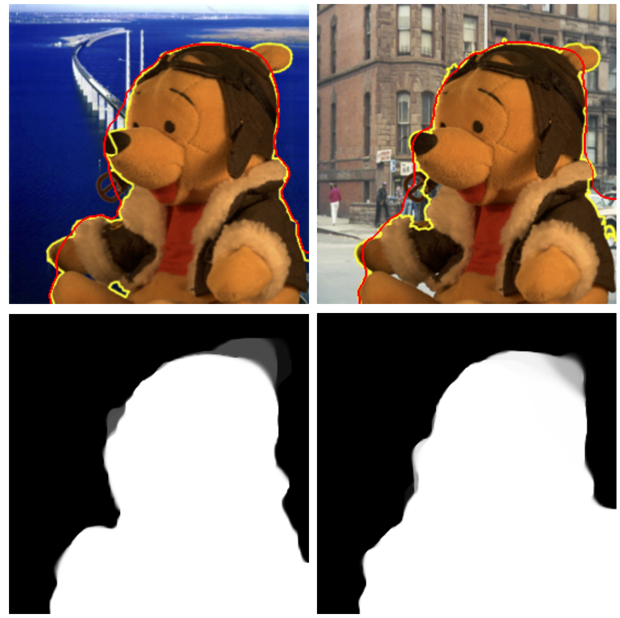

FIGURE 1 – Résultats obtenu par notre modèle (en rouge), qui permet d'approcher la méthode de [7] (en jaune). Les cartes u^k obtenues, quasiment binaires malgré la relaxation dans Δ, sont seuillées à $\frac{1}{2}$ pour définir la co-segmentation.

Discussion Notons tout d'abord qu'il est possible d'étendre ce travail pour des distances de transport non régularisés (*i.e.* $\lambda = \infty$), comme cela a été proposé pour la segmentation dans [5]. Une des limitations majeures du modèle étudié [7] est qu'il requiert une taille en pixels identique pour les deux régions segmentées, ce qui est notamment une restriction à l'invariance par changement d'échelle. Une autre généralisation intéressante de ce modèle est l'extension de la co-segmentation à plus de deux images.

Références

[1] P. Combettes, L. Condat, J.-C. Pesquet, and B. Vu. A forward-backward view of some primal-dual optimization methods in image recovery. In *IEEE ICIP*, pages 4141–4145, Oct 2014.

[2] M. Cuturi. Sinkhorn distances : Lightspeed computation of optimal transport. In *NIPS*, pages 2292–2300, 2013.

[3] M. Cuturi, G. Peyré, and A. Rolet. A smoothed dual approach for variational Wasserstein problems. *ArXiv* *1503.02533*, 2015.

[4] D. Lorenz and T. Pock. An inertial forward-backward algorithm for monotone inclusions. *JMIV*, 51(2) :311–325, 2015.

[5] J. Rabin and N. Papadakis. Convex color image segmentation with optimal transport distances. In *SSVM*, volume 9087, pages 256–269. 2015.

[6] H. Raguet, J. Fadili, and G. Peyré. A generalized forward-backward splitting. *SIIMS*, 6(3) :1199–1226, 2013.

[7] P. Swoboda and C. Schnörr. Variational image segmentation and cosegmentation with the Wasserstein distance. In *EMMCVPR*, pages 321–334, 2013.

[8] S. Vicente, V. Kolmogorov, and C. Rother. Cosegmentation revisited : Models and optimization. In *ECCV*, pages 465–479, 2010.