Compressive Sensing for Background Subtraction

Amine Hammami et Charles Dognin

ENSAE Paritech

Outline

The sparse representation

Soit X une image de taille $N_1 \times N_2$ et x un vecteur obtenu à partir de X de taille $N \times 1 (N = N_1 N_2)$.
Supposons qu'il existe $\Psi = [\psi_1, ..., \psi_N]$ une base fournissant la représentation K sparse de x:

$$x = \sum_{n=1}^{N} \theta(n)\psi_n = \sum_{l=1}^{K} \theta(n_l)\psi_{n_l} = \Psi\Theta$$

Une image est K sparse si $||\theta||_0 = K$ où l_0 compte le nombre d'éléments non nuls.

Random/Incoherent Projections

Il est impossible de mesurer les valeurs de K les plus larges.
La projection de x sur un ensemble de projecteurs linéaires $\Phi = [\phi'_1, ..., \phi'_M]'$ avec $(M < N)$ est elle mesurable:

$$y = \Phi x = \Phi \Psi \theta$$

La théorie du Compressed Sensing repose sur deux hypothèses:

- les deux bases Φ et Ψ doivent être incohérentes
- M supérieur à $\mathcal{O}(K \log(\frac{N}{K}))$

Signal Recovery via l_1 Optimization

l'optimisation l_1 peut être utilisée pour la reconstruction du signal:

$$\hat{\theta} = argmin||\theta||_1 \text{ sachant que } y = \Phi\Psi\theta$$

Dans ce papier, la méthode utilisée est Basis Pursuit Denoising:

$$\hat{\theta} = argmin||\theta||_1 + \frac{1}{2}\beta||y - \Phi\Psi\theta||_2^2$$

where $0 < \beta < \infty$

Notation

But: Retrouver la localisation, la forme et l'apparence d'un nouvel objet en utilisant des arrières plans d'entraînement.

- x_b Image arrière plan
- x_t Image test
- x_d Image soustraite

L'image soustraite est définie comme la soustraction entre l'image arrière plan et l'image test.
Le support de x_d est $\mathcal{S}_d = \{n | n = 1, ..., N; |x_d(n)| \neq 0\}$

Sparsity of Background Subtracted Images

La sparsité des images réelles est définie comme
$K_{\text{scene}} = K_b = K_t = (\lambda_0 \log N + \lambda_1)N$ où $(\lambda_0, \lambda_1) \in \mathbb{R}^2$.
La sparsité de x_d est $K_d = (\lambda_0 \log P + \lambda_1)P$ où $P = |\mathcal{S}_d|$.
Le nombre d'échantillon compressé:
$M_{\text{scene}} = M_b = M_t \approx K_{\text{scene}} \log(\frac{N}{K_{\text{scene}}}) < M_d \approx K_d \log(\frac{N}{K_d})$

The Background Constraint

Supposons un échantillon de mesures compressées $y_{bi}(M \times 1, i = 1, ..., B)$ extrait d'un ensemble d'images arrière plan x_{bi} qui suit une loi i.i.d $\mathcal{N}(y_b, \sigma^2 I)$.
Nous modélisons de plus $y_d = y_t - y_b \sim \mathcal{N}(\mu_d, \sigma^2 I)(M \times 1)$.
Nous introduisons la distance l_2 afin d'être robuste au bruit:

$$||y_{bi} - y_b||_2^2 = \sigma^2 \sum_{n=1}^{M} (\frac{y_{bi}(n) - y_b(n)}{\sigma})^2$$

Quand $M > 30$, l'application du théorème central limite donne que $||y_{bi} - y_b||_2^2 \sim \mathcal{N}(M\sigma^2, 2M\sigma^4)$.

En ayant une image test avec un nouvel objet à détecter, la distribution devient:
$||y_t - y_b||_2^2 \sim \mathcal{N}(M\sigma^2 + ||\mu_d||_2^2, 2M\sigma^4 + 4\sigma^2||\mu_d||_2^2)$.
Sachant que:

- $\frac{1}{M} << 1$
- Si $u << 1$ on a $1 + u \approx e^u$
- $\mathcal{N}(M\sigma^2, 2M\sigma^4) \approx M\sigma^2 \exp\{\sqrt{\frac{2}{M}}\mathcal{N}(0,1)\}$

nous avons finalement:

$$\log ||y_{bi} - y_b||_2^2 \sim \mathcal{N}(\mu_{bg}, \sigma_{bg}^2)$$

Object Detector Based on CS

Utilité d'un test pour voir la différence entre une image arrière plan et une image test.
Nous introduisons la distance l_2 entre y_t et y_b que nous approximons:

$$\log ||y_t - y_b||_2^2 \sim \mathcal{N}(\mu_t, \sigma_t^2)$$

Si $|\log ||y_t - y_b||_2^2 - \mu_{bg}| \geq c\sigma_{bg}$, alors un nouvel objet est détecté dans l'image test.

Foreground Reconstruction

Nous avons utilisé l'algorithme de Basis Pursuit Denoising:

$$\hat{\Theta} = argmin||\Theta||_1 + \frac{1}{2}\beta||y - \Phi\Psi\Theta||_2^2$$

où $0 < \beta < \infty$ et l'algorithme total variation optimization:

$$\text{minimiser}||\theta||_{\mathrm{TV}} = ||\nabla\theta||_1 \text{ tels que } ||\Psi\Theta - y||_2 \leq \epsilon$$

où $((\nabla\theta)_i)_j = \theta_{i+e_j} - \theta_i$

Adaptation of the Background Constraint

Deux types de changement:

- drifts comme les illuminations
- shifts des changements majeurs et brusques

- Les shifts sont très difficiles à gérer.
- Créer un algorithme robuste aux drifts $\Rightarrow$ y_b doit être continuellement mis à jour et doit faire la distinction entre un changement dans l'arrière plan et un nouvel objet à détecter:

$$y_{ma}^{\{j+1\}} = \gamma y_t^{\{j\}} + (1-\gamma) y_{ma}^{\{j\}}$$
$$y_b^{\{j+1\}} = \alpha (y_t^{\{j\}} - \hat{y}_e^{\{j\}}) + (1-\alpha) y_b^{\{j+1\}}$$

Background substraction with an SPC

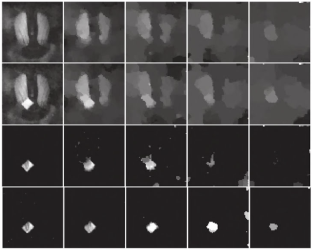

Figure 1: First Experiment's results

L'innovation est reconstruite avec l'algorithme même avec un taux inférieur à 1%

Adaptation to Illumination Changes

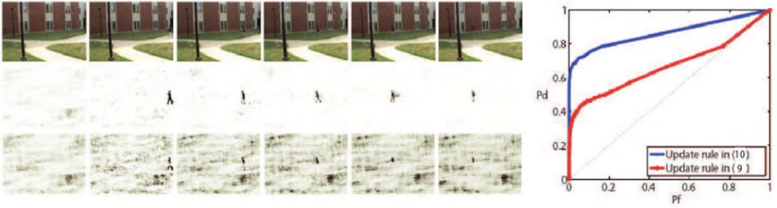

Figure 2: Adaptation to Illumination Changes

La deuxième règle de mise à jour performe mieux que la première.

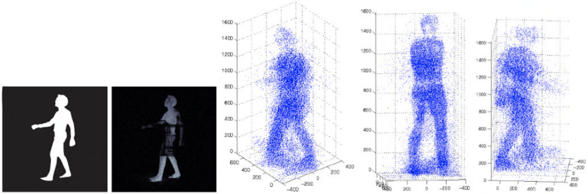

Figure 3: Silhouettes vs. Difference Images

Ils utilisent une configuration pour une reconstruction 3D en utilisant des mesures de compression.
Les figures de gauche représentent de gauche à droite la vérité de terrain et l'image 3D soustraite en arrière-plan en utilisant CS $\Rightarrow$ résultat pas très satisfaisant:ça intègre des éléments de l'arrière-plan.